

ANALYSIS II

Dietmar A. Salamon
ETH Zürich

2. April 2019

Preface

Bei diesen Notizen handelt es sich um ein Manuskript für die Vorlesung “*Analysis II*” für Studierende des zweiten Semester auf den Gebieten der Mathematik, der Physik, und der Ingenieurwissenschaften an der ETH Zürich. Vorausgesetzt werden die Inhalte der Vorlesungen “*Analysis I*” und “*Lineare Algebra I*” des ersten Semesters.

Das Manuskript wurde erstellt für die Analysis II Vorlesung am D-ITET im FS 2017. In dieser Vorlesung wurden im Kapitel 1 nur endlichdimensionale Banachräume und Matrixnormen behandelt, die Resultate von Kapitel 4 wurden ohne Beweise zusammengefasst, und Kapitel 7 über Differentialformen und den Satz von Stokes wurde nicht behandelt.

Der Anhang enthält sowohl Material, das aus anderen Vorlesungen als bekannt vorausgesetzt wird, als auch Material, das zum Inhalt dieser Vorlesung gehört. Zum letzteren gehören der Anhang A über den Abschluss und das Innere von Teilmengen eines metrischen Raumes und der Anhang B über zusammenhängende metrische Räume, während der Anhang D über die Determinante aus der *Linearen Algebra*, und Teile des Anhangs C über kompakte metrische Räume sowie der Anhang E über den Banachschen Fixpunktsatz aus der *Analysis I* vorausgesetzt werden. (Der Satz von Arzelà-Ascoli im Anhang C wurde in der Vorlesung weder behandelt, noch wird er im Text verwendet.)

4. Mai 2017

Dietmar A. Salamon

Inhaltsverzeichnis

1	Banachräume und Lineare Operatoren	1
1.1	Äquivalente Normen	1
1.2	Die Operatornorm	8
1.3	Banachalgebren	14
1.4	Die geometrische Reihe	21
1.5	Die Exponentialabbildung	23
1.6	Komposition von Potenzreihen	27
2	Differenzierbare Abbildungen	33
2.1	Grundbegriffe	33
2.2	Stetige Differenzierbarkeit	45
2.3	Die Kettenregel	49
2.4	Der Schrankensatz	60
2.5	Höhere Ableitungen	67
2.6	Die Taylorreihe	76
2.7	Lokale Extrema	87
3	Implizite Funktionen	99
3.1	Der Satz von der Umkehrabbildung	99
3.2	Der Satz von der impliziten Funktion	109
3.3	Untermannigfaltigkeiten	116
3.4	Extrema mit Nebenbedingungen	128
4	Vektorfelder und Flüsse	133
4.1	Existenz und Eindeutigkeit	133
4.2	Der Fluss eines Vektorfeldes	142
4.3	Stetigkeit des Flusses	145
4.4	Differenzierbarkeit des Flusses	149

5	Mehrfache Integrale	157
5.1	Das Riemannsches Integral	157
5.2	Jordansche Nullmengen und Stetigkeit	165
5.3	Riemannsches Summen	171
5.4	Iterierte Integrale	180
5.5	Jordan-messbare Mengen	185
5.6	Mehr über Jordan-messbare Mengen	189
5.7	Die Transformationsformel	199
5.8	Uneigentliche Integrale	218
5.9	Divergenz	224
6	Integration über Untermannigfaltigkeiten	229
6.1	Das d-dimensionale Volumen	229
6.2	Definition des Integrals	240
6.3	Beispiele	250
6.4	Der Satz von Gauß	260
6.5	Weitere Beispiele	271
7	Integration von Differentialformen	277
7.1	Alternierende Formen	277
7.2	Differentialformen	283
7.3	Orientierung	292
7.4	Integration	298
7.5	Der Satz von Stokes	303
7.6	Beispiele und Anwendungen	309
A	Abschluss und Inneres	317
A.1	Topologische Grundbegriffe	317
A.2	Die Begriffe von Abschluss und Innerem	319
B	Zusammenhängende Räume	323
B.1	Die Relativtopologie	323
B.2	Der Zusammenhangsbegriff	325
B.3	Weg-zusammenhängende Mengen	328
B.4	Zusammenhangskomponenten	330
B.5	Beispiele	331

C Kompakte metrische Räume	333
C.1 Der Kompaktheitsbegriff	333
C.2 Der Satz von Arzelà–Ascoli	337
D Die Determinante	343
D.1 Gruppen	343
D.2 Vektorräume	347
D.3 Die Determinante	350
D.4 Die allgemeine lineare Gruppe	360
E Der Banachsche Fixpunktsatz	365
Literatur	367
Index	368

Kapitel 1

Banachräume und Lineare Operatoren

1.1 Äquivalente Normen

Dieser erste Abschnitt ruft einige Definitionen und Resultate in Erinnerung, die aus der Analysis I Vorlesung bekannt sind. Sei X ein reeller Vektorraum. Eine **Norm** auf X ist eine Abbildung $X \rightarrow \mathbb{R} : x \mapsto \|x\|$ mit folgenden Eigenschaften.

(i) Für alle $x \in X$ gilt $\|x\| \geq 0$ und $\|x\| = 0 \iff x = 0$.

(ii) Für alle $x \in X$ und alle $\lambda \in \mathbb{R}$ gilt $\|\lambda x\| = |\lambda| \|x\|$.

(iii) Für alle $x, y \in X$ gilt $\|x + y\| \leq \|x\| + \|y\|$.

Ein **normierter Vektorraum** ist ein Paar $(X, \|\cdot\|)$ welches aus einem reellen Vektorraum X und einer Normfunktion $\|\cdot\|$ auf X besteht. Eine wichtige Klasse normierter Vektorräume sind solche bei denen die Norm von einem inneren Produkt erzeugt wird. Ein **inneres Produkt** auf X ist eine symmetrische bilineare Abbildung $X \times X \rightarrow \mathbb{R} : (x, y) \mapsto \langle x, y \rangle$, welche die Bedingung $\langle x, x \rangle > 0$ für alle $x \in X \setminus \{0\}$ erfüllt. Ist solch ein inneres Produkt gegeben, so definiert die Formel

$$\|x\| := \sqrt{\langle x, x \rangle} \quad \text{für } x \in X \quad (1.1.1)$$

eine Norm und es gilt die Cauchy-Schwarz-Ungleichung $|\langle x, y \rangle| \leq \|x\| \|y\|$ für alle $x, y \in X$.

Sei $(X, \|\cdot\|)$ ein normierter reeller Vektorraum. Dann definiert die Formel

$$d(x, y) := \|x - y\| \quad \text{für } x, y \in X \quad (1.1.2)$$

eine Abstandsfunktion $d : X \times X \rightarrow \mathbb{R}$, so dass X damit zu einem metrischen Raum wird. Der normierte Vektorraum $(X, \|\cdot\|)$ heißt **Banachraum** wenn der zugehörige metrische Raum (X, d) **vollständig** ist, das heißt, wenn jede Cauchy-Folge in X konvergiert. (Es sei daran erinnert, dass eine Folge $(x_n)_{n \in \mathbb{N}}$ in X eine **Cauchy-Folge** ist, wenn es zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ gibt, so dass für alle $n, m \in \mathbb{N}$ gilt: $n, m \geq n_0 \implies d(x_n, x_m) < \varepsilon$.) Ein **(reeller) Hilbertraum** ist ein Paar $(X, \langle \cdot, \cdot \rangle)$ welches aus einem reellen Vektorraum X und einem inneren Product $\langle \cdot, \cdot \rangle$ auf X besteht, so dass der zugehörige metrische Raum (X, d) mit der durch (1.1.1) und (1.1.2) definierten Abstandsfunktion vollständig ist. Also ist jeder Hilbertraum auch ein Banachraum mit der durch (1.1.1) definierten Normfunktion.

Es sei nun $(X, \|\cdot\|)$ ein normierter Vektorraum. Ist $x \in X$ und $r > 0$, so wird die Menge

$$B_r(x) := B_r(x; \|\cdot\|) := B_r(x; X, \|\cdot\|) := \{y \in X \mid \|x - y\| < r\}$$

der **(offene) Ball vom Radius r mit Mittelpunkt x** genannt. Eine Teilmenge $U \subset X$ heißt **offen**, wenn es für jedes Element $x \in U$ ein $\varepsilon > 0$ gibt so dass $B_\varepsilon(x) \subset U$ ist. Die Gesamtheit der offenen Mengen wird mit

$$\mathcal{U}(X, \|\cdot\|) := \{U \subset X \mid U \text{ ist offen bezüglich } \|\cdot\|\} \quad (1.1.3)$$

bezeichnet.

Beispiel 1.1.1. Sei \mathbb{R}^n der euklidische Raum aller Vektoren $x = (x_1, \dots, x_n)$ mit $x_i \in \mathbb{R}$. Für $1 \leq p < \infty$ definiert die Formel

$$\|x\|_p := \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}, \quad x \in \mathbb{R}^n, \quad (1.1.4)$$

eine Norm auf dem \mathbb{R}^n . Die Dreiecksungleichung für diese Norm heißt *Minkowski-Ungleichung*. Im Fall $p > 1$ verwendet ihr Beweis die *Hölder-Ungleichung* $\langle x, y \rangle = \sum_{i=1}^n x_i y_i \leq \|x\|_p \|y\|_q$ für $x, y \in \mathbb{R}^n$ und $1/p + 1/q = 1$. Für $p = \infty$ definiert die Formel

$$\|x\|_\infty := \max\{|x_1|, \dots, |x_n|\}, \quad x \in \mathbb{R}^n, \quad (1.1.5)$$

eine Norm auf \mathbb{R}^n .

Beispiel 1.1.2. Eine symmetrische Matrix $A = A^T \in \mathbb{R}^{n \times n}$, mit Koeffizienten $a_{ij} = a_{ji}$ für $i, j = 1, \dots, n$, heißt **positiv definit** wenn sie die Bedingung $x^T Ax = \sum_{i,j=1}^n x_i a_{ij} x_j > 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$ erfüllt. Jede positiv definite $(n \times n)$ -Matrix A definiert ein inneres Produkt

$$\langle x, y \rangle_A := x^T A y = \sum_{i,j=1}^n x_i a_{ij} y_j$$

auf dem \mathbb{R}^n . Die zugehörige Norm eines Vektors $x \in \mathbb{R}^n$ ist durch die Formel $\|x\|_A := \sqrt{x^T A x}$ gegeben.

Beispiel 1.1.3. Sei (M, d) ein metrischer Raum. Dann ist der Raum

$$\text{BC}(M) := \{f : M \rightarrow \mathbb{R} \mid f \text{ ist stetig und beschränkt}\}$$

mit der durch

$$\|f\|_\infty := \sup_{p \in M} |f(p)| \quad \text{für } f \in \text{BC}(M) \quad (1.1.6)$$

definierten Norm ein Banachraum. Die Vollständigkeit folgt aus der Tatsache, dass der Grenzwert einer gleichmäßig konvergierenden Folge stetiger Funktionen selbst wieder stetig ist, und dass gleichmäßige Konvergenz äquivalent ist zur Konvergenz bezüglich der Supremumsnorm. Ist (M, d) kompakt (das heißt jede Folge in M besitzt eine konvergente Teilfolge) so ist jede stetige Funktion $f : M \rightarrow \mathbb{R}$ notwendigerweise beschränkt. In diesem Fall stimmt also der Raum $\text{BC}(M)$ mit dem Raum $C(M)$ aller stetigen reellwertigen Funktionen auf M überein, und der Raum $C(M)$ ist mit der Supremumsnorm (1.1.6) ein Banachraum.

Beispiel 1.1.4. Sei $I = [a, b]$ ein kompaktes Intervall reeller Zahlen (mit $a < b$). Für $1 \leq p < \infty$ definiert die Formel

$$\|f\|_p := \left(\int_a^b |f(t)|^p dt \right)^{1/p} \quad \text{für } f \in C(I) \quad (1.1.7)$$

eine Norm auf dem Raum $C(I)$ aller stetigen reellwertigen Funktionen auf I . Der Raum $C(I)$ ist mit der Norm (1.1.7) nicht vollständig (Übung) und ist daher kein Banachraum. Im Fall $p = 2$ wird die Norm (1.1.7) durch das innere Produkt

$$\langle f, g \rangle := \int_a^b f(t)g(t) dt \quad \text{für } f, g \in C(I) \quad (1.1.8)$$

erzeugt.

Beispiel 1.1.5. Sei $I = [a, b]$ ein kompaktes Intervall reeller Zahlen (mit $a < b$) und sei $\ell \geq 0$ eine nichtnegative ganze Zahl. Dann ist der reelle Vektorraum

$$C^\ell(I) := \{f : I \rightarrow \mathbb{R} \mid f \text{ ist } \ell \text{ mal stetig differenzierbar}\}$$

mir der durch

$$\|f\|_{C^\ell} := \max_{k=0, \dots, \ell} \sup_{a \leq t \leq b} |f^{(k)}(t)|$$

für $f \in C^\ell(I)$ definierten Norm ein Banachraum. Für $\ell = 0$ ist dies der Raum $C(I)$ der stetigen Funktionen $f : I \rightarrow \mathbb{R}$ mit der Supremumsnorm wie in Beispiel 1.1.3.

Definition 1.1.6. Seien $\|\cdot\|$ und $\|\cdot\|'$ zwei Normfunktionen auf einem reellen Vektorraum X . Die Normfunktion $\|\cdot\|$ heißt **äquivalent** zu $\|\cdot\|'$ (Schreibweise $\|\cdot\| \sim \|\cdot\|'$) wenn eine Konstante $c > 0$ existiert, so dass

$$\frac{1}{c} \|x\| \leq \|x\|' \leq c \|x\| \quad \text{für alle } x \in X. \quad (1.1.9)$$

Bemerkung 1.1.7. (i) Sei X ein reeller Vektorraum. Dann definiert die “Äquivalenz von Normen” eine Äquivalenzrelation auf der Menge aller Normen auf X . Das heißt, für alle Normen $\|\cdot\|, \|\cdot\|', \|\cdot\|''$ auf X gilt

- $\|\cdot\| \sim \|\cdot\|,$
- $\|\cdot\| \sim \|\cdot\|' \iff \|\cdot\|' \sim \|\cdot\|,$
- $\|\cdot\| \sim \|\cdot\|', \|\cdot\|' \sim \|\cdot\|'' \implies \|\cdot\| \sim \|\cdot\|''.$

(ii) Seien $\|\cdot\|$ und $\|\cdot\|'$ zwei äquivalente Normen auf einem reellen Vektorraum X . Dann erzeugen die beiden Normen dieselben offenen Mengen. (Übung; siehe auch Beispiel 1.2.4).

(iii) Seien $\|\cdot\|$ und $\|\cdot\|'$ zwei äquivalente Normen auf einem reellen Vektorraum X . Da die offenen Teilmengen von X bezüglich der beiden Normen dieselben sind, folgt, dass alle *topologischen Aussagen in X* (also Aussagen die sich mit Hilfe offener Mengen formulieren lassen, zum Beispiel Konvergenz, Stetigkeit, Abgeschlossenheit, Kompaktheit) nicht davon abhängen, welche der beiden äquivalenten Normen wir dazu verwenden. Ist zum Beispiel $A \subset X$ kompakt (beziehungsweise abgeschlossen) bezüglich der Norm $\|\cdot\|$ so ist A auch kompakt (beziehungsweise abgeschlossen) bezüglich der Norm $\|\cdot\|'$. Weiterhin gilt, dass eine Folge $(x_k)_{k \in \mathbb{N}}$ in X genau dann bezüglich der Norm $\|\cdot\|$ eine Cauchy-Folge ist, beziehungsweise konvergiert, wenn sie bezüglich der Norm $\|\cdot\|'$ eine Cauchy-Folge ist, beziehungsweise konvergiert. Daher ist $(X, \|\cdot\|)$ genau dann ein Banachraum wenn $(X, \|\cdot\|')$ ein Banachraum ist.

Sei $I = [a, b] \subset \mathbb{R}$ ein kompaktes Intervall mit $a < b$. Dann ist die Menge $U := \{f \in C(I) \mid \sup_I |f| < 1\}$ offen bezüglich der Supremumsnorm $\|\cdot\|_\infty$ in Beispiel 1.1.3 aber nicht bezüglich der Norm $\|\cdot\|_p$ in Beispiel 1.1.4 für $1 \leq p < \infty$. Also sind die Normen $\|\cdot\|_\infty$ und $\|\cdot\|_p$ auf dem Raum $C(I)$ nicht äquivalent. Dies hängt damit zusammen, dass der Vektorraum der stetigen Funktionen auf einem nichttrivialen kompakten Intervall unendlichdimensional ist. Der nächste Satz zeigt, dass auf endlichdimensionalen Vektorräumen je zwei Normen äquivalent sind.

Satz 1.1.8. *Sei X ein endlichdimensionaler reeller Vektorraum. Dann sind je zwei Normen auf X äquivalent.*

Beweis. Der Beweis hat vier Schritte. Sei $n := \dim X \in \mathbb{N}$. Wir bezeichnen mit $\|\cdot\|_2$ die Euklidische Norm auf dem \mathbb{R}^n in (1.1.4) mit $p = 2$.

Schritt 1. *Sei $\|\cdot\|$ eine beliebige Norm auf dem \mathbb{R}^n . Dann existiert eine Konstante $c > 0$ so dass jedes $x \in \mathbb{R}^n$ die Ungleichung $\|x\| \leq c \|x\|_2$ erfüllt.*

Die Vektoren $e_i := (0, \dots, 0, 1, 0, \dots, 0)$ (mit der 1 an der i ten Stelle) für $i = 1, \dots, n$ bilden die Standardbasis des \mathbb{R}^n . Sei $c := (\sum_{i=1}^n \|e_i\|^2)^{1/2} > 0$. Dann gilt $x = \sum_{i=1}^n x_i e_i$ für jeden Vektor $x \in \mathbb{R}^n$ und daher

$$\|x\| \leq \sum_{i=1}^n |x_i| \|e_i\| \leq \sqrt{\sum_{i=1}^n |x_i|^2} \sqrt{\sum_{i=1}^n \|e_i\|^2} = c \|x\|_2.$$

Hier folgt die erste Ungleichung aus der Dreiecksungleichung für $\|\cdot\|$ und die zweite aus der Cauchy–Schwarz–Ungleichung für die Euklidische Norm.

Schritt 2. *Sei $\|\cdot\|$ wie in Schritt 1. Dann existiert eine Konstante $\delta > 0$ so dass jedes $x \in \mathbb{R}^n$ die Ungleichung $\delta \|x\|_2 \leq \|x\|$ erfüllt.*

Die Einheitssphäre $S^{n-1} := \{x \in \mathbb{R}^n \mid \|x\|_2 = 1\}$ ist mit der Abstandsfunktion $d(x, y) := \|x - y\|_2$ für $x, y \in S^{n-1}$ ein kompakter metrischer Raum nach dem Satz von Heine–Borel. Wir definieren die Funktion $f : S^{n-1} \rightarrow \mathbb{R}$ durch $f(x) := \|x\|$ für $x \in S^{n-1}$. Nach Schritt 1 gilt $|f(x) - f(y)| = |\|x\| - \|y\|| \leq \|x - y\| \leq c \|x - y\|_2$ für alle $x, y \in S^{n-1}$ und daher ist f stetig. Da S^{n-1} kompakt ist, existiert ein $x_0 \in S^{n-1}$ mit $\inf_{S^{n-1}} f = f(x_0) =: \delta > 0$. Für $x \in \mathbb{R}^n \setminus \{0\}$ gilt daher

$$\delta \leq f(\|x\|_2^{-1} x) = \|\|x\|_2^{-1} x\| = \|x\|_2^{-1} \|x\|.$$

Hieraus folgt die Behauptung von Schritt 2.

Schritt 3. *Je zwei Normen auf dem \mathbb{R}^n sind äquivalent.*

Nach Schritt 1 und Schritt 2 ist jede Norm auf dem \mathbb{R}^n äquivalent zur Euklidischen Norm. Daher folgt Schritt 3 aus Teil (i) von Bemerkung 1.1.7.

Schritt 4. *Je zwei Normen auf X sind äquivalent.*

Wähle eine Basis e_1, \dots, e_n von X . Dann ist die Abbildung

$$\Phi : \mathbb{R}^n \rightarrow X, \quad \Phi\xi := \sum_{i=1}^n \xi_i e_i \quad \text{für } \xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n,$$

linear und bijektiv, und ist daher ein Vektorraum-Isomorphismus. Sind $\|\cdot\|$ und $\|\cdot\|'$ zwei Normen auf X , so definieren die Formeln $\|\xi\|_\Phi := \|\Phi\xi\|$ und $\|\xi\|'_\Phi := \|\Phi\xi\|'$ für $\xi \in \mathbb{R}^n$ zwei Normen auf dem \mathbb{R}^n . Diese sind nach Schritt 3 äquivalent und daher sind auch die Normen $\|\cdot\|$ und $\|\cdot\|'$ auf X äquivalent. Damit ist Satz 1.1.8 bewiesen. \square

Satz 1.1.8 hat einige wichtige Konsequenzen.

Korollar 1.1.9. *Jeder endlichdimensionaler normierter Vektorraum ist ein Banachraum.*

Beweis. In der Analysis I Vorlesung wurde gezeigt, dass der \mathbb{R}^n mit der Euklidischen Norm ein Banachraum ist. Nach Satz 1.1.8 ist daher der \mathbb{R}^n mit jeder beliebigen Norm ein Banachraum. Da jeder endlichdimensionale normierte Vektorraum zum \mathbb{R}^n isomorph ist für ein $n \in \mathbb{N}_0$, folgt daraus die Behauptung. \square

Korollar 1.1.10. *Eine Teilmenge eines endlichdimensionalen normierten Vektorraumes ist genau dann kompakt wenn sie abgeschlossen und beschränkt ist.*

Beweis. Nach dem Satz von Heine–Borel gilt dies für den \mathbb{R}^n mit der Euklidischen Norm. Nach Satz 1.1.8 sind aber je zwei Normen auf dem \mathbb{R}^n äquivalent und daher sind alle drei Eigenschaften (Kompaktheit, Abgeschlossenheit, Beschränktheit) einer Teilmenge des \mathbb{R}^n unabhängig von der Wahl der Norm. Also gilt die Behauptung für jede beliebige Norm auf dem \mathbb{R}^n . Da jeder endlichdimensionale reelle Vektorraum zum \mathbb{R}^n für ein $n \in \mathbb{N}_0$ isomorph ist, folgt daraus die Behauptung im allgemeinen. \square

Nach Korollar 1.1.10 ist die Einheitskugel $S := \{x \in X \mid \|x\| = 1\}$ in jedem endlichdimensionalen normierten Vektorraum kompakt. Die folgende Übung zeigt, dass diese Eigenschaft genau die endlichdimensionalen normierten Vektorräume charakterisiert.

Übung 1.1.11. Sei $(X, \|\cdot\|)$ ein normierter reeller Vektorraum, so dass die Einheitskugel $S := \{x \in X \mid \|x\| = 1\}$ kompakt ist. Dann ist X endlichdimensional. **Hinweis:** Beweisen Sie folgendes.

(i) Jeder endlichdimensionale lineare Unterraum von X ist eine abgeschlossene Teilmenge von X .

(ii) Sei $Y \subset X$ eine abgeschlossene Teilmenge. Dann gilt für jedes $x \in X \setminus Y$ die Ungleichung $d(x, Y) := \inf_{y \in Y} \|x - y\| > 0$.

(iii) Sei $Y \subsetneq X$ ein abgeschlossener linearer Unterraum von X der nicht gleich dem ganzen Raum X ist. Dann existiert ein Vektor $x \in X$ so dass $\|x\| = 1$ und $d(x, Y) \geq 1/2$ ist. (Sei $x_0 \in X \setminus Y$. Dann existiert nach (ii) ein $y_0 \in Y$ so dass $\|x_0 - y_0\| \leq 2d(x_0, Y)$ ist. Nun sei $x := \|x_0 - y_0\|^{-1}(x_0 - y_0)$.)

(iv) Ist X unendlichdimensional, so existiert eine Folge $(x_n)_{n \in \mathbb{N}}$ in S so dass $\|x_n - x_m\| \geq 1/2$ ist für alle $m, n \in \mathbb{N}$ mit $m \neq n$. Daher ist S nicht kompakt.

Übung 1.1.12. Sei X ein endlichdimensionaler reeller Vektorraum. Eine Teilmenge $K \subset X$ heißt **konvex** wenn sie, für alle $x, y \in X$ und alle $t \in \mathbb{R}$, die Bedingung

$$x, y \in K, \quad 0 \leq t \leq 1 \quad \implies \quad (1-t)x + ty \in K$$

erfüllt.

(i) Sei $f : X \rightarrow \mathbb{R}$ eine Funktion. Die offene Teilmenge

$$U := \{(x, s) \in X \times \mathbb{R} \mid s > f(x)\}$$

ist genau dann konvex, wenn die Funktion f konvex ist.

(ii) Sei $U \subset X$ eine offene, beschränkte, konvexe Teilmenge so dass für alle $x \in X$ folgendes gilt:

$$x \in U \quad \iff \quad -x \in U.$$

(Nach Satz 1.1.8 hängen die Begriffe "offen" und "beschränkt" für Teilmengen von X nicht von der Norm ab, welche für die Definitionen verwendet wird.) Dann definiert die Formel

$$\|x\| := \inf \{\lambda > 0 \mid \lambda^{-1}x \in U\} \quad \text{für } x \in X$$

eine Normfunktion auf X und es gilt $U = \{x \in X \mid \|x\| < 1\}$.

Übung 1.1.13. Auf dem Raum $C([0, 1])$ betrachten wir die Normen $\|\cdot\|_p$ in Beispiel 1.1.4 für $1 \leq p \leq \infty$. Für welche p und q gibt es eine Konstante $c > 0$ so dass die Ungleichung $\|f\|_p \leq c \|f\|_q$ für alle $f \in C([0, 1])$ gilt? Zeigen Sie, dass die Normen $\|\cdot\|_p$ und $\|\cdot\|_q$ für $p \neq q$ nicht äquivalent sind.

1.2 Die Operatornorm

Seien $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ normierte reelle Vektorräume.

Definition 1.2.1. Ein linearer Operator $A : X \rightarrow Y$ heißt **beschränkt**, wenn es eine Konstante $c \geq 0$ gibt so dass die Ungleichung

$$\|Ax\|_Y \leq c \|x\|_X \quad (1.2.1)$$

für alle $x \in X$ gilt. Die kleinste solche Konstante $c \geq 0$ heißt **Norm** (beziehungsweise **Operatornorm**) von A und wird mit

$$\|A\| := \sup_{x \neq 0} \frac{\|Ax\|_Y}{\|x\|_X} \quad (1.2.2)$$

bezeichnet. (Hier ist das Supremum über alle von Null verschiedenen Vektoren $x \in X$ zu verstehen.) Die Menge der beschränkten linearen Operatoren von X nach Y wird mit $\mathcal{L}(X, Y)$ bezeichnet.

Der Begriff “Operator” wird in der mathematischen Literatur oft gleichbedeutend mit dem Begriff “Abbildung” verwendet. Die Herkunft dieser Wortwahl liegt darin begründet, dass viele der wichtigsten motivierenden Beispiele *Differentialoperatoren* oder *Integraloperatoren* aus der mathematischen Physik sind, für deren Studium die Theorie der linearen Abbildungen entwickelt wurde. Dies führt hin zum Gebiet der “*Funktionalanalysis*” welches wir in dieser Vorlesung nur am Rande berühren.

Korollar 1.2.2. Sei X endlichdimensional. Dann ist jeder lineare Operator $A : X \rightarrow Y$ beschränkt.

Beweis. Wir definieren eine Funktion

$$X \rightarrow \mathbb{R} : x \mapsto \|x\|_A := \|x\|_X + \|Ax\|_Y. \quad (1.2.3)$$

Da A linear ist, ist die Funktion (1.2.3) eine Norm. Da X endlichdimensional ist, existiert nach Satz 1.1.8 eine Konstante $c \geq 1$ so dass für jedes $x \in X$ die Ungleichung $\|x\|_A \leq c \|x\|_X$ gilt. Daraus folgt unmittelbar die gewünschte Ungleichung $\|Ax\|_Y \leq c \|x\|_X$ für alle $x \in X$. \square

Satz 1.2.3. Sei $A : X \rightarrow Y$ ein linearer Operator. Dann sind folgende Aussagen äquivalent.

- (i) A ist beschränkt.
- (ii) A ist stetig.
- (iii) A ist stetig an der Stelle $x = 0$.

Beweis. Wir zeigen, dass (ii) aus (i) folgt. Wenn A beschränkt ist, so ist A Lipschitz-stetig mit der Lipschitz-Konstanten $\|A\|$, denn es gilt

$$\begin{aligned} d_Y(Ax_0, Ax_1) &= \|Ax_0 - Ax_1\|_Y = \|A(x_0 - x_1)\|_Y \\ &\leq \|A\| \|x_0 - x_1\|_X = \|A\| d_X(x_0, x_1) \end{aligned}$$

für alle $x_0, x_1 \in X$; damit ist A auch stetig. Daß (iii) aus (ii) folgt, ergibt sich unmittelbar aus der Definition der Stetigkeit.

Wir zeigen, dass (i) aus (iii) folgt. Sei also A an der Stelle $x = 0$ stetig. Dann folgt aus der Definition der Stetigkeit mit $\varepsilon = 1$, dass ein $\delta > 0$ existiert, so dass jedes $x \in X$ mit $\|x\|_X \leq \delta$ die Ungleichung $\|Ax\|_Y \leq 1$ erfüllt. Sei nun $x \in X \setminus \{0\}$. Dann gilt

$$\left\| \frac{\delta}{\|x\|_X} x \right\|_X = \delta$$

und daher

$$1 \geq \left\| \frac{\delta}{\|x\|_X} Ax \right\|_Y = \frac{\delta}{\|x\|_X} \|Ax\|_Y.$$

Hieraus folgt $\|Ax\|_Y \leq \delta^{-1} \|x\|_X$ für alle $x \in X$. Also ist A beschränkt. \square

Beispiel 1.2.4. Seien $\|\cdot\|$ und $\|\cdot\|'$ zwei Normen auf einem reellen Vektorraum X . Wir betrachten die Aussage

$$\exists c > 0 \forall x \in X : \|x\|' \leq c \|x\|. \quad (1.2.4)$$

Diese Bedingung gilt genau dann, wenn der lineare Operator

$$\text{id} : (X, \|\cdot\|) \rightarrow (X, \|\cdot\|') \quad (1.2.5)$$

beschränkt ist. Nach Satz 1.2.3 heißt das, dass die Abbildung (1.2.5) stetig ist. Das wiederum heißt, dass das Urbild jeder offenen Menge in $(X, \|\cdot\|')$ unter der Abbildung (1.2.5) offen in $(X, \|\cdot\|)$ ist. Dies bedeutet

$$\mathcal{U}(X, \|\cdot\|') \subset \mathcal{U}(X, \|\cdot\|). \quad (1.2.6)$$

Also haben wir gezeigt, dass die Aussagen (1.2.4) und (1.2.6) äquivalent sind. Daraus folgt die Aussage

$$\|\cdot\| \sim \|\cdot\|' \iff \mathcal{U}(X, \|\cdot\|) = \mathcal{U}(X, \|\cdot\|')$$

für je zwei Normen auf einem reellen Vektorraum X (siehe auch Teil (ii) von Bemerkung 1.1.7). Nach Satz 1.1.8 heißt das wiederum, dass die Gesamtheit der offenen Mengen für einen endlichdimensionalen reellen Vektorraum X unabhängig von der Wahl der Norm ist. Damit besitzt jeder endlichdimensionale reelle Vektorraum eine *kanonische Topologie*.

Lemma 1.2.5. $\mathcal{L}(X, Y)$ ist ein normierter Vektorraum.

Beweis. Folgende drei Aussagen sind zu beweisen.

- (a) Für alle $A \in \mathcal{L}(X, Y)$ gilt $\|A\| = 0 \implies A = 0$.
- (b) Für alle $A \in \mathcal{L}(X, Y)$ und $\lambda \in \mathbb{R}$ gilt $\lambda A \in \mathcal{L}(X, Y)$ und $\|\lambda A\| = |\lambda| \|A\|$.
- (c) Für alle $A, B \in \mathcal{L}(X, Y)$ gilt $A+B \in \mathcal{L}(X, Y)$ und $\|A+B\| \leq \|A\| + \|B\|$.

Die Aussagen (a) und (b) folgen unmittelbar aus den Definitionen. Wir beweisen (c). Seien $A, B : X \rightarrow Y$ beschränkte lineare Operatoren. Dann gilt für alle $x \in X$ folgende Ungleichung

$$\begin{aligned} \|(A+B)x\|_Y &= \|Ax + Bx\|_Y \\ &\leq \|Ax\|_Y + \|Bx\|_Y \\ &\leq \|A\| \|x\|_X + \|B\| \|x\|_X \\ &= (\|A\| + \|B\|) \|x\|_X. \end{aligned}$$

Also ist $A+B$ beschränkt und für $x \in X \setminus \{0\}$ gilt

$$\frac{\|(A+B)x\|_Y}{\|x\|_X} \leq \|A\| + \|B\|.$$

Damit folgt die Dreiecksungleichung

$$\|A+B\| \leq \|A\| + \|B\|$$

aus der Definition der Operatornorm. □

Beispiel 1.2.6. Im Fall $X = \mathbb{R}^n$ und $Y = \mathbb{R}^m$ ist $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ der Raum aller linearen Abbildungen von \mathbb{R}^n nach \mathbb{R}^m (Korollar 1.2.2). Wir können ihn daher mit dem Raum $\mathbb{R}^{m \times n}$ der reellen $(m \times n)$ -Matrizen identifizieren, da jede lineare Abbildung von \mathbb{R}^n nach \mathbb{R}^m durch Multiplikation mit einer Matrix

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \in \mathbb{R}^{m \times n}$$

gegeben ist. In dieser Situation wird die Operatornorm der durch A induzierten linearen Abbildung bezüglich geeigneter Normen auf dem \mathbb{R}^n und dem \mathbb{R}^m auch eine **Matrixnorm** genannt. Hier sind drei Beispiele.

(i) Mit

$$\|x\|_1 := \sum_{j=1}^n |x_j|, \quad \|y\|_1 := \sum_{i=1}^m |y_i|$$

für $x \in \mathbb{R}^n$ und $y \in \mathbb{R}^m$ ist die Operatornorm von $A = (a_{ij}) \in \mathbb{R}^{m \times n}$ die Zahl

$$\|A\|_1 := \sup_{0 \neq x \in \mathbb{R}^n} \frac{\|Ax\|_1}{\|x\|_1} = \max_{j=1, \dots, n} \sum_{i=1}^m |a_{ij}|.$$

(ii) Mit

$$\|x\|_\infty := \max_{j=1, \dots, n} |x_j|, \quad \|y\|_\infty := \max_{i=1, \dots, m} |y_i|$$

für $x \in \mathbb{R}^n$ und $y \in \mathbb{R}^m$ ist die Operatornorm von $A = (a_{ij}) \in \mathbb{R}^{m \times n}$ die Zahl

$$\|A\|_\infty := \sup_{0 \neq x \in \mathbb{R}^n} \frac{\|Ax\|_\infty}{\|x\|_\infty} = \max_{i=1, \dots, m} \sum_{j=1}^n |a_{ij}|.$$

(iii) Mit

$$\|x\|_2 := \sqrt{\sum_{j=1}^n x_j^2}, \quad \|y\|_2 := \sqrt{\sum_{i=1}^m y_i^2}$$

für $x \in \mathbb{R}^n$ und $y \in \mathbb{R}^m$ ist die Operatornorm von $A = (a_{ij}) \in \mathbb{R}^{m \times n}$ die Wurzel des grössten Eigenwertes der Matrix $A^T A \in \mathbb{R}^{n \times n}$, das heißt

$$\|A\|_2 := \sup_{0 \neq x \in \mathbb{R}^n} \frac{\|Ax\|_2}{\|x\|_2} = \max\{\lambda \geq 0 \mid \ker(\lambda^2 \mathbb{1} - A^T A) \neq \{0\}\}.$$

(iv) Für $n > 1$ und $m > 1$ ist keine der Operatornormen in (i), (ii), und (iii) die Norm eines inneren Produktes auf dem Raum der reellen $(m \times n)$ -Matrizen. Dazu betrachten wir den Fall $m = n = 2$ und die Matrizen

$$A := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Dann haben die Matrizen A , B , $A + B$, $A - B$ alle die Norm 1 bezüglich jeder der drei Normen in (i), (ii), und (iii). Also gilt

$$\|A + B\|^2 + \|A - B\|^2 = 2 \neq 4 = 2\|A\|^2 + 2\|B\|^2$$

und damit ist die Parallelogrammidentität verletzt.

Satz 1.2.7. *Ist Y ein Banachraum so ist auch $\mathcal{L}(X, Y)$ ein Banachraum.*

Beweis. Es ist zu zeigen, dass $\mathcal{L}(X, Y)$ vollständig ist. Sei also $(A_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge in $\mathcal{L}(X, Y)$. Dann zeigt die Ungleichung

$$\|A_n x - A_m x\|_Y \leq \|A_n - A_m\| \|x\|_X,$$

dass die Folge $(A_n x)_{n \in \mathbb{N}}$ für jedes $x \in X$ eine Cauchy-Folge in Y ist. Da Y vollständig ist, konvergiert diese Folge und wir definieren $A : X \rightarrow Y$ durch

$$Ax := \lim_{n \rightarrow \infty} A_n x$$

für $x \in X$. Wir werden beweisen, dass A ein beschränkter linearer Operator ist und dass die Folge A_n in der Operatornorm gegen A konvergiert.

A ist linear. Für $x_0, x_1 \in X$ gilt

$$\begin{aligned} A(x_0 + x_1) &= \lim_{n \rightarrow \infty} A_n(x_0 + x_1) \\ &= \lim_{n \rightarrow \infty} A_n x_0 + \lim_{n \rightarrow \infty} A_n x_1 \\ &= Ax_0 + Ax_1. \end{aligned}$$

Genauso zeigt man, dass $A(\lambda x) = \lambda Ax$ ist für alle $\lambda \in \mathbb{R}$ und alle $x \in X$.

A ist beschränkt. Zunächst folgt aus der Ungleichung

$$\left| \|A_n\| - \|A_m\| \right| \leq \|A_n - A_m\|,$$

dass die Folge $(\|A_n\|)_{n \in \mathbb{N}}$ eine Cauchy-Folge reeller Zahlen ist und daher konvergiert. Sei

$$c := \lim_{n \rightarrow \infty} \|A_n\|.$$

Dann erfüllt jedes $x \in X$ die Ungleichung

$$\begin{aligned} \|Ax\|_Y &= \left\| \lim_{n \rightarrow \infty} A_n x \right\|_Y \\ &= \lim_{n \rightarrow \infty} \|A_n x\|_Y \\ &\leq \lim_{n \rightarrow \infty} \|A_n\| \|x\|_X \\ &= c \|x\|_X. \end{aligned}$$

Daher ist A beschränkt und $\|A\| \leq c$.

A_n **konvergiert gegen** A . Sei $\varepsilon > 0$. Da A_n eine Cauchy-Folge ist, gibt es ein $n_0 \in \mathbb{N}$ so dass für alle $n, m \in \mathbb{N}$ gilt:

$$m > n \geq n_0 \quad \implies \quad \|A_n - A_m\| < \varepsilon.$$

Hieraus folgt die Ungleichung $\|A_n x - A_m x\|_Y < \varepsilon \|x\|_X$ für alle $x \in X$ und $m > n \geq n_0$ und daher

$$\|A_n x - Ax\|_Y = \lim_{m \rightarrow \infty} \|A_n x - A_m x\|_Y \leq \varepsilon \|x\|_X$$

für alle $x \in X$ und alle $n \in \mathbb{N}$ mit $n \geq n_0$. Teilen wir durch $\|x\|_X$ im Fall $x \neq 0$ so ergibt sich

$$\|A_n - A\| = \sup_{x \neq 0} \frac{\|A_n x - Ax\|_Y}{\|x\|_X} \leq \varepsilon$$

für alle $n \in \mathbb{N}$ mit $n \geq n_0$. Damit ist Satz 1.2.7 bewiesen. \square

Ein wichtiger Spezialfall ist $Y = \mathbb{R}$. Der Banachraum $X^* := \mathcal{L}(X, \mathbb{R})$ heißt **Dualraum** des normierten Vektorraumes X . Satz 1.2.7 zeigt, dass dieser Dualraum stets vollständig ist, auch wenn X nicht vollständig ist.

Übung 1.2.8. (i) Sei X ein endlichdimensionaler normierter Vektorraum. Dann ist $X^* = \mathcal{L}(X, \mathbb{R})$ endlichdimensional und $\dim X^* = \dim X$.

(ii) Sei $n \in \mathbb{N}$ und seien $1 \leq p, q \leq \infty$ mit $1/p + 1/q = 1$. Für $y \in \mathbb{R}^n$ definieren wir die lineare Abbildung $\Lambda_y : \mathbb{R}^n \rightarrow \mathbb{R}$ durch

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

für $x \in \mathbb{R}^n$. Dann ist die Abbildung

$$\mathbb{R}^n \rightarrow (\mathbb{R}^n)^* : y \mapsto \Lambda_y$$

ein Vektorraum-Isomorphismus und es gilt

$$\|y\|_q = \sup_{x \neq 0} \frac{|\langle x, y \rangle|}{\|x\|_p} = \sup_{x \neq 0} \frac{|\Lambda_y(x)|}{\|x\|_p}$$

für alle $y \in \mathbb{R}^n$. (Siehe Beispiel 1.1.1 für $\|x\|_p$ und $\|y\|_q$.)

1.3 Banachalgebren

Sei $(X, \|\cdot\|)$ ein Banachraum. Dann ist auch der Raum $\mathcal{L}(X) := \mathcal{L}(X, X)$ aller beschränkten linearen Operatoren von X nach X ein Banachraum (siehe Satz 1.2.7). Dieser Raum besitzt eine Produktstruktur $(S, T) \mapsto S \circ T$, die durch die Komposition gegeben ist. Diese Produktstruktur ist im allgemeinen nicht kommutativ, und es gibt auch nicht für jeden von Null verschiedenen Operator $T \in \mathcal{L}(X)$ ein inverses Element; also ist $\mathcal{L}(X)$ kein Körper; dieser Raum erfüllt aber alle anderen Eigenschaften eines Körpers; einen solchen Raum nennt man eine Algebra. Der Begriff einer Banachalgebra ist nützlich für das Rechnen mit inversen Operatoren und Exponentialmatrizen.

Definition 1.3.1. *Eine reelle Algebra besteht aus einem reellen Vektorraum \mathcal{A} , einer Abbildung $\mathcal{A} \times \mathcal{A} \rightarrow \mathcal{A} : (x, y) \mapsto xy$, (das **Produkt** genannt) und einem Element $\mathbb{1} \in \mathcal{A}$ (das **1-Element** genannt) welche folgende Axiome erfüllen.*

(i) *Das Produkt ist assoziativ, das heißt für alle $x, y, z \in \mathcal{A}$ gilt*

$$(xy)z = x(yz).$$

(ii) *Das Produkt ist bilinear, das heißt für alle $x, y, z \in \mathcal{A}$ und alle $\lambda \in \mathbb{R}$ gilt*

$$x(y + z) = xy + xz, \quad (x + y)z = xz + yz, \quad x(\lambda z) = \lambda(xz) = (\lambda x)z.$$

(iii) *Für jedes $x \in \mathcal{A}$ gilt $\mathbb{1}x = x\mathbb{1} = x$.*

Eine reelle normierte Algebra besteht aus einer reellen Algebra \mathcal{A} und einer Norm $\mathcal{A} \rightarrow [0, \infty) : x \mapsto \|x\|$ welche die Produkt-Ungleichung

$$\|xy\| \leq \|x\| \|y\| \tag{1.3.1}$$

für alle $x, y \in \mathcal{A}$ erfüllt. Eine reelle Banachalgebra ist eine reelle normierte Algebra in der jede Cauchy-Folge konvergiert (das heißt, als normierter Vektorraum betrachtet ist \mathcal{A} vollständig, also ein Banachraum).

Man beachte, dass das 1-Element in einer reellen Algebra \mathcal{A} durch das Produkt eindeutig bestimmt ist. Ist nämlich $\mathbb{1}' \in \mathcal{A}$ ein weiteres Element welches die Bedingung $\mathbb{1}'x = x\mathbb{1}' = \mathbb{1}'$ für alle $x \in \mathcal{A}$ erfüllt, so gilt $\mathbb{1}' = \mathbb{1}'\mathbb{1} = \mathbb{1}$. Eine Algebra unterscheidet sich von einem Körper dadurch, dass sie auch eine (mit dem Produkt kompatible) Vektorraumstruktur besitzt, dass das Produkt nicht kommutativ sein muss, und dass nicht jedes von Null verschiedene Element von \mathcal{A} invertierbar sein muss.

Definition 1.3.2. Sei \mathcal{A} eine reelle Algebra. Ein Element $x \in \mathcal{A}$ heißt **invertierbar**, wenn ein Element $y \in \mathcal{A}$ existiert, welches die Bedingung

$$xy = yx = \mathbb{1} \quad (1.3.2)$$

erfüllt. Existiert ein solches Element, so ist es durch x eindeutig bestimmt (denn jedes weitere Element $y' \in \mathcal{A}$, welches der Bedingung $y'x = \mathbb{1}$ genügt, erfüllt die Gleichung $y' = y'\mathbb{1} = y'(xy) = (y'x)y = \mathbb{1}y = y$). Ist $x \in \mathcal{A}$ invertierbar so wird das eindeutige Element $y \in \mathcal{A}$, das die Bedingung (1.3.2) erfüllt, das **inverse Element von x** genannt und mit $x^{-1} := y$ bezeichnet. Die Gruppe der invertierbaren Elemente von \mathcal{A} bezeichnen wir mit

$$\mathcal{G} := \{x \in \mathcal{A} \mid \text{es gibt ein } y \in \mathcal{A} \text{ so dass } xy = yx = \mathbb{1}\}.$$

Lemma 1.3.3. Sei \mathcal{A} eine reelle Algebra mit 1-Element $\mathbb{1}$. Dann gilt folgendes.

- (i) $\mathbb{1} \in \mathcal{G}$ und $\mathbb{1}^{-1} = \mathbb{1}$.
- (ii) Ist $x \in \mathcal{G}$ so ist $x^{-1} \in \mathcal{G}$ und es gilt $(x^{-1})^{-1} = x$.
- (iii) Sind $x, y \in \mathcal{G}$ so ist $xy \in \mathcal{G}$ und es gilt $(xy)^{-1} = y^{-1}x^{-1}$.

Beweis. Diese Aussagen folgen unmittelbar aus der Definition. \square

Beispiel 1.3.4. Die reellen und die komplex Zahlen bilden Banachalgebren mit dem üblichen Produkt und der Betragsfunktion als Norm. Ein ähnliches Beispiel ist durch die Quaternionen gegeben. Die **Quaternionen** sind durch eine geeignete Produktstruktur auf dem reellen Vektorraum $\mathbb{H} \cong \mathbb{R}^4$ definiert. Wir schreiben die Elemente von \mathbb{H} als formale Summen

$$x = x_0 + \mathbf{i}x_1 + \mathbf{j}x_2 + \mathbf{k}x_3$$

mit $x_i \in \mathbb{R}$ und definieren die Produktstruktur so, dass die Gleichungen

$$\mathbf{j}\mathbf{k} = -\mathbf{k}\mathbf{j} = \mathbf{i}, \quad \mathbf{k}\mathbf{i} = -\mathbf{i}\mathbf{k} = \mathbf{j}, \quad \mathbf{i}\mathbf{j} = -\mathbf{j}\mathbf{i} = \mathbf{k}, \quad \mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1$$

gelten und 1 das 1-Element ist. Die Norm auf \mathbb{H} ist $|x| := \sqrt{\sum_{i=0}^3 x_i^2}$ und es gilt $|xy| = |x||y|$ für alle $x, y \in \mathbb{H}$. Eine bilineare Abbildung, die ebenfalls diese Eigenschaft hat aber nicht mehr assoziativ ist, gibt es noch in der Dimension acht (die Oktonionen), aber in keiner Dimension ausser 1, 2, 4, 8.

Beispiel 1.3.5. Ist X ein Banachraum so ist $\mathcal{L}(X)$ mit der Operatornorm, der Komposition, und der Identität eine Banachalgebra. Die Vollständigkeit folgt aus Satz 1.2.7. **Übung:** Die Operatornorm auf $\mathcal{L}(X)$ erfüllt die Bedingung (1.3.1).

Beispiel 1.3.6. Der Raum $\mathbb{R}^{n \times n}$ der quadratischen Matrizen ist eine Banachalgebra mit jeder Norm welche der Bedingung (1.3.1) genügt. Eine solche Norm auf dem Raum $\mathbb{R}^{n \times n}$ wird **Matrixnorm** genannt. Zum Beispiel kann dies die zu einer beliebigen Norm auf dem \mathbb{R}^n gehörige Operatornorm auf $\mathbb{R}^{n \times n}$ sein. Ein anderes Beispiel wäre die euklidische Norm die sich durch die Identifikation $\mathbb{R}^{n \times n} \cong \mathbb{R}^{n^2}$ ergibt; in diesem Fall erhält man

$$\|A\| = \left(\sum_{i,j=1}^n a_{ij}^2 \right)^{1/2}$$

für $A = (a_{ij}) \in \mathbb{R}^{n \times n}$. **Übung:** Diese Norm erfüllt die Bedingung (1.3.1) sowie $\|Ax\|_2 \leq \|A\| \|x\|_2$ für jede Matrix $A \in \mathbb{R}^{n \times n}$ und jeden Vektor $x \in \mathbb{R}^n$.

Beispiel 1.3.7. Sei (X, d) ein kompakter metrischer Raum. Dann ist $C(X)$ mit der Supremumsnorm eine Banachalgebra (siehe Beispiel 1.1.3).

Beispiel 1.3.8. Eine Folge $x = (x_0, x_1, x_2, \dots)$ reeller Zahlen heißt **absolut summierbar** wenn die Reihe $\sum_{k=0}^{\infty} x_k$ absolut konvergiert. Die Norm einer solchen absolut summierbaren Folge ist definiert durch

$$\|x\| := \sum_{k=0}^{\infty} |x_k|.$$

Wir bezeichnen mit ℓ^1 den Raum aller absolut summierbaren Folgen $(x_k)_{k \in \mathbb{N}_0}$ reeller Zahlen. Der Vektorraum ℓ^1 ist mit dieser Norm ein Banachraum. Die **Faltung** zweier Folgen $x, y \in \ell^1$ ist definiert durch

$$(x * y)_k := \sum_{i=0}^k x_i y_{k-i} \quad \text{für } k \in \mathbb{N}_0.$$

Diese Folge ist ebenfalls absolut summierbar und es gilt $\|x * y\| \leq \|x\| \|y\|$. (Dies folgt aus dem grossen Umordnungssatz in [4].) Damit ist ℓ^1 eine Banachalgebra. **Übung:** ℓ^1 ist vollständig.

Übung 1.3.9. In jeder Banachalgebra gilt $\|\mathbb{1}\| \geq 1$. Für welche der oben genannten Banachalgebren ist das Produkt kommutativ? Ist \mathcal{A} eine Banachalgebra so ist \mathcal{A}^2 mit der Normfunktion $\mathcal{A}^2 \rightarrow \mathbb{R} : (x, y) \mapsto \|x\| + \|y\|$ ein Banachraum und die Produktabbildung $\mathcal{A}^2 \rightarrow \mathcal{A} : (x, y) \mapsto xy$ ist stetig.

Reihen in Banachräumen

Die Begriffe von Konvergenz und Stetigkeit wurden in der Analysis I Vorlesung für beliebige metrische Räume eingeführt. Sie gelten also insbesondere für Funktionen und Folgen mit Werten in einem Banachraum oder einer Banachalgebra. Aufgrund der Vektorraumstruktur läßt sich der Begriff der Konvergenz einer Reihe auf Reihen mit Werten in Banachräumen übertragen. Ist $(X, \|\cdot\|)$ ein Banachraum und $(x_k)_{k \in \mathbb{N}_0}$ eine Folge in X , so nennen wir die Reihe $\sum_{k=0}^{\infty} x_k$ **konvergent**, wenn die Folge der **Partialsommen** $s_n := \sum_{k=0}^n x_k$ konvergiert; in dem Fall bezeichnen wir den Grenzwert mit

$$\sum_{k=0}^{\infty} x_k := \lim_{n \rightarrow \infty} \sum_{k=0}^n x_k.$$

Wir benutzen also die gleiche Bezeichnungsweise wie für Reihen reeller oder komplexer Zahlen, das heißt der Ausdruck “ $\sum_{k=0}^{\infty} x_k$ ” hat zwei Bedeutungen: er steht zum einen für die Folge der Partialsommen $(s_n)_{n \in \mathbb{N}_0}$ und zum anderen für deren Grenzwert, falls dieser existiert. Wir nennen die Reihe $\sum_{k=0}^{\infty} x_k$ **absolut konvergent**, wenn die Reihe $\sum_{k=0}^{\infty} \|x_k\|$ konvergiert. In diesem Fall nennen wir auch die Folge $(x_k)_{k \in \mathbb{N}_0}$ **absolut summierbar**.

Satz 1.3.10. *Sei $(X, \|\cdot\|)$ ein Banachraum und sei $(x_k)_{k \in \mathbb{N}_0}$ eine Folge in X . Konvergiert die Reihe $\sum_{k=0}^{\infty} x_k$ absolut, so konvergiert sie.*

Beweis. Sei $s_n := \sum_{k=0}^n x_k$ die Folge der Partialsommen und sei $\varepsilon > 0$. Da die Reihe $\sum_{k=0}^{\infty} x_k$ absolut konvergiert, konvergiert die Reihe $\sum_{k=0}^{\infty} \|x_k\|$ in \mathbb{R} . Das heißt, die Folge der Partialsommen $\sigma_n := \sum_{k=0}^n \|x_k\|$ der Reihe der Normen konvergiert und ist daher eine Cauchy-Folge. Also gibt es eine natürliche Zahl $n_0 \in \mathbb{N}$ so dass für alle $m, n \in \mathbb{N}$ folgendes gilt:

$$n > m \geq n_0 \quad \implies \quad \sigma_n - \sigma_m = \|x_{m+1}\| + \cdots + \|x_n\| < \varepsilon.$$

Es folgt aus der Dreiecksungleichung, dass für alle $m, n \in \mathbb{N}$ mit $n > m \geq n_0$ die folgende Ungleichung gilt

$$\|s_n - s_m\| = \|x_{m+1} + \cdots + x_n\| \leq \|x_{m+1}\| + \cdots + \|x_n\| < \varepsilon.$$

Damit ist gezeigt, dass $(s_n)_{n \in \mathbb{N}_0}$ eine Cauchy-Folge in X ist. Da X ein Banachraum (und somit vollständig) ist, konvergiert diese Folge, und das heißt, dass die Reihe $\sum_{k=0}^{\infty} x_k$ konvergiert. Damit ist Satz 1.3.10 bewiesen. \square

Potenzreihen in Banachalgebren

Sei \mathcal{A} eine reelle Banachalgebra und

$$P(z) := \sum_{k=0}^{\infty} a_k z^k \quad (1.3.3)$$

eine Potenzreihe mit reellen Koeffizienten $a_k \in \mathbb{R}$ und Konvergenzradius

$$\rho := \rho_P := \frac{1}{\limsup_{n \rightarrow \infty} |a_n|^{1/n}}.$$

Wir nehmen an, dieser Konvergenzradius sei positiv (oder sogar unendlich). Wir wollen nun die Variable z durch ein Element der Banachalgebra \mathcal{A} ersetzen. Formal macht dies Sinn, da ja in einer Banachalgebra die Produkte x^k erklärt sind. Insbesondere definieren wir $x^0 := \mathbb{1}$ für jedes $x \in \mathcal{A}$. Die daraus resultierende Reihe in \mathcal{A} ist

$$P_{\mathcal{A}}(x) := \sum_{k=0}^{\infty} a_k x^k. \quad (1.3.4)$$

Man kann nun die Frage stellen, für welche $x \in \mathcal{A}$ diese Reihe konvergiert.

Korollar 1.3.11. *Die Reihe (1.3.4) konvergiert für jedes $x \in \mathcal{A}$ mit $\|x\| < \rho$.*

Beweis. Sei $x \in \mathcal{A}$ mit $\|x\| < \rho$. Es folgt aus der Bedingung (1.3.1) in der Definition einer Banachalgebra, dass die Ungleichung

$$\|a_k x^k\| \leq |a_k| \|x\|^k$$

für jedes $k \in \mathbb{N}$ gilt. Da $\|x\| < \rho$ ist, konvergiert die Reihe $\sum_{k=0}^{\infty} |a_k| \|x\|^k$. Nach dem Majorantenkriterium konvergiert auch die Reihe $\sum_{k=0}^{\infty} \|a_k x^k\|$. Dies heißt aber genau, dass die Reihe $\sum_{k=0}^{\infty} a_k x^k$ in \mathcal{A} absolut konvergiert und somit, nach Satz 1.3.10, konvergiert. \square

Wir haben also gezeigt, dass die Formel (1.3.4) eine Abbildung

$$P_{\mathcal{A}} : \mathcal{B}_{\rho} := \{x \in \mathcal{A} \mid \|x\| < \rho\} \rightarrow \mathcal{A} \quad (1.3.5)$$

definiert, wobei ρ der Konvergenzradius von P ist. Das nächste Lemma zeigt, dass diese Abbildung stetig ist. Zunächst bemerken wir noch, dass sich der Begriff der Differenzierbarkeit, und der der Ableitung, auf Funktionen mit Werten in einem Banachraum Wort für Wort übertragen läßt.

Definition 1.3.12. Sei $I \subset \mathbb{R}$ ein Intervall und $(X, \|\cdot\|)$ ein Banachraum. Eine Funktion $f : I \rightarrow X$ heißt **differenzierbar an der Stelle** $t \in I$ mit **Ableitung** $a := f'(t) \in X$ wenn gilt: $\forall \varepsilon > 0 \exists \delta > 0 \forall h \in \mathbb{R}$:

$$0 < |h| < \delta, \quad t + h \in I \quad \Longrightarrow \quad \frac{\|f(t+h) - f(t) - ha\|}{|h|} < \varepsilon.$$

Äquivalenterweise heißt das, dass der Differenzenquotient $h^{-1}(f(t+h) - f(t))$ in X gegen a konvergiert (für $h \rightarrow 0$). Mit anderen Worten

$$f'(t) := \lim_{h \rightarrow 0} \frac{f(t+h) - f(t)}{h},$$

falls dieser Grenzwert existiert, und wenn er existiert so nennen wir f differenzierbar an der Stelle t .

Satz 1.3.13. Sei \mathcal{A} eine reelle Banachalgebra und sei (1.3.3) eine Potenzreihe mit reellen Koeffizienten und Konvergenzradius $\rho > 0$. Dann gilt folgendes

(i) Für jedes $r < \rho$ ist die Restriktion der Abbildung (1.3.5) auf die Teilmenge $\overline{B}_r := \{x \in \mathcal{A} \mid \|x\| \leq r\}$ Lipschitz-stetig; es gilt

$$\|P_{\mathcal{A}}(x) - P_{\mathcal{A}}(y)\| \leq c_r \|x - y\|, \quad c_r := \sum_{k=1}^{\infty} k |a_k| r^{k-1},$$

für $x, y \in \mathcal{A}$ mit $\|x\|, \|y\| \leq r$.

(ii) Für jedes $x \in \mathcal{A}$ ist die Funktion $t \mapsto P_{\mathcal{A}}(tx)$ auf dem offenen Intervall $(-\rho/\|x\|, \rho/\|x\|)$ differenzierbar und es gilt

$$\frac{d}{dt} P_{\mathcal{A}}(tx) = xQ_{\mathcal{A}}(tx), \quad Q(z) := P'(z) := \sum_{k=1}^{\infty} k a_k z^{k-1}, \quad (1.3.6)$$

für alle $t \in \mathbb{R}$ mit $|t| \|x\| < \rho$.

Beweis. Wir beweisen (i). Es gilt $x^k - y^k = \sum_{j=1}^k x^{k-j}(x-y)y^{j-1}$ für alle $k \in \mathbb{N}$ und alle $x, y \in \mathcal{A}$, und daher

$$\|x\| \leq r, \quad \|y\| \leq r \quad \Longrightarrow \quad \|x^k - y^k\| \leq k r^{k-1} \|x - y\|.$$

Hieraus folgt

$$\|P_{\mathcal{A}}(x) - P_{\mathcal{A}}(y)\| \leq \sum_{k=0}^{\infty} |a_k| \|x^k - y^k\| \leq \sum_{k=0}^{\infty} k |a_k| r^{k-1} \|x - y\| = c_r \|x - y\|$$

für alle $x, y \in \mathcal{A}$ mit $\|x\| \leq r$ und $\|y\| \leq r$. Damit ist (i) bewiesen.

Zum Beweis von (ii) beachten wir, dass $Q_{\mathcal{A}}(x) = \sum_{k=1}^{\infty} k a_k x^{k-1}$ ist. Daher gilt

$$\begin{aligned}
& \|P_{\mathcal{A}}((t+h)x) - P_{\mathcal{A}}(tx) - hxQ_{\mathcal{A}}(tx)\| \\
&= \left\| \sum_{k=2}^{\infty} ((t+h)^k - t^k - kht^{k-1}) a_k x^k \right\| \\
&\leq \sum_{k=2}^{\infty} |(t+h)^k - t^k - kht^{k-1}| |a_k| \|x\|^k \\
&\leq \sum_{k=2}^{\infty} \sum_{j=2}^k \binom{k}{j} |h|^j |t|^{k-j} |a_k| \|x\|^k \\
&= \sum_{k=2}^{\infty} \left((|t|+|h|)^k - |t|^k - k|h||t|^{k-1} \right) |a_k| \|x\|^k \\
&= |p(|t|+|h|) - p(|t|) - |h|\dot{p}(|t|)|.
\end{aligned}$$

Hier bezeichnet p die Potenzreihe

$$p(\lambda) := \sum_{k=0}^{\infty} |a_k| \|x\|^k \lambda^k$$

mit dem Konvergenzradius $\rho/\|x\|$. Sei $t \in \mathbb{R}$ mit $|t| < \rho/\|x\|$. Dann existieren zwei Konstanten $\delta > 0$ und $c > 0$ so dass $|t| + \delta < \rho/\|x\|$ ist und jedes $h \in \mathbb{R}$ mit $|h| < \delta$ die Ungleichung

$$|p(|t|+|h|) - p(|t|) - |h|\dot{p}(|t|)| \leq c|h|^2$$

erfüllt. Daraus folgt

$$\|P_{\mathcal{A}}((t+h)x) - P_{\mathcal{A}}(tx) - hxQ_{\mathcal{A}}(tx)\| \leq c|h|^2$$

für jedes $h \in \mathbb{R}$ mit $|h| < \delta$. Damit haben wir gezeigt, dass

$$\lim_{h \rightarrow 0} \left\| \frac{P_{\mathcal{A}}((t+h)x) - P_{\mathcal{A}}(tx)}{h} - xQ_{\mathcal{A}}(tx) \right\| = 0$$

ist und daraus folgt (ii). Damit ist Satz 1.3.13 bewiesen. \square

Betrachten wir den Spezialfall $\mathcal{A} := \mathbb{R}^{n \times n}$ mit einer Matrixnorm $\|\cdot\|$. Dann folgt aus Satz 1.3.13, dass die Matrix-Reihe $P(A) := \sum_{k=0}^{\infty} a_k A^k$ für jede Matrix $A \in \mathbb{R}^{n \times n}$ mit $\|A\| < \rho$ in dieser Norm konvergiert. Während die Schlussfolgerung (Konvergenz) nach Satz 1.1.8 unabhängig von der Wahl der Norm ist, benötigen wir bei der Voraussetzung ($\|A\| < \rho$) eine Matrixnorm.

1.4 Die geometrische Reihe

Die **geometrische Reihe** hat den Konvergenzradius $\rho = 1$ und stellt die Funktion

$$\sum_{k=0}^{\infty} z^k = \frac{1}{1-z} \quad \text{für } z \in \mathbb{C} \text{ mit } |z| < 1$$

dar. Nach Satz 1.3.13 ergibt sich daraus folgendes Resultat.

Satz 1.4.1. *Sei \mathcal{A} eine reelle Banachalgebra mit $\|\mathbb{1}\| = 1$ und sei $\mathcal{G} \subset \mathcal{A}$ die Gruppe der invertierbaren Elemente. Dann gilt folgendes.*

(i) *Sei $x \in \mathcal{A}$ mit $\|x\| < 1$. Dann ist $\mathbb{1} - x \in \mathcal{G}$ und*

$$(\mathbb{1} - x)^{-1} = \sum_{k=0}^{\infty} x^k. \quad (1.4.1)$$

(ii) *Sei $x \in \mathcal{G}$ und $y \in \mathcal{A}$ so dass*

$$\|x - y\| \|x^{-1}\| < 1. \quad (1.4.2)$$

Dann ist $y \in \mathcal{G}$ und

$$y^{-1} = x^{-1} \sum_{k=0}^{\infty} (\mathbb{1} - yx^{-1})^k, \quad \|x^{-1} - y^{-1}\| \leq \frac{\|x^{-1}\|^2 \|x - y\|}{1 - \|x^{-1}\| \|x - y\|}. \quad (1.4.3)$$

(iii) *Die Gruppe \mathcal{G} der invertierbaren Elemente in \mathcal{A} ist eine offene Teilmenge von \mathcal{A} und die Abbildung $\mathcal{G} \rightarrow \mathcal{G} : x \mapsto x^{-1}$ ist stetig.*

Beweis. Sei $x \in \mathcal{A}$ mit $\|x\| < 1$. Dann gilt

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

Daher konvergiert Reihe $\sum_{k=0}^{\infty} x^k$ nach Satz 1.3.10. Für alle $n \in \mathbb{N}$ gilt

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

und daraus folgt $(\mathbb{1} - x) \sum_{k=0}^{\infty} x^k = \sum_{k=0}^{\infty} x^k (\mathbb{1} - x) = \mathbb{1}$. Also ist $\mathbb{1} - x \in \mathcal{G}$ und $(\mathbb{1} - x)^{-1} = \sum_{k=0}^{\infty} x^k$ und damit ist (i) bewiesen.

Nun seien $x \in \mathcal{G}$ und $y \in \mathcal{A}$ gegeben mit $\|x - y\| \|x^{-1}\| < 1$. Dann gilt

$$\|\mathbb{1} - yx^{-1}\| = \|(x - y)x^{-1}\| \leq \|x - y\| \|x^{-1}\| < 1.$$

Nach Teil (i) folgt daraus $yx^{-1} \in \mathcal{G}$ und $(yx^{-1})^{-1} = \sum_{k=0}^{\infty} (\mathbb{1} - yx^{-1})^k$. Daher ist $y \in \mathcal{G}$ und es gilt

$$y^{-1} = x^{-1}(yx^{-1})^{-1} = x^{-1} \sum_{k=0}^{\infty} (\mathbb{1} - yx^{-1})^k$$

und

$$\begin{aligned} \|x^{-1} - y^{-1}\| &= \left\| x^{-1} \sum_{k=1}^{\infty} (\mathbb{1} - yx^{-1})^k \right\| \\ &\leq \|x^{-1}\| \sum_{k=1}^{\infty} \|\mathbb{1} - yx^{-1}\|^k \\ &= \|x^{-1}\| \frac{\|\mathbb{1} - yx^{-1}\|}{1 - \|\mathbb{1} - yx^{-1}\|} \\ &\leq \frac{\|x^{-1}\|^2 \|x - y\|}{1 - \|x^{-1}\| \|x - y\|}. \end{aligned}$$

Damit ist Teil (ii) bewiesen und Teil (iii) folgt unmittelbar aus (ii). Damit ist Satz 1.4.1 bewiesen. \square

Beispiel 1.4.2. Wir betrachten die Banachalgebra

$$\mathcal{A} = \mathbb{R}^{n \times n}$$

aller reellen $n \times n$ -Matrizen. Jede Matrix $A \in \mathbb{R}^{n \times n}$ bestimmt eine lineare Abbildung $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ durch $T_A x := Ax$ für $x \in \mathbb{R}^n$. Nach Definition ist eine Matrix $A \in \mathbb{R}^{n \times n}$ genau dann ein invertierbares Element von \mathcal{A} , wenn eine Matrix $B \in \mathbb{R}^{n \times n}$ existiert, welche die Gleichungen $AB = BA = \mathbb{1}$ erfüllt. Solch eine Matrix B existiert genau dann wenn die lineare Abbildung T_A bijektiv ist. In diesem Fall induziert die inverse Matrix $A^{-1} = B$ die Umkehrabbildung $T_A^{-1} = T_{A^{-1}}$. Aus der *Linearen Algebra* wissen wir, dass die invertierbaren Matrizen gerade die mit Determinante ungleich Null sind (siehe Anhang D). Daher ist die Gruppe der invertierbaren Elemente in $\mathcal{A} = \mathbb{R}^{n \times n}$ durch $\text{GL}(n, \mathbb{R}) := \{A \in \mathbb{R}^{n \times n} \mid \det(A) \neq 0\}$ gegeben. Diese ist die **allgemeine lineare Gruppe**. Nach Satz 1.4.1 ist dies eine offene Teilmenge von $\mathbb{R}^{n \times n}$ und die Abbildung $\text{GL}(n, \mathbb{R}) \rightarrow \text{GL}(n, \mathbb{R}) : A \mapsto A^{-1}$ ist stetig.

1.5 Die Exponentialabbildung

Die Potenzreihe $\exp(z) := \sum_{k=0}^{\infty} z^k/k!$ hat den Konvergenzradius $\rho = \infty$ und induziert damit eine Exponentialabbildung auf jeder Banachalgebra.

Definition 1.5.1. Sei \mathcal{A} eine reelle Banachalgebra. Die **Exponentialabbildung auf \mathcal{A}** ist die durch

$$\exp(x) := e^x := \sum_{k=0}^{\infty} \frac{x^k}{k!} \quad \text{für } x \in \mathcal{A} \quad (1.5.1)$$

definierte Abbildung

$$\exp : \mathcal{A} \rightarrow \mathcal{A}.$$

Nach Satz 1.3.13 konvergiert die Reihe in (1.5.1) absolut für jedes $x \in \mathcal{A}$.

Satz 1.5.2. Sei \mathcal{A} eine reelle Banachalgebra. Dann ist die Exponentialabbildung $\exp : \mathcal{A} \rightarrow \mathcal{A}$ in (1.5.1) stetig und hat folgende Eigenschaften.

(i) Für jedes $x \in \mathcal{A}$ und jede Folge $(x_n)_{n \in \mathbb{N}}$ in \mathcal{A} gilt

$$\lim_{n \rightarrow \infty} x_n = x \quad \implies \quad \exp(x) = \lim_{n \rightarrow \infty} \left(\mathbb{1} + \frac{x_n}{n} \right)^n. \quad (1.5.2)$$

(ii) Es gilt $\exp(0) = \mathbb{1}$ und, für alle $x, y \in \mathcal{A}$,

$$xy = yx \quad \implies \quad \exp(x + y) = \exp(x) \exp(y). \quad (1.5.3)$$

(iii) Für jedes $x \in \mathcal{A}$ gilt

$$\exp(x) \in \mathcal{G}, \quad \exp(x)^{-1} = \exp(-x). \quad (1.5.4)$$

(iv) Für jedes $x \in \mathcal{A}$ und jedes $y \in \mathcal{G}$ gilt

$$\exp(yxy^{-1}) = y \exp(x) y^{-1}. \quad (1.5.5)$$

(v) Für jedes $x \in \mathcal{A}$ ist die Funktion $\mathbb{R} \rightarrow \mathcal{A} : t \mapsto \exp(tx)$ stetig differenzierbar und ihre Ableitung ist

$$\frac{d}{dt} \exp(tx) = x \exp(tx). \quad (1.5.6)$$

Beweis. Die Stetigkeit der Exponentialabbildung folgt aus Satz 1.3.13 und die Aussage (i) beweist man wie im Fall $\mathcal{A} = \mathbb{R}$ (siehe [4]). Hier ist noch einmal das Argument zum Beweis von (i). Sei $\varepsilon > 0$ gegeben und sei $N \in \mathbb{N}$ so gewählt, dass $\|x_n - x\| \leq 1$ ist für jede natürliche Zahl $n \geq N$ und

$$\sum_{k=N}^{\infty} \frac{(\|x\| + 1)^k}{k!} < \frac{\varepsilon}{3}. \quad (1.5.7)$$

Dann gilt $\lim_{n \rightarrow \infty} \binom{n}{k} \frac{1}{n^k} = \frac{1}{k!}$ für jedes $k \in \{0, 1, \dots, N-1\}$. Daher existiert eine ganze Zahl $n_0 \geq N$ so dass für alle $n \in \mathbb{N}$ folgendes gilt:

$$n \geq n_0 \quad \implies \quad \sum_{k=0}^{N-1} \left\| \binom{n}{k} \frac{x_n^k}{n^k} - \frac{x^k}{k!} \right\| < \frac{\varepsilon}{3}. \quad (1.5.8)$$

Sei nun $n \in \mathbb{N}$ mit $n \geq n_0$. Dann ist $n \geq N$ und es gilt

$$\begin{aligned} \left\| \left(\mathbb{1} + \frac{x_n}{n} \right) - e^x \right\| &= \left\| \sum_{k=0}^n \binom{n}{k} \frac{x_n^k}{n^k} - \sum_{k=0}^{\infty} \frac{x^k}{k!} \right\| \\ &\leq \sum_{k=0}^{N-1} \left\| \binom{n}{k} \frac{x_n^k}{n^k} - \frac{x^k}{k!} \right\| + \sum_{k=N}^n \binom{n}{k} \frac{\|x_n\|^k}{n^k} + \sum_{k=N}^{\infty} \frac{\|x\|^k}{k!} \\ &\leq \sum_{k=0}^{N-1} \left\| \binom{n}{k} \frac{x_n^k}{n^k} - \frac{x^k}{k!} \right\| + 2 \sum_{k=N}^{\infty} \frac{(\|x\| + 1)^k}{k!} \\ &< \varepsilon. \end{aligned}$$

Hier folgt der vorletzte Schritt aus den Ungleichungen $\|x_n\| \leq \|x\| + 1$ für $n \geq N$ und $\binom{n}{k} \frac{1}{n^k} \leq \frac{1}{k!}$ für $n \in \mathbb{N}$ mit $n \geq k$. Der letzte Schritt folgt aus (1.5.7) und (1.5.8). Damit ist (i) bewiesen.

Die Aussage (ii) beweist man ebenfalls genau wie im Fall $\mathcal{A} = \mathbb{R}$ (siehe [4]). Man kann diese Aussage jedoch auch direkt aus (i) herleiten, denn wenn $x, y \in \mathcal{A}$ kommutieren so dass $xy = yx$ ist, dann gilt

$$\left(\mathbb{1} + \frac{x}{n} \right)^n \left(\mathbb{1} + \frac{y}{n} \right)^n = \left(\mathbb{1} + \frac{x + y + xy/n}{n} \right)^n$$

für jedes $n \in \mathbb{N}$. Beim Grenzübergang $n \rightarrow \infty$ ergibt sich daraus nach (i) die gewünschte Gleichung $\exp(x) \exp(y) = \exp(x + y)$.

Teil (iii) folgt unmittelbar aus Teil (ii) mit $y = -x$, Teil (iv) folgt aus der Gleichung $(yxy^{-1})^k = yx^ky^{-1}$ für $y \in \mathcal{G}$, und Teil (v) folgt aus Satz 1.3.13. Damit ist Satz 1.5.2 bewiesen. \square

Lineare Differentialgleichungssysteme

Die Exponentialabbildung auf der Banachalgebra $\mathcal{A} = \mathbb{R}^{n \times n}$ aller $(n \times n)$ -Matrizen spielt eine wichtige Rolle bei der Lösung eines Systems von n linearen Differentialgleichungen erster Ordnung in n Variablen. Ein solches Differentialgleichungssystem hat die Form

$$\begin{aligned} \dot{x}_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n, & x_1(0) &= x_{01}, \\ \dot{x}_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n, & x_2(0) &= x_{02}, \\ &\vdots & & \\ \dot{x}_n &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n, & x_n(0) &= x_{0n}. \end{aligned} \tag{1.5.9}$$

Hier sind die $a_{ij} \in \mathbb{R}$ konstante reelle Koeffizienten und, für jedes i , ist die Variable x_i eine stetig differenzierbare Abbildung $x_i : \mathbb{R} \rightarrow \mathbb{R}$. Wir stellen uns das Argument dieser Abbildung als Zeitvariable t vor und bezeichnen mit $\dot{x}_i = dx_i/dt$ die Ableitung der Funktion x_i nach der Zeit t . In der physikalischen Interpretation bezeichnet das n -Tupel

$$x(t) := (x_1(t), \dots, x_n(t)) \in \mathbb{R}^n$$

den Ortsvektor eines Teilchens (oder mehrerer Teilchen gleichzeitig) zum Zeitpunkt t und die Ableitung $\dot{x}(t) = (\dot{x}_1(t), \dots, \dot{x}_n(t)) \in \mathbb{R}^n$ den Geschwindigkeitsvektor zum Zeitpunkt t . Fasst man die konstanten Koeffizienten zu einer Matrix $A = (a_{ij})_{i,j=1}^n \in \mathbb{R}^{n \times n}$ zusammen so läßt sich das Differentialgleichungssystem (1.5.9) auch in der Kurzform

$$\dot{x} = Ax, \quad x(0) = x_0, \tag{1.5.10}$$

schreiben. Hier bezeichnet $x_0 := (x_{01}, \dots, x_{0n}) \in \mathbb{R}^n$ den **Anfangsvektor** (oder den Ortsvektor zum Zeitpunkt $t = 0$). Es folgt nun aus der Gleichung (1.5.6), dass die eindeutige Lösung der Differentialgleichung (1.5.10) durch die Formel

$$x(t) := e^{At}x_0 \tag{1.5.11}$$

gegeben ist, wobei

$$e^{At} = \sum_{k=0}^{\infty} \frac{A^k t^k}{k!} \in \mathbb{R}^{n \times n}$$

die Exponentialmatrix bezeichnet. Die Konvergenz dieser Reihe folgt aus der Konvergenz der Exponentialabbildung für beliebige Banachalgebren \mathcal{A} im Spezialfall $\mathcal{A} = \mathbb{R}^{n \times n}$ (siehe Satz 1.5.2).

Beispiel 1.5.3. Für

$$\Lambda := \text{diag}(\lambda_1, \dots, \lambda_n)$$

gilt $\Lambda^k = \text{diag}(\lambda_1^k, \dots, \lambda_n^k)$ und damit

$$e^\Lambda = \sum_{k=0}^{\infty} \frac{\Lambda^k}{k!} = \text{diag}(e^{\lambda_1}, \dots, e^{\lambda_n}).$$

Beispiel 1.5.4. Für $\lambda \in \mathbb{R}$ gilt

$$A = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} \implies e^{At} = e^{\lambda t} \begin{pmatrix} 1 & t & t^2/2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}.$$

Übung: Verwenden Sie diese Formel zur Bestimmung der Lösungen eines geeigneten Systems von drei Differentialgleichungen erster Ordnung in drei Variablen. Verallgemeinern Sie die Formel in diesem Beispiel auf vier und fünf Variablen, dann auf Matrizen beliebiger Grösse.

Beispiel 1.5.5. Für

$$H := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

gilt $H^{2k} = (-1)^k \mathbb{1}$ und damit

$$\begin{aligned} e^{Ht} &= \left(1 - \frac{t^2}{2!} + \frac{t^4}{4!} \mp \dots\right) \mathbb{1} + \left(t - \frac{t^3}{3!} + \frac{t^5}{5!} \mp \dots\right) H \\ &= \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}. \end{aligned}$$

(Siehe [4] für die Potenzreihendarstellungen von Cosinus und Sinus.) Da die Matrizen $a\mathbb{1}$ und bH kommutieren, gilt für die Matrix

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

die Gleichung

$$e^{tA} = e^{at} \begin{pmatrix} \cos(bt) & -\sin(bt) \\ \sin(bt) & \cos(bt) \end{pmatrix}$$

für alle $t \in \mathbb{R}$.

1.6 Komposition von Potenzreihen

Seien $(a_k)_{k \geq 1}$ und $(b_\ell)_{\ell \geq 0}$ zwei Folgen reeller Zahlen, so dass die Potenzreihen

$$P(z) := \sum_{k=1}^{\infty} a_k z^k, \quad Q(z) := \sum_{\ell=1}^{\infty} b_\ell z^\ell$$

positive Konvergenzradien haben, das heißt

$$\rho_P := \frac{1}{\limsup_{k \rightarrow \infty} |a_k|^{1/k}} > 0, \quad \rho_Q := \frac{1}{\limsup_{\ell \rightarrow \infty} |b_\ell|^{1/\ell}} > 0.$$

Insbesondere gilt $P(0) = 0$. Sei $0 < \rho \leq \rho_P$ so gewählt, dass

$$z \in \mathbb{C}, \quad |z| < \rho \quad \Longrightarrow \quad \sum_{k=1}^{\infty} |a_k| |z|^k < \rho_Q. \quad (1.6.1)$$

Ist $z \in \mathbb{C}$ mit Betrag $|z| < \rho \leq \rho_P$ so gilt $|P(z)| < \rho_Q$ nach (1.6.1). Damit ist die Komposition $Q \circ P : \{z \in \mathbb{C} \mid |z| < \rho\} \rightarrow \mathbb{C}$, wohldefiniert. Der folgende Satz besagt, dass sich diese Komposition als Potenzreihe darstellen läßt.

Satz 1.6.1. Für $k \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$ sei

$$c_k := \sum_{\ell=0}^k b_\ell \ell! \sum_{\substack{\sum_i m_i = \ell \\ \sum_i i m_i = k}} \prod_{\substack{i \in \mathbb{N} \\ m_i > 0}} \frac{a_i^{m_i}}{m_i!}. \quad (1.6.2)$$

Die zweite Summe in (1.6.2) ist über alle Folgen $(m_i)_{i \in \mathbb{N}}$ nichtnegativer ganzer Zahlen zu verstehen, die die Gleichungen $\sum_i i m_i = k$ und $\sum_i m_i = \ell$ erfüllen. Da $m_i \leq k$ ist für alle i und $m_i = 0$ ist für $i > k$, ist diese Summe endlich und verschwindet für $\ell > k$. Sei

$$R(z) := \sum_{k=0}^{\infty} c_k z^k \quad (1.6.3)$$

die Potenzreihe mit den Koeffizienten c_k . Dann gilt folgendes.

(i) Der Konvergenzradius $\rho_R := (\limsup_{k \rightarrow \infty} |c_k|^{1/k})^{-1}$ der Potenzreihe R ist grösser oder gleich ρ und $R(z) = Q(P(z))$ für alle $z \in \mathbb{C}$ mit $|z| < \rho$.

(ii) Sei \mathcal{A} eine reelle Banachalgebra. Dann erfüllt jedes $x \in \mathcal{A}$ mit $\|x\| < \rho$ die Ungleichung $\|P_{\mathcal{A}}(x)\| < \rho_Q$ und die Gleichung

$$R_{\mathcal{A}}(x) = Q_{\mathcal{A}}(P_{\mathcal{A}}(x)). \quad (1.6.4)$$

Beweis. Für $k \in \mathbb{N}_0$ und $N \in \mathbb{N}$ seien die reellen Koeffizienten $c_{k,N}$ durch

$$c_{k,N} := \sum_{\ell=0}^N b_\ell \ell! \sum_{\substack{m_1+\dots+m_N=\ell \\ \sum_{i=1}^N i m_i=k}} \frac{a_1^{m_1} a_2^{m_2} \dots a_N^{m_N}}{m_1! \dots m_N!}$$

gegeben. Dann gilt $c_{k,N} = c_k$ für $k \leq N$ und $c_{k,N} = 0$ für $k > N^2$ (weil jedes N -Tupel $(m_1, \dots, m_N) \in \mathbb{N}_0^N$ mit $\sum_i m_i = \ell \leq N$ die Ungleichung $k = \sum_i i m_i \leq N \sum_i m_i = N\ell \leq N^2$ erfüllt). Insbesondere ist die Funktion

$$R_N(z) := \sum_{k=0}^{\infty} c_{k,N} z^k$$

ein Polynom. Für $N \in \mathbb{N}$ seien die Polynome P_N und Q_N durch

$$P_N(z) := \sum_{k=1}^N a_k z^k, \quad Q_N(z) := \sum_{\ell=0}^N b_\ell z^\ell$$

für $z \in \mathbb{C}$ definiert. Dann gilt für alle $z \in \mathbb{C}$ die Gleichung

$$\begin{aligned} R_N(z) &= \sum_{k=0}^{\infty} c_{k,N} z^k \\ &= \sum_{k=0}^{\infty} \sum_{\ell=0}^N b_\ell \ell! \sum_{\substack{m_1+\dots+m_N=\ell \\ \sum_{i=1}^N i m_i=k}} \frac{a_1^{m_1} a_2^{m_2} \dots a_N^{m_N}}{m_1! \dots m_N!} z^k \\ &= \sum_{\ell=0}^N b_\ell \sum_{m_1+\dots+m_N=\ell} \frac{\ell!}{m_1! \dots m_N!} (a_1 z)^{m_1} (a_2 z^2)^{m_2} \dots (a_N z^N)^{m_N} \\ &= \sum_{\ell=0}^N b_\ell (a_1 z + a_2 z^2 + \dots + a_N z^N)^\ell \\ &= Q_N(P_N(z)). \end{aligned}$$

Sei nun $0 < r < \rho$. Dann gilt $s := \sum_{k=1}^{\infty} |a_k| r^k < \rho_Q$. Sei außerdem $z \in \mathbb{C}$ mit $|z| \leq r$. Dann konvergiert $P_N(z)$ gegen $P(z)$ und es ist $|P_N(z)| \leq s$ für alle $N \in \mathbb{N}$. Da Q_N auf dem Ball $\{z \in \mathbb{C} \mid |z| \leq s\}$ gleichmäßig gegen Q konvergiert, folgt daraus, dass $Q_N(P_N(z))$ gegen $Q(P(z))$ konvergiert, und daher gilt

$$\lim_{N \rightarrow \infty} R_N(z) = Q(P(z)) \quad \text{für alle } z \in \mathbb{C} \text{ mit } |z| \leq r. \quad (1.6.5)$$

Ersetzen wir die Koeffizienten a_k und b_ℓ durch ihre Beträge, so erhalten wir

$$|c_k| \leq \bar{c}_k := \sum_{\ell=0}^k |b_\ell| \ell! \sum_{\substack{\sum_i m_i = \ell \\ \sum_i i m_i = k}} \prod_{\substack{i \in \mathbb{N} \\ m_i > 0}} \frac{a_i^{m_i}}{m_i!},$$

$$|c_{k,N}| \leq \bar{c}_{k,N} := \sum_{\ell=0}^N |b_\ell| \ell! \sum_{\substack{m_1 + \dots + m_N = \ell \\ \sum_{i=1}^N i m_i = k}} \frac{|a_1|^{m_1} |a_2|^{m_2} \dots |a_N|^{m_N}}{m_1! \dots m_N!}.$$

Da $\bar{c}_{k,N} \leq \bar{c}_k$ ist für alle k, N sowie $\bar{c}_{k,N} = \bar{c}_k$ ist für alle $k \leq N$ und $\bar{c}_{k,N} = 0$ ist für alle $k > N^2$, folgt daraus, für jedes $N' \in \mathbb{N}$,

$$N' \leq \sqrt{N} \quad \Longrightarrow \quad \sum_{k=0}^{\infty} \bar{c}_{k,N'} r^k \leq \sum_{k=0}^N \bar{c}_k r^k \leq \sum_{k=0}^{\infty} \bar{c}_{k,N} r^k. \quad (1.6.6)$$

Hier konvergieren die beiden äusseren Terme gegen

$$\sum_{\ell=0}^{\infty} |b_\ell| \left(\sum_{k=1}^{\infty} |a_k| r^k \right)^\ell = \sum_{\ell=0}^{\infty} |b_\ell| s^\ell$$

für $N, N' \rightarrow \infty$, nach dem gleichen Argument wie oben. Also konvergiert die Differenz $\sum_{k>N} \bar{c}_{k,N} r^k$ der letzten beiden Terme in (1.6.6) gegen Null, und für $z \in \mathbb{C}$ mit $|z| \leq r$ gilt daher

$$\left| R_N(z) - \sum_{k=0}^N c_k z^k \right| = \left| \sum_{k=N+1}^{\infty} c_k z^k \right| \leq \sum_{k=N+1}^{\infty} \bar{c}_{k,N} r^k \xrightarrow{N \rightarrow \infty} 0. \quad (1.6.7)$$

Es folgt aus (1.6.5) und (1.6.7), dass die Gleichung

$$R(z) = \sum_{k=0}^{\infty} c_k z^k = \lim_{N \rightarrow \infty} \sum_{k=0}^N c_k z^k = \lim_{N \rightarrow \infty} R_N(z) = Q(P(z))$$

für jedes $z \in \mathbb{C}$ mit $|z| \leq r$ gilt. Da r als beliebige reelle Zahl $0 < r < \rho$ gewählt war, folgt daraus, dass der Konvergenzradius der Potenzreihe R die Ungleichung $\rho_R \geq \rho$ erfüllt und dass $R(z) = Q(P(z))$ ist für jedes $z \in \mathbb{C}$ mit $|z| < \rho$. Damit ist (i) bewiesen. Teil (ii) folgt aus dem gleichen Argument, indem man $z \in \mathbb{C}$ durch $x \in \mathcal{A}$ und den Betrag einer komplexen Zahl durch die Norm in \mathcal{A} ersetzt. Damit ist Satz 1.6.1 bewiesen. \square

Beispiel 1.6.2 (Logarithmus). Die Reihe

$$L(z) := \sum_{k=1}^{\infty} \frac{(-1)^{k-1} z^k}{k} = z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} \pm \dots$$

hat den Konvergenzradius $\rho_L = 1$ und stellt die Funktion

$$L(z) = \log(1+z)$$

für $z \in \mathbb{C}$ mit $|z| < 1$ dar. Das heißt, L erfüllt die Gleichung

$$\exp(L(z)) = 1+z \quad \text{für alle } z \in \mathbb{C} \text{ mit } |z| < 1.$$

(Siehe [2, Seite 44, Übung 1.7.18].) Nach Teil (i) von Satz 1.6.1 folgt daraus zunächst, dass die Koeffizienten $c_0 = c_1 = 1$ und $c_k = 0$ für $k \geq 2$ der Potenzreihe $1+z$ durch die Formel (1.6.2) mit $a_i = (-1)^{i-1}/i$ und $b_\ell = 1/\ell!$ gegeben sind. Das heißt, für alle $k \in \mathbb{N}_0$ gilt

$$\sum_{\ell=0}^k (-1)^{k-\ell} \sum_{\substack{\sum_i m_i = \ell \\ \sum_i i m_i = k}} \prod_{\substack{i \in \mathbb{N} \\ m_i > 0}} \frac{1}{m_i! i^{m_i}} = \begin{cases} 1, & \text{falls } k = 0, 1, \\ 0, & \text{falls } k \geq 2. \end{cases} \quad (1.6.8)$$

Weiterhin folgt aus Teil (ii) von Satz 1.6.1 dass, wenn \mathcal{A} eine Banachalgebra ist, jedes $x \in \mathcal{A}$ mit $\|x\| < 1$ die Gleichung

$$\exp\left(\sum_{k=1}^{\infty} \frac{(-1)^{k-1} x^k}{k}\right) = \mathbb{1} + x \quad (1.6.9)$$

erfüllt. Insbesondere gilt die Gleichung (1.6.9) für jede Matrix $x \in \mathbb{R}^{n \times n}$ mit $\|x\| < 1$, wobei $\|\cdot\|$ eine beliebige Matrixnorm auf $\mathbb{R}^{n \times n}$ ist.

Schlussbemerkungen

Der Satz über die Existenz und Eindeutigkeit von Lösungen gewöhnlicher Differentialgleichungen wird in Kapitel 4 in viel grösserer Allgemeinheit (auch für nichtlineare Differentialgleichungen) bewiesen. Formeln wie in den Beispielen im Abschnitt 1.5 erlauben es uns im Prinzip, ein lineares Differentialgleichungssystem beliebiger Ordnung explizit zu lösen (indem wir es zunächst in ein Differentialgleichungssystem erster Ordnung verwandeln, anschließend

die Eigenwert-Zerlegung, das heißt die Jordansche Normalform, der Matrix A bestimmen, und schließlich die Exponentialmatrix der Jordanschen Normalform ausrechnen). Bei nichtlinearen Differentialgleichungen ist es aber nur in den seltensten Fällen möglich, explizite Lösungen hinzuschreiben, und man kann oft bestenfalls hoffen, etwas über das *qualitative Verhalten* der Lösungen aussagen zu können. Dies führt hin zur qualitativen Theorie **dynamischer Systeme**, einem ganz eigenen und spannenden Gebiet der mathematischen Forschung.

Ein gänzlich anderes Thema, das hier kurz angeschnitten wurde, ist das der Banachräume und der linearen Operatoren. Diese Begriffe stehen am Anfang der **Funktionalanalysis**, einem wichtigen Gebiet der Mathematik, das erst im 5. Semester vertieft wird, und in weiten Teilen sowohl der Mathematik (Partielle Differentialgleichungen bis hin zur Topologie und Geometrie) als auch der Physik (zum Beispiel Quantenmechanik) eine zentrale Rolle spielt.

Schließlich haben wir hier einen kurzen Einblick in die Banachalgebren erhalten, die im Schnittpunkt zwischen Analysis, Topologie und Algebra stehen, und auch ein ganz eigenes Gebiet der mathematischen Forschung sind.

Kapitel 2

Differenzierbare Abbildungen

Dieses Kapitel beginnt in Abschnitt 2.1 mit einer Einführung in den Begriff der Differenzierbarkeit für Funktionen von mehreren reellen Variablen, oder allgemeiner für Abbildungen zwischen endlichdimensionalen Banachräumen. Das Thema des Abschnitts 2.3 sind die Grundlegenden Rechenregeln für die Ableitung, wie die Leibnizregel und die Kettenregel in höheren Dimensionen. Abschnitt 2.4 behandelt den Schrankensatz (eine Art höherdimensionaler Mittelwertsatz) und einige seiner Konsequenzen. Anschließend betrachten wir höhere Ableitungen (Abschnitt 2.5), diskutieren Taylor-Reihen (Abschnitt 2.6), und untersuchen lokale Extrema (Abschnitt 2.7).

2.1 Grundbegriffe

Zu Beginn dieser einführenden Betrachtungen sei daran erinnert, dass eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ an einer Stelle $a \in \mathbb{R}$ differenzierbar genannt wird, wenn der Grenzwert

$$A := \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} \quad (2.1.1)$$

existiert. In präziser mathematischer Formulierung ist dies die Aussage

$$\exists A \in \mathbb{R} \forall \varepsilon > 0 \exists \delta > 0 \forall h \in \mathbb{R} \left(0 < |h| < \delta \implies \left| \frac{f(a+h) - f(a)}{h} - A \right| < \varepsilon \right). \quad (2.1.2)$$

Mit anderen Worten, es existiert eine reelle Zahl A , so dass für jede noch so kleine reelle Zahl $\varepsilon > 0$ eine reelle Zahl $\delta > 0$ existiert, so dass jede reelle Zahl h mit $0 < |h| < \delta$ die Ungleichung $\left| A - \frac{f(a+h) - f(a)}{h} \right| < \varepsilon$ erfüllt.

Die Frage, um die es in diesem Abschnitt gehen soll ist, wie man diesen Begriff der Differenzierbarkeit aus der Analysis I auf Funktionen erweitern kann, die auf einer offenen Teilmenge eines endlichdimensionalen Vektorraumes definiert sind. In der Tat gibt es drei solche Verallgemeinerungen, die wir im folgenden diskutieren werden.

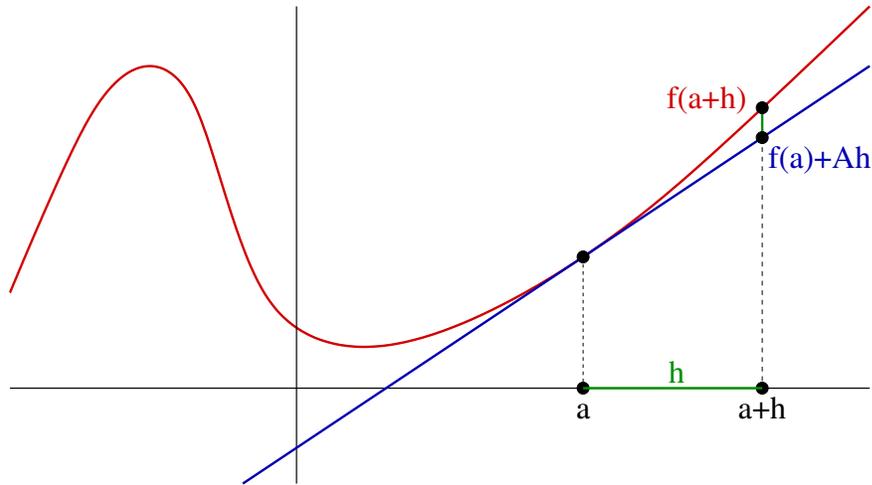


Abbildung 2.1: Die Ableitung als lineare Approximation.

Zunächst ist zu beachten, dass man durch Vektoren nicht dividieren kann, so dass sich die rechte Seite in (2.1.1) nicht auf Funktionen auf Vektorräumen übertragen läßt. Jedoch können wir die Ungleichung in (2.1.2) so umformen, dass sie auch für Elemente eines normierten Vektorraumes Sinn macht:

$$\begin{aligned} \left| \frac{f(a+h) - f(a)}{h} - A \right| &= \left| \frac{f(a+h) - f(a) - Ah}{h} \right| \\ &= \frac{|f(a+h) - f(a) - Ah|}{|h|}. \end{aligned} \quad (2.1.3)$$

Die Bedingung (2.1.2) für die Differenzierbarkeit sagt nun, dass dieser Ausdruck beliebig klein wird für hinreichend kleine $|h|$. Dies hat die geometrische Bedeutung, dass die Funktion $h \mapsto f(a) + Ah$ als Tangente an den Graphen von f an der Stelle a interpretiert werden kann, und eine gute erste Näherung an die Funktion $h \mapsto f(a+h)$ darstellt, in dem Sinne, dass die Differenz dieser beiden Funktionen für kleine Werte von h klein ist im Vergleich zu $|h|$ (siehe Abbildung 2.1).

Der Begriff der Differenzierbarkeit

Suchen wir für Funktionen auf offenen Teilmengen eines normierten Vektorraumes nach einer analogen Näherung, so ist die Zahl $A = f'(a)$ im eindimensionalen Fall zu ersetzen durch eine lineare Abbildung. Für endlichdimensionale Vektorräume X, Y erinnern wir an die Bezeichnung $\mathcal{L}(X, Y)$ für die Menge aller linearen Abbildungen von X nach Y .

Definition 2.1.1 (Differenzierbarkeit).

Seien $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ endlichdimensionale normierte Vektorräume, sei $U \subset X$ eine offene Menge, sei $f : U \rightarrow Y$ eine Abbildung, und sei $a \in U$. Die Abbildung f heißt **differenzierbar an der Stelle a** , wenn eine lineare Abbildung $A \in \mathcal{L}(X, Y)$ existiert, so dass für jede reelle Zahl $\varepsilon > 0$, eine reelle Zahl $\delta > 0$ existiert, so dass für alle $h \in X$ folgende Aussage gilt:

$$0 < \|h\|_X < \delta \quad \Longrightarrow \quad a + h \in U \text{ und } \|f(a+h) - f(a) - Ah\|_Y < \varepsilon \|h\|_X. \quad (2.1.4)$$

Als logische Formel geschrieben heißt das

$$\forall \varepsilon > 0 \exists \delta > 0 \forall h \in X \left(0 < \|h\|_X < \delta \Longrightarrow a + x \in U \text{ und } \frac{\|f(a+h) - f(a) - Ah\|_Y}{\|h\|_X} < \varepsilon \right). \quad (2.1.5)$$

Zu dieser Definition sei angemerkt, dass die aus der Analysis I bekannte Definition der Differenzierbarkeit von Funktionen einer reellen Variablen als der Spezialfall $X = Y = \mathbb{R}$ auftritt, denn in diesem Fall ist eine lineare Abbildung von X nach Y stets durch Multiplikation mit einer reellen Zahl A gegeben. Zweitens sei daran erinnert, dass eine lineare Abbildung $A : X \rightarrow Y$ zwischen endlichdimensionalen normierten Vektorräumen nach Korollar 1.2.2 stets beschränkt ist, was heißt, dass ihre Operatornorm

$$\|A\| = \sup_{0 \neq x \in X} \frac{\|Ax\|_Y}{\|x\|_X}$$

endlich ist. Damit läßt sich die Definition 2.1.1 wörtlich auf den Fall übertragen, dass X und Y beliebige Banachräume sind; es ist lediglich darauf zu achten, dass $\mathcal{L}(X, Y)$ dann den Raum aller *beschränkten* linearen Operatoren von X nach Y bezeichnet (Definition 1.2.1). Drittens sei noch einmal betont, dass in der Definition der Differenzierbarkeit die *Existenz* einer linearen Abbildung $A : X \rightarrow Y$ verlangt wird, die der Bedingung (2.1.5) genügt. Dies wirft die Frage auf, ob *nur ein* solches A existiert. Lemma 2.1.3 gibt eine positive Antwort auf diese Frage.

Übung 2.1.2. Seien X und Y endlichdimensionale normierte Vektorräume, sei $U \subset X$ eine offene Menge, sei $f : U \rightarrow Y$ eine Abbildung, und sei $a \in U$. Dann hängt die Aussage “ f ist differenzierbar an der Stelle a ” nicht von der Wahl der Normen auf X und Y ab. **Hinweis:** Satz 1.1.8.

Lemma 2.1.3. Seien X, Y, U, f, a wie in Definition 2.1.1. Dann existiert höchstens eine lineare Abbildung $A \in \mathcal{L}(X, Y)$, die (2.1.5) erfüllt.

Beweis. Seien $A, B : X \rightarrow Y$ zwei lineare Abbildungen, die (2.1.5) erfüllen, und sei $\varepsilon > 0$ gegeben. Dann existiert nach Voraussetzung ein $\delta > 0$ mit

$$B_\delta(a) = \{x \in X \mid \|x - a\|_X < \delta\} \subset U,$$

so dass jeder Vektor $h \in X$ mit $0 < \|h\|_X < \delta$ die Ungleichungen

$$\begin{aligned} \|f(a+h) - f(a) - Ah\|_Y &< \varepsilon \|h\|_X, \\ \|f(a+h) - f(a) - Bh\|_Y &< \varepsilon \|h\|_X \end{aligned}$$

erfüllt. Für $h \in X$ mit $0 < \|h\|_X < \delta$ folgt daraus

$$\begin{aligned} \|Ah - Bh\|_Y &= \|Ah + f(a) - f(a+h) + f(a+h) - f(a) - Bh\|_Y \\ &\leq \|f(a+h) - f(a) - Ah\|_Y + \|f(a+h) - f(a) - Bh\|_Y \\ &< 2\varepsilon \|h\|_X. \end{aligned}$$

Damit erhalten wir die Ungleichung $\|h\|_X^{-1} \|(A - B)h\|_Y < 2\varepsilon$ für alle $h \in X$ mit $0 < \|h\|_X < \delta$ und deshalb durch reskalieren sogar für alle $h \in X \setminus \{0\}$. Bilden wir nun das Supremum über alle $h \in X \setminus \{0\}$, so erhalten wir

$$\|A - B\| = \sup_{0 \neq h \in X} \frac{\|(A - B)h\|_Y}{\|h\|_X} \leq 2\varepsilon$$

für jedes $\varepsilon > 0$. Daraus folgt $\|A - B\| = 0$ und daher $A = B$. Damit ist Lemma 2.1.3 bewiesen. \square

Definition 2.1.4 (Ableitung). Seien X, Y endlichdimensionale normierte Vektorräume, sei $U \subset X$ eine offene Menge, und sei $f : U \rightarrow Y$ eine an der Stelle $a \in U$ differenzierbare Abbildung. Dann wird die eindeutige lineare Abbildung $A : X \rightarrow Y$, die (2.1.5) erfüllt, die **Ableitung von f an der Stelle a** genannt und mit $df(a) := A : X \rightarrow Y$ bezeichnet.¹

¹Andere in der Literatur übliche Bezeichnungen für die Ableitung sind $Df(a)$, $D_a f$, $d_a f$, $T_a f$, $f'(a)$. In diesem Manuskript wird in der Regel die Notation $df(a)$ verwendet.

Lemma 2.1.5 (Differenzierbare Abbildungen sind stetig).

Seien X, Y endlichdimensionale normierte Vektorräume, sei $U \subset X$ eine offene Menge, und sei $f : U \rightarrow Y$ eine an der Stelle $a \in U$ differenzierbare Abbildung. Dann ist f an der Stelle a stetig.

Beweis. Sei $A := df(a) \in \mathcal{L}(X, Y)$ und sei $\varepsilon > 0$ gegeben. Wähle $\delta > 0$ so dass (2.1.4) gilt und definiere

$$\rho := \min \left\{ \delta, \frac{\varepsilon}{2\|A\|}, \frac{1}{2} \right\} > 0.$$

Nun sei $h \in X$ mit $\|h\|_X < \rho \leq \delta$. Nach (2.1.4) ist dann $a + h \in U$ und erfüllt die Ungleichung

$$\begin{aligned} \|f(a+h) - f(a)\|_Y &\leq \|f(a+h) - f(a) - Ah\|_Y + \|Ah\|_Y \\ &\leq \varepsilon \|h\|_X + \|A\| \|h\|_X \\ &< \varepsilon \rho + \|A\| \rho \\ &\leq \varepsilon. \end{aligned}$$

Damit ist Lemma 2.1.5 bewiesen. \square

Beispiel 2.1.6. Sind X, Y endlichdimensionale Vektorräume und ist $U \subset X$ eine offene Menge, so ist jede konstante Abbildung $f : U \rightarrow Y$ an jeder Stelle $a \in U$ differenzierbar und hat die Ableitung $f'(a) = 0$. Die Umkehrung gilt nur für zusammenhängende offene Mengen (Anhang B) und wird im Abschnitt 2.4 bewiesen.

Beispiel 2.1.7. Seien eine Matrix $A \in \mathbb{R}^{m \times n}$ und ein Vektor $b \in \mathbb{R}^m$ gegeben. Diese bestimmen eine *affine Abbildung* $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ durch

$$f(x) := Ax + b \quad \text{für } x \in \mathbb{R}^n.$$

In diesem Fall gilt $f(a+h) - f(a) - Ah = 0$ für alle $a, h \in \mathbb{R}^n$. Daher ist f an jeder Stelle $a \in \mathbb{R}^n$ differenzierbar und hat dort die Ableitung

$$df(a) = A.$$

Hier ist hier anzumerken, dass wir die Matrix $A \in \mathbb{R}^{m \times n}$ mit der linearen Abbildung $\mathbb{R}^n \rightarrow \mathbb{R}^m : \xi \mapsto A\xi$ identifizieren (Multiplikation von Matrix und Vektor). Eine solche Identifikation können wir immer dann vornehmen, wenn unsere Vektorräume X und Y kanonische Basen besitzen. Wir werden jedoch auch Beispiele sehen, in denen das nicht der Fall ist. Es bleibt das Grundprinzip bestehen, das besagt “die Ableitung ist eine lineare Abbildung” (die oft, aber nicht immer, mit einer Matrix identifiziert werden kann).

Beispiel 2.1.8. Sei $Q = Q^T \in \mathbb{R}^{n \times n}$ eine symmetrische Matrix und definiere die Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}$ durch

$$f(x) := \frac{1}{2} x^T Q x = \frac{1}{2} \langle x, Qx \rangle \quad \text{für } x \in \mathbb{R}^n.$$

Hier ist x^T der aus x durch Transposition gewonnenen Zeilenvektor und

$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i$$

für $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ und $y = (y_1, \dots, y_n) \in \mathbb{R}^n$ das standard innere Produkt auf dem \mathbb{R}^n . Für $a, h \in \mathbb{R}^n$ gilt

$$\begin{aligned} f(a+h) - f(a) &= \frac{1}{2} (a+h)^T Q (a+h) - \frac{1}{2} a^T Q a \\ &= \frac{1}{2} a^T Q h + \frac{1}{2} h^T Q a + \frac{1}{2} h^T Q h \\ &= a^T Q h + \frac{1}{2} h^T Q h \end{aligned}$$

und daher, nach der Cauchy–Schwarz-Ungleichung,

$$\begin{aligned} |f(a+h) - f(a) - a^T Q h| &= \frac{1}{2} |h^T Q h| \\ &= \frac{1}{2} |\langle h, Qh \rangle| \\ &\leq \frac{1}{2} \|h\| \|Qh\| \\ &\leq \frac{\|Q\| \|h\|}{2} \|h\|. \end{aligned}$$

Ist also $\varepsilon > 0$ gegeben, so erfüllt

$$\delta := \frac{2\varepsilon}{\|Q\|} > 0$$

die Bedingung (2.1.4) mit

$$A := a^T Q \in \mathbb{R}^{1 \times n}$$

(verstanden als lineare Abbildung von \mathbb{R}^n nach \mathbb{R}). Die Abbildung f ist daher differenzierbar mit der Ableitung

$$df(a) = a^T Q$$

für alle $a \in \mathbb{R}^n$.

Richtungsableitungen

Wir befassen uns nun mit den am Anfang bereits angekündigten weiteren zwei Möglichkeiten, die aus der Analysis I bekannte Definition der Differenzierbarkeit auf Funktionen von mehreren Variablen zu erweitern. Dies sind die Richtungsableitung und die partiellen Ableitungen.

Definition 2.1.9 (Richtungsableitung). *Seien X, Y endlichdimensionale normierte Vektorräume, sei $U \subset X$ eine offene Menge, und seien eine Abbildung $f : U \rightarrow Y$ und Elemente $a \in U$ und $\xi \in X$ gegeben. Die **Richtungsableitung von f an der Stelle a in Richtung ξ** ist der Grenzwert*

$$\partial_\xi f(a) := \lim_{t \rightarrow 0} \frac{f(a + t\xi) - f(a)}{t} \in Y, \quad (2.1.6)$$

falls dieser existiert.

Wir betrachten nun den Spezialfall $X = \mathbb{R}^n$ und $Y = \mathbb{R}^m$ und bezeichnen mit $e_1, \dots, e_n \in \mathbb{R}^n$ die Standardbasis des \mathbb{R}^n , das heißt

$$e_i := (0, \dots, 0, 1, 0, \dots, 0), \quad i = 1, \dots, n,$$

(n Einträge mit der 1 an der i ten Stelle).

Definition 2.1.10 (Partielle Ableitung). *Sei $U \subset \mathbb{R}^n$ eine offene Menge und seien eine Abbildung $f : U \rightarrow \mathbb{R}^m$ und ein Element $a \in U$ gegeben. Die Abbildung f heißt **an der Stelle a partiell differenzierbar**, wenn die Richtungsableitungen*

$$\begin{aligned} \partial_i f(a) &:= \partial_{e_i} f(a) \\ &:= \frac{\partial f}{\partial x_i}(a) \\ &= \lim_{t \rightarrow 0} \frac{f(a + te_i) - f(a)}{t} \\ &= \left. \frac{d}{dt} \right|_{t=0} f(a_1, \dots, a_{i-1}, a_i + t, a_{i+1}, \dots, a_n) \end{aligned} \quad (2.1.7)$$

für $i = 1, \dots, n$ existieren. In diesem Fall werden die Vektoren $\partial_i f(a) \in \mathbb{R}^m$ die **partiellen Ableitungen von f an der Stelle a** genannt.

Lemma 2.1.11. *Seien X, Y endlichdimensionale normierte Vektorräume, sei $U \subset X$ eine offene Menge, und sei $f : U \rightarrow Y$ differenzierbar an der Stelle $a \in U$. Dann existiert die Richtungsableitung $\partial_\xi f(a)$ für jeden Vektor $\xi \in X$ und ist durch*

$$\partial_\xi f(a) = df(a)\xi \quad (2.1.8)$$

gegeben.

Beweis. Sei $A := df(a) : X \rightarrow Y$ und sei $\xi \in X$ gegeben. Für $\xi = 0$ folgt die Existenz der Richtungsableitung ebenso wie die Gleichung $\partial_\xi f(a) = 0 = A\xi$ unmittelbar aus der Definition. Nehmen wir also an, dass $\xi \neq 0$ ist. Zu zeigen ist dann die Aussage

$$\lim_{t \rightarrow 0} \frac{f(a + t\xi) - f(a)}{t} = A\xi. \quad (2.1.9)$$

Sei also $\varepsilon > 0$ gegeben und sei $\tilde{\varepsilon} := \varepsilon / \|\xi\|_X$. Dann existiert nach Definition 2.1.1 eine Zahl $\tilde{\delta} > 0$, so dass jeder Vektor $h \in X$ mit $0 < \|h\|_X < \tilde{\delta}$ die Bedingung $a + h \in U$ und die Ungleichung

$$\|f(a + h) - f(a) - Ah\|_Y < \tilde{\varepsilon} \|h\|_X$$

erfüllt. Nun sei $\delta := \tilde{\delta} / \|\xi\|_X$ und sei $t \in \mathbb{R}$ mit $0 < |t| < \delta$. Dann gilt

$$0 < \|t\xi\|_X = |t| \|\xi\|_X < \delta \|\xi\|_X = \tilde{\delta},$$

daher ist $a + t\xi \in U$ und

$$\|f(a + t\xi) - f(a) - tA\xi\|_Y < \tilde{\varepsilon} \|t\xi\|_X = |t| \varepsilon,$$

und daraus folgt

$$\left\| \frac{f(a + t\xi) - f(a)}{t} - A\xi \right\|_Y < \varepsilon.$$

Damit sind (2.1.9) und Lemma 2.1.11 bewiesen. \square

Lemma 2.1.11 zeigt, dass aus der Differenzierbarkeit die Existenz der Richtungsableitungen folgt, und aus der Existenz der Richtungsableitungen folgt offensichtlich die partielle Differenzierbarkeit. Ebenso wissen wir dass aus der Differenzierbarkeit die Stetigkeit folgt (Lemma 2.1.5). Die folgenden Beispiele zeigen, dass sich keine dieser Implikationen umkehren läßt. Auch wissen wir bereits aus der Analysis I, dass aus der Stetigkeit nicht einmal die partielle Differenzierbarkeit folgt. (Die stetige Funktion $\mathbb{R}^n \rightarrow \mathbb{R} : x \mapsto \|x\|$ ist an der Stelle $a = 0$ nicht partiell differenzierbar.)

Beispiel 2.1.12. Dieses Beispiel zeigt, dass die Existenz der Richtungsableitungen nicht die Differenzierbarkeit impliziert. Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ die Funktion

$$f(x, y) := \frac{x^2 y}{x^2 + y^2} \quad \text{für } (x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$$

mit $f(0, 0) := 0$. Diese Funktion erfüllt die Bedingung

$$f(t\xi, t\eta) = tf(\xi, \eta)$$

für alle $\zeta = (\xi, \eta) \in \mathbb{R}^2$ und alle $t \in \mathbb{R}$. Daher existieren alle Richtungsableitungen von f an der Stelle $a = 0 \in \mathbb{R}^2$ und es gilt $\partial_\zeta f(0) = f(\zeta)$ für $\zeta \in \mathbb{R}^2$. Da die Abbildung $\mathbb{R}^2 \rightarrow \mathbb{R} : \zeta \mapsto \partial_\zeta f(0) = f(\zeta)$ nicht linear ist, kann f nach Lemma 2.1.11 an der Stelle $a = 0$ nicht differenzierbar sein.

Beispiel 2.1.13. Dieses Beispiel zeigt, dass aus der Existenz der Richtungsableitungen nicht einmal die Stetigkeit folgt. Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ die Funktion

$$f(x, y) := \frac{x^2 + y^2}{y} \quad \text{falls } y \neq 0,$$

mit $f(x, 0) := 0$. Diese Funktion erfüllt die Bedingung $f(t\zeta) = tf(\zeta)$, für alle $\zeta \in \mathbb{R}^2$ und alle $t \in \mathbb{R}$, so dass die Richtungsableitungen von f an der Stelle $a = 0$ alle existieren. Jedoch ist $f(\varepsilon, \varepsilon^3) = \varepsilon^{-1} + \varepsilon^3 \geq \varepsilon^{-1}$ für alle $\varepsilon > 0$, und daher ist f an der Stelle $a = 0$ unstetig. Nach Lemma 2.1.3 ist f an der Stelle $a = 0$ auch nicht differenzierbar.

Beispiel 2.1.14. Dieses Beispiel zeigt, dass aus der partiellen Differenzierbarkeit nicht die Existenz der Richtungsableitungen folgt. Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ die Funktion

$$f(x, y) := \frac{xy}{\sqrt{x^2 + y^2}} \quad \text{für } (x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$$

mit $f(0, 0) := 0$. Diese Funktion erfüllt die Bedingung

$$f(t\xi, t\eta) = |t|f(\xi, \eta)$$

für alle $\zeta = (\xi, \eta) \in \mathbb{R}^2$ und alle $t \in \mathbb{R}$. Daher existiert die Richtungsableitung von f an der Stelle $a = 0 \in \mathbb{R}^2$ in Richtung von $\zeta = (\xi, \eta) \in \mathbb{R}^2$ genau dann wenn $f(\xi, \eta) = 0$ ist. Dies ist der Fall für die Basisvektoren $e_1 = (1, 0)$ und $e_2 = (0, 1)$. Damit ist f an der Stelle $a = 0$ partiell differenzierbar, jedoch existieren nicht alle Richtungsableitungen von f an der Stelle $a = 0$.

Übung: Die Funktion f ist stetig.

Die Jacobi-Matrix

Nehmen wir nun an, dass $U \subset \mathbb{R}^n$ eine offene Menge und $f : U \rightarrow \mathbb{R}^m$ eine an der Stelle $x_0 \in U$ differenzierbare Abbildung ist. Dann ist $df(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine lineare Abbildung. Wie wir aus der Linearen Algebra wissen, ist eine jede solche Abbildung durch Multiplikation mit einer Matrix $A \in \mathbb{R}^{m \times n}$ gegeben. Dies wirft die Frage auf, welche Matrix die Ableitung $df(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ unserer Funktion repräsentiert, und wie wir sie berechnen können. Zur Beantwortung dieser Frage benötigen wir das folgende Lemma.

Lemma 2.1.15. *Sei X ein endlichdimensionaler normierter Vektorraum, sei $U \subset X$ eine offene Menge, sei*

$$f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} : U \rightarrow \mathbb{R}^m \quad (2.1.10)$$

eine Abbildung, und sei $x_0 \in U$ und $\xi \in X$. Dann gilt folgendes.

(i) *Die Richtungsableitung $\partial_\xi f(x_0)$ existiert genau dann, wenn die Richtungsableitungen $\partial_\xi f_j(x_0)$ für $j = 1, \dots, m$ existieren. Ist dies der Fall, so gilt*

$$\partial_\xi f(x_0) = \begin{pmatrix} \partial_\xi f_1(x_0) \\ \vdots \\ \partial_\xi f_m(x_0) \end{pmatrix} \in \mathbb{R}^m. \quad (2.1.11)$$

(ii) *Die Abbildung $f : U \rightarrow \mathbb{R}^m$ ist genau dann an der Stelle x_0 differenzierbar, wenn die Funktionen $f_j : U \rightarrow \mathbb{R}$ für $j = 1, \dots, m$ an der Stelle x_0 differenzierbar sind. Ist dies der Fall, so gilt*

$$df(x_0) = \begin{pmatrix} df_1(x_0) \\ \vdots \\ df_m(x_0) \end{pmatrix} : \mathbb{R}^n \rightarrow \mathbb{R}^m. \quad (2.1.12)$$

Beweis. Wir beweisen Teil (i). Dazu nehmen wir zunächst an, dass die Richtungsableitung von f an der Stelle x_0 in Richtung ξ existiert, und wir bezeichnen diese Richtungsableitung mit

$$a := \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} := \partial_\xi f(x_0) \in \mathbb{R}^m.$$

Dann gilt

$$\left| \frac{f_j(x_0 + t\xi) - f_j(x_0)}{t} - a_j \right| \leq \left\| \frac{f(x_0 + t\xi) - f(x_0)}{t} - a \right\|_{\mathbb{R}^m}$$

für alle $t \in \mathbb{R} \setminus \{0\}$ und alle $j \in \{1, \dots, m\}$. Da die rechte Seite nach Voraussetzung gegen Null konvergiert für $t \rightarrow 0$, folgt daraus

$$\lim_{t \rightarrow 0} \left| \frac{f_j(x_0 + t\xi) - f_j(x_0)}{t} - a_j \right| = 0 \quad \text{für } j = 1, \dots, m.$$

Daher existiert für jedes j die Richtungsableitung von f_j an der Stelle x_0 in Richtung von ξ und es gilt

$$\partial_\xi f_j(x_0) = \lim_{t \rightarrow 0} \frac{f_j(x_0 + t\xi) - f_j(x_0)}{t} = a_j.$$

Nehmen wir umgekehrt an, dass die Richtungsableitung $a_j := \partial_\xi f_j(x_0)$ für alle $j = 1, \dots, m$ existiert, und definieren wir $a := (a_1, \dots, a_m) \in \mathbb{R}^m$, so gilt

$$\left\| \frac{f(x_0 + t\xi) - f(x_0)}{t} - a \right\|_{\mathbb{R}^m} = \sqrt{\sum_{i=1}^m \left| \frac{f_i(x_0 + t\xi) - f_i(x_0)}{t} - a_i \right|^2}.$$

Nach Voraussetzung konvergiert die rechte Seite dieser Gleichung gegen Null für $t \rightarrow 0$, und daraus folgt die Existenz der Richtungsableitung $\partial_\xi f(x_0) = a$. Damit ist Teil (i) bewiesen.

Wir beweisen Teil (ii) mit dem gleichen Argument. Ist f an der Stelle x_0 differenzierbar mit der Ableitung $A := df(x_0) : X \rightarrow \mathbb{R}^m$, und bezeichnen wir mit $A_j : X \rightarrow \mathbb{R}$ die Komposition von A mit der Projektion von \mathbb{R}^m auf die j te Koordinate, so folgt die Differenzierbarkeit von f_j an der Stelle x_0 mit $df_j(x_0) = A_j : X \rightarrow \mathbb{R}$ aus der Ungleichung

$$\frac{|f_j(x_0 + h) - f_j(x_0) - A_j h|}{\|h\|_X} \leq \frac{\|f(x_0 + h) - f(x_0) - Ah\|_{\mathbb{R}^m}}{\|h\|_X}.$$

Setzen wir umgekehrt voraus, dass f_j an der Stelle x_0 differenzierbar ist mit $A_j := df_j(x_0)$, und definieren wir die lineare Abbildung $A : X \rightarrow \mathbb{R}^m$ durch $A\xi := (A_1\xi, \dots, A_m\xi)$ für $\xi \in X$, so folgt die Differenzierbarkeit von f an der Stelle x_0 aus der Gleichung

$$\frac{\|f(x_0 + h) - f(x_0) - Ah\|_{\mathbb{R}^m}^2}{\|h\|_X^2} = \sum_{j=1}^m \left(\frac{|f_j(x_0 + h) - f_j(x_0) - A_j h|}{\|h\|_X} \right)^2.$$

Damit ist Lemma 2.1.15 bewiesen. \square

Lemma 2.1.16 (Jacobi-Matrix).

Sei $U \subset \mathbb{R}^n$ eine offene Menge und sei $f : U \rightarrow \mathbb{R}^m$ eine Abbildung mit den Koordinaten f_1, \dots, f_m , die an der Stelle $x_0 \in U$ differenzierbar ist. Dann ist die Ableitung $df(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ durch die Multiplikation mit der Matrix

$$Df(x_0) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \frac{\partial f_1}{\partial x_2}(x_0) & \cdots & \frac{\partial f_1}{\partial x_n}(x_0) \\ \frac{\partial f_2}{\partial x_1}(x_0) & \frac{\partial f_2}{\partial x_2}(x_0) & \cdots & \frac{\partial f_2}{\partial x_n}(x_0) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x_0) & \frac{\partial f_m}{\partial x_2}(x_0) & \cdots & \frac{\partial f_m}{\partial x_n}(x_0) \end{pmatrix} \in \mathbb{R}^{m \times n} \quad (2.1.13)$$

der partiellen Ableitungen gegeben.

Proof. Bezeichnen wir mit e_1, \dots, e_n die Standardbasis des \mathbb{R}^n , so erfüllt jeder Vektor $\xi = \sum_{i=1}^n \xi_i e_i \in \mathbb{R}^n$ die Gleichung

$$df(x_0)\xi = \sum_{i=1}^n \xi_i df(x_0)e_i = \sum_{i=1}^n \xi_i \partial_{e_i} f(x_0) = \sum_{i=1}^n \xi_i \frac{\partial f}{\partial x_i}(x_0) = Df(x_0)\xi.$$

Hier folgt die erste Gleichheit aus der Linearität von $df(x_0)$, die zweite aus Lemma 2.1.11, die dritte aus der Definition der partiellen Ableitungen, und die vierte aus Teil (i) von Lemma 2.1.15 sowie der Definition der Matrix $Df(x_0)$ in (2.1.13). Damit ist Lemma 2.1.16 bewiesen. \square

Definition 2.1.17. In der Situation von Lemma 2.1.16 wird die durch Gleichung (2.1.13) definierte Matrix $Df(x_0) \in \mathbb{R}^{m \times n}$ der partiellen Ableitungen von f die **Jacobi-Matrix von f an der Stelle $x_0 \in U$** genannt.

Wir merken an, dass das Symbol $Df(x_0)$ für die Jacobi-Matrix hier um der Klarheit willen verwendet wurde, zur formalen Unterscheidung zwischen der Ableitung als linearer Abbildung $df(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ und der Jacobi-Matrix der partiellen Ableitungen $Df(x_0) \in \mathbb{R}^{m \times n}$. Da Lemma 2.1.16 nun bewiesen ist, wird diese Unterscheidung jedoch hinfällig, und wir werden im folgenden zu unserem Standardsymbol $df(x_0)$ für die Ableitung zurückkehren, sei es als Jacobi-Matrix, oder als lineare Abbildung. Denn es wird im allgemeinen nicht nötig sein, zwischen einer linearen Abbildung von \mathbb{R}^n nach \mathbb{R}^m und der sie repräsentierenden Matrix in der Notation zu unterscheiden.

2.2 Stetige Differenzierbarkeit

In Abschnitt 2.1 haben wir gesehen, dass jede differenzierbare Funktion auf einer offenen Teilmenge des \mathbb{R}^n partiell differenzierbar ist, während die Umkehrung nicht gilt. Setzt man aber voraus, dass die partiellen Ableitungen überall existieren und zudem stetig sind, so zeigt der folgende Satz, dass unsere Funktion dann auch differenzierbar ist.

Satz 2.2.1. *Sei $U \subset \mathbb{R}^n$ eine offene Teilmenge und sei $x_0 \in U$ gegeben. Weiterhin sei $f : U \rightarrow \mathbb{R}^m$ eine Abbildung, die an jeder Stelle in U partiell differenzierbar ist, und deren partielle Ableitungen $\partial_i f : U \rightarrow \mathbb{R}^m$ für $i = 1, \dots, n$ an der Stelle x_0 stetig sind. Dann ist f an der Stelle x_0 differenzierbar.*

Definition 2.2.2. (i) *Sei $U \subset \mathbb{R}^n$ eine offene Teilmenge. Eine Abbildung $f : U \rightarrow \mathbb{R}^m$ heißt **stetig differenzierbar** wenn sie partiell differenzierbar ist und ihre partiellen Ableitungen $\partial_i f : U \rightarrow \mathbb{R}^m$ für $i = 1, \dots, n$ stetig sind.*
(ii) *Seien X, Y endlichdimensionale normierte Vektorräume und sei $U \subset X$ eine offene Menge. Eine Abbildung $f : U \rightarrow Y$ heißt **stetig differenzierbar** wenn sie differenzierbar und ihre Ableitung $df : U \rightarrow \mathcal{L}(X, Y)$ stetig ist.*

Übung 2.2.3. Für $X = \mathbb{R}^n$ und $Y = \mathbb{R}^m$ sind die beiden Definitionen in Definition 2.2.2 zueinander äquivalent.

Beweis von Satz 2.2.1. Nach Lemma 2.1.15 genügt es, den Satz für $m = 1$ zu beweisen. Sei also $f : U \rightarrow \mathbb{R}$ eine partiell differenzierbare Funktion, deren partielle Ableitungen alle an der Stelle x_0 stetig sind. Zu zeigen ist, dass f an der Stelle x_0 differenzierbar ist. Dazu verwenden wir die Euklidische Norm

$$\|x\| := \sqrt{x_1^2 + \dots + x_n^2} \quad \text{für } x = (x_1, \dots, x_n) \in \mathbb{R}^n,$$

und bezeichnen mit e_1, \dots, e_n die Standardbasis des \mathbb{R}^n .

Sei $\varepsilon > 0$ gegeben. Da die Funktionen $\partial_i f : U \rightarrow \mathbb{R}$ an der Stelle x_0 stetig sind, existiert eine Konstante $\delta > 0$, so dass für alle $h \in \mathbb{R}^n$ folgendes gilt:

$$\|h\| < \delta \quad \implies \quad \begin{aligned} &x_0 + h \in U \text{ und, für } i = 1, \dots, n, \\ &|\partial_i f(x_0 + h) - \partial_i f(x_0)| < \varepsilon/n. \end{aligned} \quad (2.2.1)$$

Sei ein Vektor $h = (h_1, \dots, h_n) \in \mathbb{R}^n$ mit $0 < \|h\| < \delta$ gegeben. Dann definieren wir die Vektoren $x_1, \dots, x_n \in \mathbb{R}^n$ durch

$$x_k := x_0 + h_1 e_1 + \dots + h_k e_k \quad \text{für } k = 1, \dots, n.$$

Man beachte dass die x_i hier alles Vektoren im \mathbb{R}^n sind, und *nicht* die Koordinaten eines Vektors. Sie erfüllen die Gleichungen

$$x_k = x_{k-1} + h_k e_k \quad \text{für } k = 2, \dots, n, \quad x_n = x_0 + h. \quad (2.2.2)$$

Da $\|x_k - x_0\| \leq \|h\| < \delta$ ist, gilt $x_k \in U$ nach (2.2.1) und daher, nach (2.2.2),

$$f(x_0 + h) - f(x_0) = \sum_{k=1}^n (f(x_k) - f(x_{k-1})). \quad (2.2.3)$$

Ebenso gilt $\|x_{k-1} + th_k e_k - x_0\| < \delta$ für alle $t \in [0, 1]$ und alle $k \in \{1, \dots, n\}$. Nach (2.2.1) folgt daraus $x_{k-1} + th_k e_k \in U$ und

$$|\partial_k f(x_{k-1} + th_k e_k) - \partial_k f(x_0)| < \frac{\varepsilon}{n} \quad (2.2.4)$$

für $0 \leq t \leq 1$ und $k = 1, \dots, n$. Für jedes solche k definieren wir die Funktion $\phi_k : [0, 1] \rightarrow \mathbb{R}$ durch

$$\phi_k(t) := f(x_{k-1} + th_k e_k) \quad \text{für } 0 \leq t \leq 1.$$

Nach dem Mittelwertsatz existiert für jedes $k \in \{1, \dots, n\}$ ein $t_k \in (0, 1)$ mit $\phi'_k(t_k) = \phi_k(1) - \phi_k(0)$. Mit $\phi_k(0) = f(x_{k-1})$ und $\phi_k(1) = f(x_k)$ gilt

$$h_k \partial_k f(x_{k-1} + t_k h_k e_k) = \phi'_k(t_k) = f(x_k) - f(x_{k-1}). \quad (2.2.5)$$

Aus (2.2.3), (2.2.4), und (2.2.5) ergibt sich

$$\begin{aligned} & \left| f(x_0 + h) - f(x_0) - \sum_{k=1}^n h_k \partial_k f(x_0) \right| \\ &= \left| \sum_{k=1}^n (f(x_{k+1}) - f(x_k) - h_k \partial_k f(x_0)) \right| \\ &= \left| \sum_{k=1}^n h_k (\partial_k f(x_{k-1} + t_k h_k e_k) - \partial_k f(x_0)) \right| \\ &\leq \sum_{k=1}^n |h_k| |\partial_k f(x_{k-1} + t_k h_k e_k) - \partial_k f(x_0)| \\ &< \sum_{k=1}^n \frac{\varepsilon |h_k|}{n} \leq \varepsilon \|h\|. \end{aligned}$$

Damit folgt die Behauptung von Satz 2.2.1 aus der Definition der Differenzierbarkeit in Definition 2.1.1. \square

Mit Hilfe von Satz 2.2.1 können wir für viele Beispiele von Funktionen von mehreren Variablen die Differenzierbarkeit auf die aus der Analysis I bekannte Differenzierbarkeit von Funktionen einer reellen Variablen zurückführen.

Beispiel 2.2.4. Die durch

$$f(x, y) := x^y = e^{y \log(x)}$$

definierte Funktion $f : U := (0, \infty) \times \mathbb{R} \rightarrow \mathbb{R}$ definierte Funktion ist stetig differenzierbar, denn nach Resultaten aus Analysis I ist sie partiell differenzierbar, und ihre partiellen Ableitungen sind durch

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= e^{y \log(x)} \frac{y}{x} = yx^{y-1}, \\ \frac{\partial f}{\partial y}(x, y) &= e^{y \log(x)} \log(x) = \log(x)x^y \end{aligned}$$

für $x > 0$ und $y \in \mathbb{R}$ gegeben. Diese Funktionen $\frac{\partial f}{\partial x} : U \rightarrow \mathbb{R}$ und $\frac{\partial f}{\partial y} : U \rightarrow \mathbb{R}$ sind stetig, ebenfalls nach Beispielen aus Analysis I und einem Resultat aus Analysis I, welches besagt, dass die Summe, das Produkt, und der Quotient zweier stetiger Funktionen (auf beliebigen metrischen Räumen) wieder stetig sind. Daraus folgt nach Satz 2.2.1, dass unsere Funktion $f : U \rightarrow \mathbb{R}$ stetig differenzierbar ist. Nach Lemma 2.1.5 wird ihre Ableitung $df(x, y) : \mathbb{R}^2 \rightarrow \mathbb{R}$ (als lineare Abbildung) durch die, ebenfalls mit $df(x, y)$ bezeichnete, Jacobi-Matrix

$$df(x, y) = (yx^{y-1}, \log(x)x^y) \in \mathbb{R}^{1 \times 2}$$

repräsentiert.

Beispiel 2.2.5. In der gleichen Weise sieht man, dass die durch

$$f(x) := \begin{pmatrix} x_1 + x_2 + x_3 \\ x_1x_2 + x_3x_1 + x_2x_3 \\ x_1x_2x_3 \end{pmatrix}$$

für $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ definierte Abbildung $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ stetig differenzierbar ist. Ihre Jacobi-Matrix an der Stelle $x \in \mathbb{R}^3$ ist

$$df(x) = \begin{pmatrix} 1 & 1 & 1 \\ x_2 + x_3 & x_3 + x_1 & x_1 + x_2 \\ x_2x_3 & x_3x_1 & x_1x_2 \end{pmatrix}.$$

Dies ist nur *ein* Beispiel. Nach Satz 2.2.1 und den entsprechenden Resultaten aus Analysis I ist jede polynomiale Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ stetig differenzierbar.

Der nächste Satz betrifft Folgen stetig differenzierbarer Funktionen und besagt dass, wenn sie ebenso wie ihre partiellen Ableitungen gleichmäßig konvergiert, ihr Grenzwert wieder stetig differenzierbar ist. Anwendungen dieses wichtigen Satzes finden sich in den Abschnitten 2.3, 2.5 und 2.6.

Satz 2.2.6 (C^1 -Konvergenz).

Sei $U \subset \mathbb{R}^n$ offen und sei $f_k : U \rightarrow \mathbb{R}$ eine Folge stetig differenzierbarer Funktionen, so dass die Funktionenfolgen $(f_k)_{k \in \mathbb{N}}$ und $(\partial_i f_k)_{k \in \mathbb{N}}$ für $i = 1, \dots, n$ auf jeder kompakten Teilmenge von U gleichmäßig konvergieren. Seien

$$f(x) := \lim_{k \rightarrow \infty} f_k(x), \quad g_i(x) := \lim_{k \rightarrow \infty} \partial_i f_k(x) \quad \text{für } x \in U \text{ und } i = 1, \dots, n.$$

Dann ist $f : U \rightarrow \mathbb{R}$ stetig differenzierbar mit $\partial_i f = g_i$ für $i = 1, \dots, n$.

Beweis. Nach Voraussetzung gilt

$$\lim_{k \rightarrow \infty} \sup_{x \in K} |f(x) - f_k(x)| = 0, \quad \lim_{k \rightarrow \infty} \sup_{x \in K} |g_i(x) - \partial_i f_k(x)| = 0 \quad (2.2.6)$$

für $i = 1, \dots, n$ und jede kompakte Teilmenge $K \subset U$. Da der Grenzwert einer gleichmäßig konvergenten Folge stetiger Funktionen selbst wieder stetig ist (Analysis I), sind die Funktionen $f, g_1, \dots, g_n : U \rightarrow \mathbb{R}$ stetig. Es bleibt zu zeigen, dass die partielle Ableitung $\partial_i f(x)$ für alle $x \in U$ und alle i existiert und gleich $g_i(x)$ ist. Dazu wählen wir $x \in U$ und $\varepsilon > 0$ mit $\overline{B}_\varepsilon(x) \subset U$. Sei $i \in \{1, \dots, n\}$ und sei $e_i \in \mathbb{R}^n$ der Standardbasisvektor. Nach Voraussetzung ist die Funktion $[-\varepsilon, \varepsilon] \rightarrow \mathbb{R} : t \mapsto f_k(x + te_i)$ stetig differenzierbar mit Ableitung $\frac{d}{dt} f_k(x + te_i) = \partial_i f_k(x + te_i)$, und daraus folgt

$$f_k(x + te_i) = \int_0^t \partial_i f_k(x + se_i) ds \quad (2.2.7)$$

für alle $k \in \mathbb{N}$ und alle $t \in [-\varepsilon, \varepsilon]$. Außerdem gilt nach [5, Satz 3.1]

$$\begin{aligned} \left| \int_0^t g_i(x + se_i) - \int_0^t \partial_i f_k(x + se_i) ds \right| &\leq \int_0^t |g_i(x + se_i) - \partial_i f_k(x + se_i)| ds \\ &\leq \varepsilon \sup_{y \in \overline{B}_\varepsilon(x)} |\partial_i f_k(y) - g_i(y)| \end{aligned}$$

für $t \in [-\varepsilon, \varepsilon]$ und alle $k \in \mathbb{N}$. Nach (2.2.6) konvergiert die rechte Seite dieser Ungleichung gegen Null für $k \rightarrow \infty$, und nach (2.2.7) folgt daraus

$$f(x + te_i) = \int_0^t g_i(x + se_i) ds \quad \text{für } -\varepsilon \leq t \leq \varepsilon. \quad (2.2.8)$$

Die Funktion $[-\varepsilon, \varepsilon] \rightarrow \mathbb{R} : t \mapsto f(x + te_i)$ ist daher stetig differenzierbar und es gilt $\partial_i f(x) = \frac{d}{dt} \Big|_{t=0} f(x + te_i) = g_i(x)$. Damit ist Satz 2.2.6 bewiesen. \square

2.3 Die Kettenregel

Wir wenden uns nun den grundlegenden Rechenregeln der Differentialrechnung zu, die insbesondere die Leibnizregel und die Kettenregel beinhalten. Hier sind die Hauptresultate dieses Abschnitts.

Satz 2.3.1 (Rechenregeln). *Sei $(X, \|\cdot\|)$ ein endlichdimensionaler normierter Vektorraum, sei $U \subset X$ eine offenen Teilmenge, sei $x_0 \in U$, und seien $f, g : U \rightarrow \mathbb{R}$ Funktionen, die an der Stelle x_0 differenzierbar sind. Dann gilt folgendes.*

(i) *Die Summe $f + g : U \rightarrow \mathbb{R}$ ist an der Stelle x_0 differenzierbar, und es gilt*

$$d(f + g)(x_0)\xi = df(x_0)\xi + dg(x_0)\xi \quad (2.3.1)$$

für alle $\xi \in X$.

(ii) *Das Produkt $fg : U \rightarrow \mathbb{R}$ ist an der Stelle x_0 differenzierbar, und ihre Ableitung ist durch die **Leibnizregel***

$$d(fg)(x_0)\xi = g(x_0)df(x_0)\xi + f(x_0)dg(x_0)\xi \quad (2.3.2)$$

für alle $\xi \in X$ gegeben.

(iii) *Ist $g(x) \neq 0$ für alle $x \in U$, so ist der Quotient $f/g : U \rightarrow \mathbb{R}$ an der Stelle x_0 differenzierbar, und es gilt*

$$d\left(\frac{f}{g}\right)(x_0)\xi = \frac{g(x_0)df(x_0)\xi - f(x_0)dg(x_0)\xi}{g(x_0)^2} \quad (2.3.3)$$

für alle $\xi \in X$.

Beweis. Siehe Seite 50. □

Satz 2.3.2 (Kettenregel). *Seien $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$, $(Z, \|\cdot\|_Z)$ endlichdimensionale Vektorräume, seien $U \subset X$ und $V \subset Y$ offenen Teilmengen, sei $f : U \rightarrow V$ eine an der Stelle $x_0 \in U$ differenzierbare Abbildung, und sei $g : V \rightarrow Z$ eine an der Stelle $y_0 := f(x_0) \in V$ differenzierbare Abbildung. Dann ist die Komposition $g \circ f : U \rightarrow Z$ an der Stelle x_0 differenzierbar, und ihre Ableitung an der Stelle x_0 ist die Komposition der Ableitungen von f und g , das heißt*

$$d(g \circ f)(x_0) = dg(f(x_0)) \circ df(x_0) : X \rightarrow Z. \quad (2.3.4)$$

Beweis. Siehe Seite 52. □

Den Beweis von Satz 2.3.1 kann man auf dreierlei verschieden Weisen führen. Erstens könnte man den Beweis des entsprechenden Satzes für Funktionen einer reellen Variablen auf Funktionen mehrerer Variablen übertragen. Zum zweiten könnte man ohne Beschränkung der Allgemeinheit annehmen, dass $X = \mathbb{R}^n$ ist und dann, *mit* Beschränkung der Allgemeinheit, zusätzlich voraussetzen, dass f und g stetig differenzierbar sind, was für die allermeisten Anwendungen ausreicht. In dieser Situation folgen die Behauptungen von Satz 2.3.1 unmittelbar aus Satz 2.2.1 und den entsprechenden Resultaten über die Summe, das Produkt, und den Quotienten von Funktionen einer reellen Variablen in Analysis I. Insbesondere ergeben sich die Formeln in Satz 2.3.1 dann aus der Betrachtung der Funktionen $t \mapsto f(x_0 + t\xi)$ und $t \mapsto g(x_0 + t\xi)$. Die dritte Methode ist es, Satz 2.3.1 auf die Kettenregel in Satz 2.3.2 zurückzuführen, und diese Methode werden wir hier verwenden.

Satz 2.3.2 \implies *Satz 2.3.1*. Die durch

$$F(x) := \begin{pmatrix} f(x) \\ g(x) \end{pmatrix} \quad \text{für } x \in U$$

definierte Abbildung $F : U \rightarrow \mathbb{R}^2$ ist nach Lemma 2.1.15 an der Stelle x_0 differenzierbar und ihre Ableitung $dF(x_0) : X \rightarrow \mathbb{R}^2$ ist

$$dF(x_0)\xi = \begin{pmatrix} df(x_0)\xi \\ dg(x_0)\xi \end{pmatrix} \quad \text{für } \xi \in X. \quad (2.3.5)$$

Wir beweisen Teil (i). Die durch

$$h(y, z) := y + z$$

definierte Funktion $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ ist nach Satz 2.2.1 differenzierbar und ihre Ableitung $dh(y, z) : \mathbb{R}^2 \rightarrow \mathbb{R}$ an der Stelle $(y, z) \in \mathbb{R}^2$ ist durch

$$dh(y, z)(\eta, \zeta) = \eta + \zeta$$

für $(\eta, \zeta) \in \mathbb{R}^2$ gegeben. Daraus folgt nach Satz 2.3.2, dass die Funktion

$$f + g = h \circ F : U \rightarrow \mathbb{R}$$

an der Stelle x_0 differenzierbar ist mit der Ableitung

$$\begin{aligned} d(f + g)(x_0)\xi &= dh(F(x_0))dF(x_0)\xi \\ &= dh(f(x_0), g(x_0))(df(x_0)\xi, dg(x_0)\xi) \\ &= df(x_0)\xi + dg(x_0)\xi \end{aligned}$$

für $\xi \in X$.

Wir beweisen Teil (ii). Die durch

$$h(y, z) := yz$$

definierte Funktion $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ ist nach Satz 2.2.1 differenzierbar und ihre Ableitung $dh(y, z) : \mathbb{R}^2 \rightarrow \mathbb{R}$ an der Stelle $(y, z) \in \mathbb{R}^2$ ist durch

$$dh(y, z)(\eta, \zeta) = z\eta + y\zeta$$

für $(\eta, \zeta) \in \mathbb{R}^2$ gegeben. Daraus folgt nach Satz 2.3.2, dass die Funktion

$$fg = h \circ F : U \rightarrow \mathbb{R}$$

an der Stelle x_0 differenzierbar ist mit der Ableitung

$$\begin{aligned} d(f + g)(x_0)\xi &= dh(F(x_0))dF(x_0)\xi \\ &= dh(f(x_0), g(x_0))(df(x_0)\xi, dg(x_0)\xi) \\ &= g(x_0)df(x_0)\xi + f(x_0)dg(x_0)\xi \end{aligned}$$

für $\xi \in X$.

Wir beweisen Teil (iii). Da $g(x) \neq 0$ ist für alle $x \in U$, gilt

$$F(U) \subset V := \mathbb{R} \times (\mathbb{R} \setminus \{0\})$$

Nun ist die durch

$$h(y, z) := \frac{y}{z}$$

definierte Funktion $h : V \rightarrow \mathbb{R}$ nach Satz 2.2.1 differenzierbar und ihre Ableitung $dh(y, z) : \mathbb{R}^2 \rightarrow \mathbb{R}$ an der Stelle $(y, z) \in \mathbb{R}^2$ ist durch

$$dh(y, z)(\eta, \zeta) = \frac{\eta}{z} - \frac{y\zeta}{z^2}$$

für $(\eta, \zeta) \in \mathbb{R}^2$ gegeben. Daraus folgt nach Satz 2.3.2, dass die Funktion

$$fg = h \circ F : U \rightarrow \mathbb{R}$$

an der Stelle x_0 differenzierbar ist mit der Ableitung

$$\begin{aligned} d(f + g)(x_0)\xi &= dh(F(x_0))dF(x_0)\xi \\ &= dh(f(x_0), g(x_0))(df(x_0)\xi, dg(x_0)\xi) \\ &= \frac{g(x_0)df(x_0)\xi - f(x_0)dg(x_0)\xi}{g(x_0)^2} \end{aligned}$$

für $\xi \in X$. Damit ist Satz 2.3.1 bewiesen. \square

Der folgende Beweis von Satz 2.3.2 läßt sich Wort für Wort auf differenzierbare Abbildungen zwischen Banachräumen erweitern.

Beweis von Satz 2.3.2. Seien $A \in \mathcal{L}(X, Y)$ und $B \in \mathcal{L}(Y, Z)$ die linearen Abbildungen

$$A := df(x_0), \quad B := dg(y_0).$$

Zu zeigen ist, dass die Komposition

$$g \circ f : U \rightarrow Z$$

an der Stelle $x_0 \in U$ differenzierbar ist mit der Ableitung

$$d(g \circ f)(x_0) = BA \in \mathcal{L}(X, Z).$$

Sei dazu $\varepsilon > 0$ gegeben. Da g an der Stelle $y_0 = f(x_0) \in V$ differenzierbar ist mit $dg(y_0) = B$, existiert eine Konstante $\rho > 0$ mit

$$B_\rho(y_0; Y) = \{y \in Y \mid \|y - y_0\|_Y < \rho\} \subset V, \quad (2.3.6)$$

so dass jedes Element $y \in B_\rho(y_0; Y)$ die Ungleichung

$$\|g(y) - g(y_0) - B(y - y_0)\|_Z \leq \frac{\varepsilon}{2(1 + \|A\|)} \|y - y_0\|_Y \quad (2.3.7)$$

erfüllt. Da f an der Stelle $x_0 \in U$ differenzierbar ist mit $df(x_0) = A$, existiert eine Konstante $\delta > 0$ mit

$$B_\delta(x_0; X) = \{x \in X \mid \|x - x_0\|_X < \delta\} \subset U, \quad (2.3.8)$$

so dass jedes Element $x \in B_\delta(x_0; X)$ die Ungleichung

$$\|f(x) - f(x_0) - A(x - x_0)\|_Y \leq \frac{\varepsilon}{2(\varepsilon + \|B\|)} \|x - x_0\|_X \quad (2.3.9)$$

erfüllt. Wählen wir dieses $\delta > 0$ hinreichend klein, so gilt

$$(1 + \|A\|)\delta < \rho. \quad (2.3.10)$$

Nun sei $x \in X$ fest gewählt mit $\|x - x_0\|_X < \delta$. Dann ist $x \in U$ nach (2.3.8) und aus (2.3.9) und (2.3.10) folgt mit $f(x_0) = y_0$ die Ungleichung

$$\begin{aligned} \|f(x) - y_0\|_Y &\leq \|f(x) - f(x_0) - A(x - x_0)\|_Y + \|A(x - x_0)\|_Y \\ &\leq \frac{\varepsilon}{2(\varepsilon + \|B\|)} \|x - x_0\|_X + \|A\| \|x - x_0\|_X \\ &\leq (1 + \|A\|) \|x - x_0\|_X \\ &< (1 + \|A\|)\delta \\ &< \rho. \end{aligned} \quad (2.3.11)$$

Daraus folgt $f(x) \in V$ nach (2.3.6) und

$$\begin{aligned}
& \|g(f(x)) - g(f(x_0)) - BA(x - x_0)\|_Z \\
& \leq \|g(f(x)) - g(y_0) - B(f(x) - y_0)\|_Z + \|B(f(x) - f(x_0)) - A(x - x_0)\|_Z \\
& \leq \frac{\varepsilon}{2} \frac{\|f(x) - y_0\|_Y}{1 + \|A\|} + \|B\| \|f(x) - f(x_0) - A(x - x_0)\|_Y \\
& \leq \left(\frac{\varepsilon}{2} + \frac{\varepsilon \|B\|}{2(\varepsilon + \|B\|)} \right) \|x - x_0\|_X \\
& \leq \varepsilon \|x - x_0\|_X.
\end{aligned}$$

Hier folgt der zweite Schritt aus (2.3.7) und der dritte Schritt aus (2.3.9) und (2.3.11). Damit folgt die Behauptung von Satz 2.3.2 aus der Definition der Differenzierbarkeit in Definition 2.1.1. \square

Korollar 2.3.3. *Seien $U \subset \mathbb{R}^n$ und $V \subset \mathbb{R}^m$ offenen Mengen und seien*

$$f = (f_1, \dots, f_m) : U \rightarrow V, \quad g = (g_1, \dots, g_\ell) : V \rightarrow \mathbb{R}^\ell$$

stetig differenzierbare Abbildungen. Dann ist $g \circ f : U \rightarrow \mathbb{R}^\ell$ stetig differenzierbar, und es gilt

$$\frac{\partial(g_k \circ f)}{\partial x_i}(x) = \sum_{j=1}^m \frac{\partial g_k}{\partial y_j}(f(x)) \frac{\partial f_j}{\partial x_i}(x) \quad (2.3.12)$$

für $x \in U$, $i = 1, \dots, n$, und $k = 1, \dots, \ell$.

Beweis. Sei $e_i \in \mathbb{R}^n$ der i te Basisvektor in der Standardbasis. Dann gilt

$$\begin{aligned}
\frac{\partial(g_k \circ f)}{\partial x_i}(x) &= d(g_k \circ f)(x)e_i \\
&= dg_k(f(x))df(x)e_i \\
&= \left(\frac{\partial g_k}{\partial x_1}(f(x)) \cdots \frac{\partial g_k}{\partial x_m}(f(x)) \right) \begin{pmatrix} df_1(x)e_i \\ \vdots \\ df_m(x)e_i \end{pmatrix} \\
&= \sum_{j=1}^m \frac{\partial g_k}{\partial y_j}(f(x)) \frac{\partial f_j}{\partial x_i}(x).
\end{aligned}$$

Hier folgt die erste Gleichheit aus Lemma 2.1.11, die zweite aus Satz 2.3.2, die dritte aus Lemma 2.1.15 und Lemma 2.1.16, und die vierte aus Lemma 2.1.11. Damit ist die Gleichung (2.3.12) bewiesen, und daraus folgt sofort die Stetigkeit der partiellen Ableitungen von $g \circ f$. \square

Die Gleichung (2.3.12) sagt, dass die Jacobi-Matrix von $g \circ f$ an der Stelle x das Produkt der Jacobi-Matrix von g an der Stelle $f(x)$ mit der Jacobi-Matrix von f an der Stelle x ist. Diese Gleichung kann auch in der Kurzform

$$\frac{\partial z_k}{\partial x_i} = \sum_{j=1}^m \frac{\partial z_k}{\partial y_j} \frac{\partial y_j}{\partial x_i} \quad (2.3.13)$$

geschrieben werden. Hier steht y_j für $f_j(x)$ und z_k für $g_k(y)$. Diese Schreibweise ist dann nützlich, wenn die y_j als Funktionen der Variablen x_1, \dots, x_n und die z_k als Funktionen der Variablen y_1, \dots, y_m verstanden werden.

Beispiele

Beispiel 2.3.4 (Gradient). Sei $U \subset \mathbb{R}^n$ eine offene Menge und $f : U \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion. Dann ist die Ableitung von f an der Stelle $x \in U$ eine lineare Abbildung $df(x) : \mathbb{R}^n \rightarrow \mathbb{R}$, welche durch die Jacobi-Matrix $df(x) = (\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x)) \in \mathbb{R}^{1 \times n}$ (einen Zeilenvektor) repräsentiert wird. Es ist oft nützlich, die Ableitung als einen Spaltenvektor zu schreiben. Dies entspricht dem Übergang von einem linearen Funktional auf dem \mathbb{R}^n zu einem Vektor im \mathbb{R}^n . Im Fall der Ableitung von f wird dieser Vektor der **Gradient von f an der Stelle x** genannt und mit

$$\nabla f(x) := \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix} \quad (2.3.14)$$

bezeichnet. Diese Schreibweise ist besonders nützlich in Verbindung mit dem standard inneren Produkt $\langle \xi, \eta \rangle = \sum_{i=1}^n \xi_i \eta_i$ zweier Vektoren $\xi, \eta \in \mathbb{R}^n$. Die Richtungsableitung von f an der Stelle $x \in U$ in Richtung $\xi \in \mathbb{R}^n$ ist dann

$$df(x)\xi = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x)\xi_i = \langle \nabla f(x), \xi \rangle. \quad (2.3.15)$$

Ist nun $x : I \rightarrow U$ eine C^1 -Kurve, so ergibt die Kettenregel die Formel

$$\frac{d}{dt} f(x(t)) = df(x(t))\dot{x}(t) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x(t))\dot{x}_i(t) = \langle \nabla f(x(t)), \dot{x}(t) \rangle. \quad (2.3.16)$$

(Hier hat der Vektor $x = (x_1, \dots, x_n)$ zwei Bedeutungen! Zum einen bezeichnet er die unabhängigen Variablen in der Menge $U \subset \mathbb{R}^n$, und zum anderen bezeichnet er die Abbildung $I \rightarrow U : t \mapsto x(t) = (x_1(t), \dots, x_n(t))$.)

Beispiel 2.3.5. Sei $f : U := (0, \infty) \times \mathbb{R} \rightarrow \mathbb{R}$ die Funktion

$$f(x, y) := x^y \quad \text{für } x > 0 \text{ und } y \in \mathbb{R}$$

und sei die Kurve $I = (0, \infty) \rightarrow U : t \mapsto (x(t), y(t))$ durch

$$x(t) := y(t) := t \quad \text{für } t > 0$$

gegeben. Dann ist $f(x(t), y(t)) = t^t$ für $t > 0$ und nach Beispiel 2.2.4 gilt

$$\frac{\partial f}{\partial x}(x, y) = yx^{y-1}, \quad \frac{\partial f}{\partial y}(x, y) = \log(x)x^y$$

für alle $x > 0$ und alle $y \in \mathbb{R}$. Daher ergibt Gleichung (2.3.16) die Formel

$$\frac{d}{dt}t^t = \frac{\partial f}{\partial x}(x(t), y(t))\dot{x}(t) + \frac{\partial f}{\partial y}(x(t), y(t))\dot{y}(t) = (1 + \log(t))t^t. \quad (2.3.17)$$

Beispiel 2.3.6. In der allgemeinen Situation von Beispiel 2.3.4 ist es oft interessant, die Lösungen $x : I \rightarrow U$ der Differentialgleichung

$$\dot{x}(t) = -\nabla f(x(t)) \quad (2.3.18)$$

zu betrachten. Für diese Lösungen hat die Gleichung (2.3.16) die Form

$$\frac{d}{dt}f(x(t)) = \langle \nabla f(x(t)), \dot{x}(t) \rangle = -\|\nabla f(x(t))\|^2. \quad (2.3.19)$$

Nach dem Mittelwertsatz folgt daraus, dass die Funktion $I \rightarrow \mathbb{R} : t \mapsto f(x(t))$ monoton fallend ist. **Übung:** Die Funktion $I \rightarrow \mathbb{R} : t \mapsto f(x(t))$ ist entweder auf ganz I strikt monoton fallend, oder die Lösung $I \rightarrow U : t \mapsto x(t)$ ist konstant. Schließen Sie daraus, dass die Differentialgleichung (2.3.18) keine nichtkonstanten periodischen Lösungen haben kann.

Beispiel 2.3.7. Eine andere interessante Situation entsteht in Beispiel 2.3.4, wenn die Funktion $f : U \rightarrow \mathbb{R}$ entlang einer Kurve $x : I \rightarrow U$ konstant ist, so dass $f(x(t)) = c$ ist für alle $t \in I$ und eine Konstante $c \in \mathbb{R}$. In dem Fall folgt aus (2.3.16) die Gleichung

$$0 = \frac{d}{dt}f(x(t)) = \langle \nabla f(x(t)), \dot{x}(t) \rangle.$$

Das heißt, der Geschwindigkeitsvektor $\dot{x}(t)$ steht senkrecht auf dem Gradienten von f an der Stelle $x(t)$.

Beispiel 2.3.8. Eine weitere Anwendung von Beispiel 2.3.4 ergibt sich aus den **Newtonschen Gleichungen**

$$\ddot{x}(t) = -\nabla V(x(t)). \quad (2.3.20)$$

Hier ist $U \subset \mathbb{R}^n$ eine offene Menge und $V : U \rightarrow \mathbb{R}$ eine glatte Potentialfunktion. Ihr Gradient ist das *Kraftfeld* und die Gleichung (2.3.20) ist das Newtonsche Gesetz “*Kraft gleich Masse mal Beschleunigung*” (mit Masse gleich 1). Die *Energie* einer Lösung $x : I \rightarrow U$ zum Zeitpunkt t ist die Summe aus *kinetischer* und *potentieller* Energie und ist durch

$$E(t) := \frac{1}{2} \|\dot{x}(t)\|^2 + V(x(t)) \quad (2.3.21)$$

Nach der Kettenregel in (2.3.16) ist die Zeitableitung der Energie

$$\dot{E}(t) = \langle \dot{x}(t), \ddot{x}(t) \rangle + \langle \nabla V(x(t)), \dot{x}(t) \rangle = \langle \dot{x}(t), \ddot{x}(t) + \nabla V(x(t)) \rangle = 0.$$

Also ist die Energie entlang jeder Lösung von (2.3.20) konstant. In dem Spezialfall $n = 3$ mit $U = \mathbb{R}^3 \setminus \{0\}$ und

$$V(x) = -\frac{1}{\|x\|} \quad \nabla V(x) = \frac{x}{\|x\|^3}$$

erhält man das **Keplerproblem**.

Beispiel 2.3.9. Lemma 2.1.11 kann auch als Spezialfall der Kettenregel hergeleitet werden. Seien also X, Y endlichdimensionale normierte Vektorräume, sei $U \subset X$ eine offene Menge, und sei $f : U \rightarrow Y$ eine Abbildung, die an der Stelle $x \in U$ differenzierbar ist. Weiterhin sei $\hat{x} \in X$ gegeben und sei $I \subset \mathbb{R}$ ein offenes Intervall mit $0 \in I$, so dass

$$\gamma(t) := x + t\hat{x} \in U$$

ist für alle $t \in I$. Nach Beispiel 2.1.7 ist die Abbildung $\gamma : I \rightarrow U$ differenzierbar mit $\dot{\gamma}(t) = \hat{x}$ für alle $t \in I$. Nach Satz 2.3.2 ist daher die Abbildung $f \circ \gamma : I \rightarrow Y$ an der Stelle $t = 0$ differenzierbar und es gilt

$$\left. \frac{d}{dt} \right|_{t=0} f(x + t\hat{x}) = \left. \frac{d}{dt} \right|_{t=0} (f \circ \gamma)(t) = df(\gamma(0))\dot{\gamma}(0) = df(x)\hat{x}. \quad (2.3.22)$$

Dies ist die Gleichung (2.1.8) in Lemma 2.1.11 mit $a = x$ und $\xi = \hat{x}$.

Die folgenden Beispiele zeigt, dass es durchaus sinnvoll sein kann, mit allgemeinen endlichdimensionalen Vektorräumen zu arbeiten, anstatt mit dem \mathbb{R}^n . Man entledigt sich dadurch der Notwendigkeit, eine Basis zu wählen, und manchmal ist es auch sinnvoll, eine andere als die Euklidische Norm zu verwenden. So sind etwa die Normen auf $\mathbb{R}^{m \times n}$ in Beispiel 1.2.6 nicht die Normen von inneren Produkten.

Beispiel 2.3.10. Sei der Raum $X := \mathbb{R}^{n \times n}$ mit einer der Normen in Beispiel 1.2.6 versehen und sei $g : X \times X \rightarrow X$ die Abbildung

$$g(A, B) = AB \quad \text{für } A, B \in \mathbb{R}^{n \times n}.$$

Diese Abbildung ist differenzierbar nach Beispiel 2.1.8 und Lemma 2.1.15. Ihre Ableitung $dg(A, B) \in \mathcal{L}(X \times X, X)$ an der Stelle $(A, B) \in X \times X$ läßt sich mit Hilfe von Beispiel 2.3.9 bestimmen. Für $(\widehat{A}, \widehat{B}) \in X \times X$ gilt

$$\begin{aligned} dg(A, B)(\widehat{A}, \widehat{B}) &= \left. \frac{d}{dt} \right|_{t=0} f(A + t\widehat{A}, B + t\widehat{B}) \\ &= \left. \frac{d}{dt} \right|_{t=0} (A + t\widehat{A})(B + t\widehat{B}) \\ &= \left. \frac{d}{dt} \right|_{t=0} (AB + t(\widehat{A}B + A\widehat{B}) + t^2\widehat{A}\widehat{B}) \\ &= \widehat{A}B + A\widehat{B}. \end{aligned}$$

Beispiel 2.3.11. Sei $X := \mathbb{R}^{n \times n}$. Dann ist die Abbildung $\phi_k : X \rightarrow X$ mit

$$\phi_k(A) := A^k \quad \text{für } A \in X$$

stetig differenzierbar und ihre Ableitung $d\phi_k(A) \in \mathcal{L}(X, X)$ ist

$$d\phi_k(A)\widehat{A} = \sum_{j=0}^{k-1} A^j \widehat{A} A^{k-1-j} \quad \text{für } A, \widehat{A} \in X. \quad (2.3.23)$$

für $A, \widehat{A} \in X$ gegeben. **Beweis:** Für $k = 1$ erfüllt $\phi_1 = \text{id} : X \rightarrow X$ die Behauptung nach Beispiel 2.1.7. Erfüllt f_k die Behauptung für ein $k \in \mathbb{N}$, ist $\psi_k : X \rightarrow X \times X$ durch $\psi_k(A) := (A, \phi_k(A))$ für $A \in X$ definiert, und ist $g : X \times X \rightarrow X$ die Abbildung aus Beispiel 2.3.10, so ist $\phi_{k+1} = g \circ \psi_k$ nach Satz 2.3.2 stetig differenzierbar mit $d\psi_{k+1}(A)\widehat{A} = dg(\psi_k(A))d\psi_k(A)\widehat{A}$ für $A, \widehat{A} \in X$. Daraus folgt auch die Formel (2.3.23) für $k + 1$ (Übung).

Potenzreihen von Matrizen

Sei a_0, a_1, a_2, \dots eine Folge reeller Zahlen mit

$$\rho := \frac{1}{\limsup_{k \rightarrow \infty} |a_k|^{1/k}} > 0. \quad (2.3.24)$$

Sei $\|\cdot\|$ eine Matrixnorm auf dem Raum $X := \mathbb{R}^{n \times n}$ (siehe Beispiel 1.3.6). Nun sei $U := \{A \in \mathbb{R}^{n \times n} \mid \|A\| < \rho\}$, und sei $f : U \rightarrow X$ die Abbildung

$$f(A) := \sum_{k=0}^{\infty} a_k A^k \quad (2.3.25)$$

für $A \in U$. Dass diese Reihe für jedes $A \in U$ absolut konvergiert wurde in Korollar 1.3.11 bewiesen, und dass die durch (2.3.25) definierte Abbildung $f : U \rightarrow X$ stetig ist, wurde in Satz 1.3.13 bewiesen.

Satz 2.3.12. *Die Abbildung $f : U \rightarrow X$ in (2.3.25) ist stetig differenzierbar und ihre Ableitung $df(A) \in \mathcal{L}(X, X)$ an der Stelle $A \in U$ ist*

$$df(A)\hat{A} := \sum_{k=1}^{\infty} a_k \sum_{j=0}^{k-1} A^j \hat{A} A^{k-1-j} \quad \text{für } \hat{A} \in X. \quad (2.3.26)$$

Beweis. Für $\ell \in \mathbb{N}$ sei $f_\ell : U \rightarrow X$ die Abbildung

$$f_\ell(A) := \sum_{k=0}^{\ell} a_k \phi_k(A) = \sum_{k=0}^{\ell} a_k A^k \quad \text{für } A \in U.$$

Nach Beispiel 2.3.11 und Satz 2.3.1 ist f_ℓ stetig differenzierbar mit

$$df_\ell(A) = \sum_{k=1}^{\ell} a_k \Phi_k(A), \quad \Phi_k(A)\hat{A} := \sum_{j=0}^{k-1} A^j \hat{A} A^{k-1-j}.$$

Für $0 < r < \rho$ sei $B_r := \{A \in X \mid \|A\| \leq r\}$. Dann gilt

$$\|\Phi_k(A)\hat{A}\| \leq \sum_{j=0}^{k-1} \|A^j \hat{A} A^{k-1-j}\| \leq k \|A\|^{k-1} \|\hat{A}\| \leq k r^{k-1} \|\hat{A}\|$$

für $A \in B_r, \hat{A} \in X$, und daher $\sum_{k=1}^{\infty} \|a_k \Phi_k(A)\|_{\mathcal{L}(X, X)} \leq \sum_{k=1}^{\infty} k |a_k| r^{k-1} < \infty$ für $A \in B_r$ nach (2.3.24). Damit folgt aus (dem Beweis von) Satz 1.3.10, dass die Abbildungsfolge $df_\ell : U \rightarrow \mathcal{L}(X, X)$ auf jeder kompakten Teilmenge von U gleichmäßig konvergiert. Da auch $f_\ell : U \rightarrow X$ auf jeder kompakten Teilmenge von U gleichmäßig gegen f konvergiert, ist $f : U \rightarrow X$ nach Satz 2.2.6 stetig differenzierbar mit der Ableitung (2.3.26). \square

Beispiel 2.3.13. Nach Satz 2.3.12 ist die Exponentialabbildung

$$\exp : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}, \quad \exp(A) = \sum_{k=0}^{\infty} \frac{A^k}{k!}, \quad (2.3.27)$$

stetig differenzierbar mit der Ableitung

$$d \exp(A) \widehat{A} = \sum_{k=1}^{\infty} \frac{1}{k!} \sum_{j=0}^{k-1} A^j \widehat{A} A^{k-1-j} \quad (2.3.28)$$

für $A, \widehat{A} \in \mathbb{R}^{n \times n}$. Insbesondere gilt

$$d \exp(0) \widehat{A} = \widehat{A}$$

für alle $\widehat{A} \in \mathbb{R}^{n \times n}$. In Korollar 4.4.4 werden wir sehen dass die Exponentialabbildung glatt ist. Die höheren Ableitungen werden im Abschnitt 2.5 eingeführt.

Beispiel 2.3.14. Die Abbildung

$$\text{Inv} : \text{GL}(n, \mathbb{R}) \rightarrow \text{GL}(n, \mathbb{R}), \quad \text{Inv}(A) := A^{-1} \quad (2.3.29)$$

ist stetig differenzierbar mit der Ableitung

$$d \text{Inv}(A) \widehat{A} = -A^{-1} \widehat{A} A^{-1} \quad (2.3.30)$$

für $A \in \text{GL}(n, \mathbb{R})$ und $\widehat{A} \in \mathbb{R}^{n \times n}$. **Beweis:** Für $A \in \text{GL}(n, \mathbb{R})$ und $X \in \mathbb{R}^{n \times n}$ mit $\|X\| \|A^{-1}\| < 1$ gilt nach Teil (ii) von Satz 1.4.1

$$A + X \in \text{GL}(n, \mathbb{R}), \quad (A + X)^{-1} = A^{-1} \sum_{k=0}^{\infty} (-XA^{-1})^k.$$

Das heißt $\text{Inv}(A + X) = A^{-1} f(XA^{-1}) = L \circ f \circ R(X)$ für $\|X\| < \|A^{-1}\|^{-1}$, wobei die linearen Abbildungen $L, R : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$ durch $L(Z) := A^{-1}Z$ und $R(X) := XA^{-1}$ und die Abbildung $f : \{Y \in \mathbb{R}^{n \times n} \mid \|Y\| < 1\} \rightarrow \mathbb{R}^{n \times n}$ durch $f(Y) := \sum_{k=0}^{\infty} (-Y)^k$ definiert sind. Nach Satz 2.3.12 ist f stetig differenzierbar mit $df(0)\widehat{Y} = -\widehat{Y}$ für $\widehat{Y} \in \mathbb{R}^{n \times n}$. Damit folgt die Behauptung aus der Kettenregel in Satz 2.3.2 für die Abbildung $L \circ f \circ R$.

2.4 Der Schrankensatz

Wir nehmen in diesem Abschnitt an, dass X und Y endlichdimensionale normierte Vektorräume sind. Für lineare Operatoren $A \in \mathcal{L}(X, Y)$ verwenden wir die stets die Operatornorm

$$\|A\| := \|A\|_{\mathcal{L}(X, Y)} := \sup_{0 \neq x \in X} \frac{\|Ax\|_Y}{\|x\|_X}.$$

(Siehe Abschnitt 1.2.) Der Schrankensatz kann als Erweiterung des Mittelwertsatzes auf differenzierbare Abbildungen zwischen endlichdimensionalen normierten Vektorräumen betrachtet werden.

Satz 2.4.1 (Schrankensatz). *Sei $U \subset X$ offen und $f : U \rightarrow Y$ eine stetig differenzierbare Abbildung.*

(i) *Sei $K \subset U$ eine kompakte konvexe Teilmenge und $c_K := \sup_{x \in K} \|df(x)\|$. Dann gilt*

$$\|f(x_0) - f(x_1)\|_Y \leq c_K \|x_0 - x_1\|_X \quad (2.4.1)$$

für alle $x, x' \in K$.

(ii) *Sei $K \subset U$ eine kompakte Teilmenge. Dann existiert ein $c > 0$ so dass die Ungleichung*

$$\|f(x_0) - f(x_1)\|_Y \leq c \|x_0 - x_1\|_X \quad (2.4.2)$$

für alle $x, x' \in K$ erfüllt ist.

Beweis. Siehe Seite 64. □

Der Beweis von Satz 2.4.1 benötigt die Definition des Integrals einer stetigen Funktion auf einem kompakten Intervall mit Werten in einem endlichdimensionalen normierten Vektorraum X . Im Fall $X = \mathbb{R}^n$ ist die Idee, dass man das Integral einfach Koordinatenweise definiert, was bedeutet, dass das Integral einer stetigen Funktion $x = (x_1, \dots, x_n) : [a, b] \rightarrow \mathbb{R}^n$ der Vektor ist, dessen i te Koordinate durch das Integral der Funktion $x_i : [a, b] \rightarrow \mathbb{R}$ gegeben ist. Ist unser Vektorraum noch nicht mit einer Basis versehen, so kann man für diese Definition eben zunächst eine Basis wählen. Dann ist jedoch zu zeigen, dass die dadurch gegebene Definition des Integrals nicht von der Wahl der Basis abhängt. Eine andere Möglichkeit für die Definition des Integrals einer vektorwertigen Funktion wäre es, dieses erneut über die Konvergenz der Riemannschen Summen zu konstruieren. Will man beide Methoden verwenden, so ist ebenfalls zu zeigen, dass diese zum selben Ergebnis führen. Das ist nicht weiter schwierig und wird im folgenden Lemma bewiesen.

Lemma 2.4.2. *Sei X ein n -dimensionaler normierter Vektorraum mit einer Basis e_1, \dots, e_n , seien $a < b$ reelle Zahlen, und sei $\xi : [a, b] \rightarrow X$ eine stetige Abbildung. Dann sind für jedes $x \in X$ die folgenden Aussagen äquivalent.*

(i) *Für jedes lineare Funktional $\Lambda : X \rightarrow \mathbb{R}$ gilt*

$$\Lambda(x) = \int_a^b \Lambda(\xi(t)) dt. \quad (2.4.3)$$

(ii) *Sind $\xi_1, \dots, \xi_n : [a, b] \rightarrow \mathbb{R}$ die durch $\xi(t) =: \sum_{i=1}^n \xi_i(t)e_i$ für $t \in [a, b]$ definierten stetigen Funktionen, so gilt*

$$x = \sum_{i=1}^n \left(\int_a^b \xi_i(t) dt \right) e_i. \quad (2.4.4)$$

(iii) *Der Vektor x ist der Grenzwert*

$$x = \lim_{N \rightarrow \infty} \sum_{k=1}^N \frac{b-a}{N} \xi \left(a + k \frac{b-a}{N} \right). \quad (2.4.5)$$

Außerdem existiert genau ein Vektor $x \in X$, der diese äquivalenten Bedingungen erfüllt.

Definition 2.4.3. *Der eindeutige Vektor $x \in X$ in Lemma 2.4.2 wird das Integral von ξ über $[a, b]$ genannt und mit $\int_a^b \xi(t) dt := x$ bezeichnet. Das heißt, das Integral ist durch die Bedingung*

$$\Lambda \left(\int_a^b \xi(t) dt \right) = \int_a^b \Lambda(\xi(t)) dt \quad (2.4.6)$$

für jede lineare Abbildung $\Lambda : X \rightarrow \mathbb{R}$ charakterisiert.

Beweis von Lemma 2.4.2. Wir beweisen (i) \implies (ii). Seien $\Lambda_i : X \rightarrow \mathbb{R}$ die durch $\Lambda_i(e_j) := \delta_{ij}$ für $i, j = 1, \dots, n$ definierten linearen Funktionale. Dann gilt $\Lambda_i(\xi(t)) = \Lambda_i(\sum_{j=1}^n \xi_j(t)e_j) = \xi_i(t)$ für alle t und i und daher, nach (i),

$$x = \sum_{i=1}^n \Lambda_i(x)e_i = \sum_{i=1}^n \left(\int_a^b \Lambda_i(\xi(t)) dt \right) e_i = \sum_{i=1}^n \left(\int_a^b \xi_i(t) dt \right) e_i.$$

Damit ist gezeigt, daß der Vektor x die Bedingung (ii) erfüllt.

Wir beweisen (ii) \implies (iii). In der Situation von (ii) gilt

$$x_i := \int_a^b \xi_i(t) dt = \lim_{N \rightarrow \infty} \sum_{k=1}^N \frac{b-a}{N} \xi_i \left(a + k \frac{b-a}{N} \right)$$

für $i = 1, \dots, n$ nach [5, Satz 4.1]. Da $\xi(t) = \sum_{i=1}^n \xi_i(t) e_i$ ist für alle t nach Voraussetzung, und $x = \sum_{i=1}^n x_i e_i$ ist nach (ii), folgt daraus

$$\begin{aligned} x &= \sum_{i=1}^n \left(\lim_{N \rightarrow \infty} \sum_{k=1}^N \frac{b-a}{N} \xi_i \left(a + k \frac{b-a}{N} \right) \right) e_i \\ &= \lim_{N \rightarrow \infty} \sum_{k=1}^N \sum_{i=1}^n \frac{b-a}{N} \xi_i \left(a + k \frac{b-a}{N} \right) e_i \\ &= \lim_{N \rightarrow \infty} \sum_{k=1}^N \frac{b-a}{N} \xi \left(a + k \frac{b-a}{N} \right), \end{aligned}$$

und daher erfüllt der Vektor x die Bedingung (iii).

Wir beweisen (iii) \implies (i). Da jedes lineare Funktional $\Lambda : X \rightarrow \mathbb{R}$ stetig ist nach Korollar 1.2.2, folgt aus (iii) die Gleichung

$$\begin{aligned} \Lambda(x) &= \Lambda \left(\lim_{N \rightarrow \infty} \sum_{k=1}^N \frac{b-a}{N} \xi \left(a + k \frac{b-a}{N} \right) \right) \\ &= \lim_{N \rightarrow \infty} \Lambda \left(\sum_{k=1}^N \frac{b-a}{N} \xi \left(a + k \frac{b-a}{N} \right) \right) \\ &= \lim_{N \rightarrow \infty} \sum_{k=1}^N \frac{b-a}{N} \Lambda \left(\xi \left(a + k \frac{b-a}{N} \right) \right) \\ &= \int_a^b \Lambda(\xi(t)) dt. \end{aligned}$$

Hier folgt die letzte Gleichheit aus [5, Satz 4.1]. Dies zeigt, dass der Vektor x die Bedingung (i) erfüllt.

Die Existenz und Eindeutigkeit eines Vektors $x \in X$, der diese drei äquivalenten Bedingungen erfüllt, folgt sofort aus Teil (ii). Damit ist Lemma 2.4.2 bewiesen. \square

Im folgenden bezeichnen wir für $a < b$ mit $C([a, b], X)$ den Vektorraum der stetigen Funktionen $x : [a, b] \rightarrow X$.

Lemma 2.4.4 (Eigenschaften des Integrals). Für $a < b$ gilt folgendes.

(i) Die Abbildung $C([a, b], X) \rightarrow X : x \mapsto \int_a^b x(t) dt$ ist linear.

(ii) Für alle $x \in C([a, b], X)$ und alle $A \in \mathcal{L}(X, Y)$ gilt

$$\int_a^b Ax(t) dt = A \int_a^b x(t) dt. \quad (2.4.7)$$

(iii) Für alle $x \in C([a, b], X)$ gilt

$$\left\| \int_a^b x(t) dt \right\|_X \leq \int_a^b \|x(t)\|_X dt \quad (2.4.8)$$

(iv) Für alle $x \in C([a, b], X)$ und alle $c \in (a, b)$ gilt

$$\int_a^c x(t) dt + \int_c^b x(t) dt = \int_a^b x(t) dt \quad (2.4.9)$$

(v) Ist $x : [a, b] \rightarrow X$ stetig differenzierbar, so gilt

$$\int_a^b \dot{x}(t) dt = x(b) - x(a). \quad (2.4.10)$$

(vi) Ist $\xi : [a, b] \rightarrow X$ stetig so ist die Funktion

$$x(t) := \int_a^t \xi(s) ds, \quad a \leq t \leq b, \quad (2.4.11)$$

stetig differenzierbar mit $\dot{x}(t) = \xi(t)$ für alle $t \in [0, 1]$.

Beweis. Teil (i) folgt direkt aus der Definition des Integrals.

Für Teil (ii) wählen wir ein lineares Funktional $\Lambda : Y \rightarrow \mathbb{R}$. Dann gilt

$$\Lambda \left(\int_a^b Ax(t) dt \right) = \int_a^b \Lambda(Ax(t)) dt = \Lambda \left(A \int_a^b x(t) dt \right)$$

und damit folgt (ii) aus der Eindeutigkeit in Lemma 2.4.2.

Zum Beweis von (iii) verwenden wir die Ungleichung

$$\left\| \sum_{k=1}^N \frac{b-a}{N} \xi \left(a + k \frac{b-a}{N} \right) \right\| \leq \sum_{k=1}^N \frac{b-a}{N} \left\| \xi \left(a + k \frac{b-a}{N} \right) \right\|$$

für $N \in \mathbb{N}$. Daraus folgt (2.4.8) mit $N \rightarrow \infty$ nach Teil (iii) von Lemma 2.4.2.

Die Aussagen in (iv), (v), und (vi) folgen direkt aus den entsprechenden Aussagen für die Integrale reellwertiger stetiger Funktionen und Teil (i) von Lemma 2.4.2. Damit ist Lemma 2.4.4 bewiesen. \square

Beweis von Satz 2.4.1. Wir beweisen Teil (i). Sei also $K \subset U$ eine kompakte konvexe Teilmenge und sei $c_K := \sup_{x \in K} \|df(x)\|$. Dann ist $c_K < \infty$ da K kompakt ist. Seien $x_0, x_1 \in K$. Dann ist $x_0 + t(x_1 - x_0) \in K$, da K konvex ist. Wir definieren nun die Funktion $y : [0, 1] \rightarrow Y$ durch

$$y(t) := f(x_0 + t(x_1 - x_0)) \quad \text{für } 0 \leq t \leq 1.$$

Diese Funktion ist stetig differenzierbar nach Satz 2.3.2 mit der Ableitung

$$\dot{y}(t) = df(x_0 + t(x_1 - x_0))(x_1 - x_0).$$

Daraus folgt

$$\|\dot{y}(t)\|_Y \leq \|df(x_0 + t(x_1 - x_0))\| \|x_1 - x_0\|_X \leq c_K \|x_1 - x_0\|_X. \quad (2.4.12)$$

Außerdem folgt daraus nach Teil (v) von Lemma 2.4.4 die Gleichung

$$f(x_1) - f(x_0) = y(1) - y(0) = \int_0^1 \dot{y}(t) dt.$$

Daraus wiederum folgt nach Teil (iii) von Lemma 2.4.4 die Ungleichung

$$\|f(x_1) - f(x_0)\|_Y \leq \int_0^1 \|\dot{y}(t)\|_Y dt \leq c_K \|x_1 - x_0\|_X.$$

Hier folgt der letzte Schritt aus (2.4.12) und damit ist Teil (i) bewiesen.

Wir beweisen Teil (ii) mit einem indirekten Argument. Sei also $K \subset U$ eine kompakte Teilmenge, so dass die Behauptung von Teil (ii) nicht gilt. Dann existieren zwei Folgen $(a_k)_{k \in \mathbb{N}}$ und $(b_k)_{k \in \mathbb{N}}$ mit

$$\|f(a_k) - f(b_k)\|_Y > k \|a_k - b_k\|_X \quad \text{für alle } k \in \mathbb{N}. \quad (2.4.13)$$

Da K kompakt ist, können wir durch Übergang zu einer geeigneten Teilfolge annehmen, dass die Grenzwerte

$$a := \lim_{k \rightarrow \infty} a_k, \quad b := \lim_{k \rightarrow \infty} b_k \quad (2.4.14)$$

existieren. Diese Grenzwerte liegen notwendigerweise in K und wir zeigen, dass sie übereinstimmen.

Dazu betrachten wir die Differenz $a_k - b_k$. Nach (2.4.13) gilt

$$\begin{aligned} \|a_k - b_k\|_X &< \frac{1}{k} \|f(a_k) - f(b_k)\|_Y \\ &\leq \frac{1}{k} (\|f(a_k)\|_Y + \|f(b_k)\|_Y) \\ &\leq \frac{2}{k} \sup_{x \in K} \|f(x)\|_Y \end{aligned}$$

Da K kompakt ist, haben wir

$$\sup_{x \in K} \|f(x)\| < \infty$$

und daher

$$\|a - b\|_X = \lim_{k \rightarrow \infty} \|a_k - b_k\|_X = 0.$$

Also gilt

$$a = b \in K. \quad (2.4.15)$$

Wir wählen nun ein $\varepsilon > 0$ mit

$$B := \{x \in X \mid \|x - a\|_X \leq \varepsilon\} \subset U.$$

Dann ist B eine kompakte Teilmenge von U . Daher ist

$$c_B := \sup_{x \in B} \|df(x)\| < \infty$$

und nach Teil (i) gilt

$$\|f(x) - f(x')\|_Y \leq c_B \|x - x'\|_X \quad \text{für alle } x, x' \in B. \quad (2.4.16)$$

Nach (2.4.14) und (2.4.15) existiert eine Konstante $k_0 \in \mathbb{N}$, so dass $a_k, b_k \in B$ sind für alle $k \in \mathbb{N}$ mit $k \geq k_0$. Daraus folgt nach (2.4.13) und (2.4.16) die Ungleichung

$$k \|a_k - b_k\|_X < \|f(a_k) - f(b_k)\|_Y \leq c_B \|a_k - b_k\|_X$$

für jede ganze Zahl $k > k_0$. Für $k \geq \max\{k_0, c_B\}$ ist dies ein Widerspruch. Damit ist Satz 2.4.1 bewiesen. \square

Mit Hilfe des Schrankensatzes sind wir nun in der Lage, die Umkehrung der Beobachtung, dass konstante Funktionen die Ableitung Null haben, zu beweisen. Wir benötigen dazu den Zusammenhangsbegriff im Anhang B.

Satz 2.4.5. *Sei $U \subset X$ eine nichtleere offene zusammenhängende Teilmenge und sei $f : U \rightarrow Y$ eine differenzierbare Abbildung mit $df(x) = 0$ für alle $x \in U$. Dann ist f konstant.*

Beweis. Ist U konvex, so folgt die Aussage direkt aus der Ungleichung (2.4.1) in Satz 2.4.1 mit $c_K = 0$. Andernfalls argumentieren wir wie folgt.

Sei $x_0 \in U$ und $y_0 := f(x_0)$. Zu zeigen ist, dass $f(x) = y_0$ ist für alle $x \in U$. Dazu definieren wir die Menge

$$A := \{x \in U \mid f(x) = y_0\}$$

Diese Menge ist offensichtlich nichtleer, da x_0 ein Element von A ist. Ausserdem ist f stetig nach Lemma 2.1.5, und daher ist $A = f^{-1}(y_0)$ eine U -abgeschlossene Teilmenge von U . Mit anderen Worten, A ist abgeschlossen bezüglich der Relativtopologie von U (siehe Lemma B.1.2).

Wir zeigen nun, dass A eine offene Teilmenge von U ist. Sei also $x \in A$. Da U offen ist existiert eine Konstante $\varepsilon > 0$ mit

$$B_\varepsilon(x) = \{x' \in X \mid \|x - x'\| < \varepsilon\} \subset U.$$

Ist $x' \in B_\varepsilon(x)$, so ist $x + t(x' - x) \in B_\varepsilon(x) \subset U$ für alle $t \in [0, 1]$. Daraus folgt nach Teil (v) von Lemma 2.4.4 die Gleichung

$$\begin{aligned} f(x') - y_0 &= f(x') - f(x) \\ &= \int_0^1 \frac{d}{dt} f(x + t(x' - x)) dt \\ &= \int_0^1 df(x + t(x' - x))(x' - x) dt \\ &= 0 \end{aligned}$$

Hier folgt der erste Schritt aus der Voraussetzung $x \in A$, der zweite Schritt aus der Differenzierbarkeit der Komposition $[0, 1] \rightarrow Y : t \mapsto f(x + t(x' - x))$ in Satz 2.3.2 sowie Teil (v) von Lemma 2.4.4, der dritte Schritt aus der Formel für die Ableitung in Satz 2.3.2, und der letzte Schritt aus der Tatsache, dass die Ableitung von f nach Voraussetzung verschwindet. Damit haben wir gezeigt, dass $B_\varepsilon(x) \subset A$ ist. Also ist A eine nichtleere Teilmenge von U , die sowohl offen als auch abgeschlossen bezüglich der Relativtopologie von U ist. Da U zusammenhängend ist, ist $A = U$, was Satz 2.4.5 beweist. \square

2.5 Höhere Ableitungen

Sei $U \subset \mathbb{R}^n$ eine offene Teilmenge und sei $f : U \rightarrow \mathbb{R}^m$ eine auf ganz U partiell differenzierbare Abbildung. Dann sind die partiellen Ableitungen

$$\partial_i f = \frac{\partial f}{\partial x_i} : U \rightarrow \mathbb{R}^m, \quad i = 1, \dots, n,$$

selbst wieder Abbildungen von U nach \mathbb{R}^m und wir können den Fall betrachten, dass diese wieder partiell differenzierbar sind. Dazu rufen wir zunächst noch einmal an die Definition der partiellen Ableitung nach der i ten Variablen x_i in Erinnerung. An der Stelle $x = (x_1, \dots, x_n) \in U$ ist diese durch

$$\begin{aligned} \partial_i f(x) &= \frac{\partial f}{\partial x_i}(x) \\ &= \lim_{t \rightarrow 0} \frac{f(x + te_i) - f(x)}{t} \\ &= \left. \frac{d}{dt} \right|_{t=0} f(x + te_i) \\ &= \left. \frac{d}{dt} \right|_{t=0} f(x_1, \dots, x_{i-1}, x_i + t, x_{i+1}, \dots, x_n) \end{aligned}$$

gegeben, wobei $e_i \in \mathbb{R}^n$ den i ten Standardbasisvektor bezeichnet. Das heißt, man erhält die i te partielle Ableitung an der Stelle x dadurch, dass man die Koordinaten x_j für $j \neq i$ festhält, und die dadurch entstehende Funktion von einer reellen Variablen an der Stelle x_i differenziert.

Definition 2.5.1. Eine Abbildung $f : U \rightarrow \mathbb{R}^m$ auf einer offenen Teilmenge $U \subset \mathbb{R}^n$ heißt **zweimal partiell differenzierbar** wenn sie partiell differenzierbar ist und ihre partiellen Ableitungen $\partial_j f : U \rightarrow \mathbb{R}^m$ für $j = 1, \dots, n$ wieder partiell differenzierbar sind. Ist eine solche Abbildung $f : U \rightarrow \mathbb{R}^m$ zweimal partiell differenzierbar, so werden die durch

$$\begin{aligned} \partial_i \partial_j f(x) &:= \frac{\partial^2 f}{\partial x_i \partial x_j}(x) \\ &:= \frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_j}(x) \\ &:= \left. \frac{d}{dt} \right|_{t=0} \frac{\partial f}{\partial x_j}(x_1, \dots, x_{i-1}, x_i + t, x_{i+1}, \dots, x_n) \end{aligned} \tag{2.5.1}$$

für $x \in U$ und $i, j = 1, \dots, n$ definierten Abbildungen $\partial_i \partial_j f : U \rightarrow \mathbb{R}^m$ die **zweiten partiellen Ableitungen von f** genannt.

An diese Definition schließt sich natürlicherweise die Frage an, welchen Effekt die Vertauschung der Reihenfolge auf die partiellen Ableitungen hat. Hierzu ein Beispiel.

Beispiel 2.5.2. Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ die durch $f(0, 0) = 0$ und

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

für $x^2 + y^2 \neq 0$ definierte Funktion. Diese Funktion ist partiell differenzierbar. Ihre partiellen Ableitungen verschwinden an der Stelle $x = y = 0$, und sind durch

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= \frac{(3x^2y - y^3)(x^2 + y^2) - 2x(x^3y - xy^3)}{(x^2 + y^2)^2} \\ &= \frac{3x^4y + 3x^2y^3 - x^2y^3 - y^5 - 2x^4y + 2x^2y^3}{(x^2 + y^2)^2} \\ &= \frac{x^4y + 4x^2y^3 - y^5}{(x^2 + y^2)^2} \\ \frac{\partial f}{\partial y}(x, y) &= \frac{x^5 - 4x^3y^2 - xy^4}{(x^2 + y^2)^2}. \end{aligned}$$

für $x^2 + y^2 \neq 0$ gegeben. Hier folgt die zweite Formel aus der ersten durch Vertauschung von x und y . Die partiellen Ableitungen $\partial f/\partial x$ und $\partial f/\partial y$ sind (auch an der Stelle $x = y = 0$) stetig, und daher ist f stetig differenzierbar. Insbesondere gilt

$$\frac{\partial f}{\partial x}(x, 0) = 0, \quad \frac{\partial f}{\partial x}(0, y) = -y, \quad \frac{\partial f}{\partial y}(x, 0) = x, \quad \frac{\partial f}{\partial y}(0, y) = 0.$$

Daraus folgt, dass f (auch an der Stelle $x = y = 0$) zweimal partiell differenzierbar ist mit

$$\frac{\partial^2 f}{\partial x \partial x}(0, 0) = 0, \quad \frac{\partial^2 f}{\partial y \partial x}(0, 0) = -1, \quad \frac{\partial^2 f}{\partial x \partial y}(0, 0) = 1, \quad \frac{\partial^2 f}{\partial y \partial y}(0, 0) = 0.$$

Die gemischten zweiten partiellen Ableitungen stimmen also an der Stelle $x = y = 0$ nicht überein. Darüber hinaus ergibt die Berechnung der zweiten partiellen Ableitungen auf $\mathbb{R}^2 \setminus \{(0, 0)\}$, dass diese an der Stelle $x = y = 0$ nicht stetig sind (Übung).

Der folgende Satz von Hermann Amandus Schwarz² zeigt, dass die Asymmetrie der gemischten zweiten partiellen Ableitungen in Beispiel 2.5.2 ein Sonderfall ist, der nicht auftreten kann, wenn die zweiten partiellen Ableitungen nicht nur überall existieren, sondern auch stetig sind.

Satz 2.5.3 (Schwarz). *Sei $U \subset \mathbb{R}^n$ eine offene Teilmenge, sei $f : U \rightarrow \mathbb{R}^m$ eine partiell differenzierbare Abbildung, und seien Indizes $i, j \in \{1, \dots, n\}$ mit $i \neq j$ und ein Element $x_0 \in U$ gegeben. Wir nehmen an, dass die Abbildung $\partial f / \partial x_j : U \rightarrow \mathbb{R}^m$ nach der Variablen x_i überall partiell differenzierbar ist, und dass die Abbildung*

$$\frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_j} : U \rightarrow \mathbb{R}^m$$

an der Stelle x_0 stetig ist. Dann ist die Abbildung $\partial f / \partial x_i : U \rightarrow \mathbb{R}^m$ nach der Variablen x_j an der Stelle x_0 partiell differenzierbar und es gilt

$$\frac{\partial}{\partial x_j} \frac{\partial f}{\partial x_i}(x_0) = \frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_j}(x_0).$$

Beweis. Nach Lemma 2.1.15 genügt es, den Fall $m = 1$ zu betrachten. Außerdem betreffen die Behauptungen des Satzes nur die Einschränkung der Abbildung f auf die Teilmenge von U in der alle Variablen bis auf x_i und x_j festgehalten sind. Daher genügt es ebenfalls, den Fall $n = 2$ zu betrachten. Wir nehmen also neu an, dass $U \subset \mathbb{R}^2$ eine offenen Menge ist und verwenden die Bezeichnung (x, y) für die Variablen in U . Weiter nehmen wir an, dass $f : U \rightarrow \mathbb{R}$ eine partiell differenzierbare Funktion ist, dass die zweite partielle Ableitung

$$\frac{\partial^2 f}{\partial x \partial y}(x, y) = \lim_{h \rightarrow 0} \frac{1}{h} \left(\frac{\partial f}{\partial y}(x + h, y) - \frac{\partial f}{\partial y}(x, y) \right) \quad (2.5.2)$$

für alle $(x, y) \in U$ existiert, und dass die Funktion $\frac{\partial^2 f}{\partial x \partial y} : U \rightarrow \mathbb{R}$ an der Stelle $(x_0, y_0) \in U$ stetig ist. Ihren Wert an dieser Stelle bezeichnen wir mit

$$a_0 := \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0). \quad (2.5.3)$$

² Hermann Amandus Schwarz war ein Schüler von Ernst Eduard Kummer und Karl Weierstraß. Er war von 1869 bis 1875 Professor an der ETH Zürich und ihm wurde im Jahr 1914 die Ehrendoktorwürde der ETH Zürich verliehen.

Zu zeigen ist die Gleichung

$$\lim_{h \rightarrow 0} \frac{1}{h} \left(\frac{\partial f}{\partial x}(x_0, y_0 + h) - \frac{\partial f}{\partial x}(x_0, y_0) \right) = a_0. \quad (2.5.4)$$

Sei $\varepsilon > 0$ gegeben. Da die Funktion $\partial^2 f / \partial x \partial y : U \rightarrow \mathbb{R}$ an der Stelle (x_0, y_0) stetig ist und dort den Wert a_0 hat, existiert eine Konstante $\delta > 0$, so dass für alle $(x, y) \in \mathbb{R}^n$ folgende Aussage gilt:

$$\begin{aligned} |x - x_0| < \delta, \\ |y - y_0| < \delta \end{aligned} \quad \implies \quad (x, y) \in U \text{ und } \left| \frac{\partial^2 f}{\partial x \partial y}(x, y) - a_0 \right| < \varepsilon. \quad (2.5.5)$$

Mit dieser Wahl von δ beweisen wir in drei Schritten, dass jedes $y \in \mathbb{R}$ mit $0 < |y - y_0| < \delta$ die Ungleichung

$$\left| \frac{\partial f}{\partial x}(x_0, y) - \frac{\partial f}{\partial x}(x_0, y_0) - a_0(y - y_0) \right| \leq \varepsilon |y - y_0| \quad (2.5.6)$$

erfüllt. Da $\varepsilon > 0$ beliebig gewählt war, wäre damit die gewünschte Gleichung (2.5.4) bewiesen.

Schritt 1. Für alle $x, y \in \mathbb{R}$ mit $0 < |x - x_0| < \delta$ und $|y - y_0| < \delta$ gilt

$$\left| \frac{\partial f}{\partial y}(x, y) - \frac{\partial f}{\partial y}(x_0, y) - a_0(x - x_0) \right| < \varepsilon |x - x_0|. \quad (2.5.7)$$

Wir halten zwei reelle Zahlen x, y fest mit

$$0 < |x - x_0| < \delta, \quad |y - y_0| < \delta.$$

Wir nehmen zunächst an, dass $x_0 < x < x_0 + \delta$ ist. Dann ist $(\xi, y) \in U$ für alle $\xi \in [x_0, x]$ nach (2.5.5) und die Funktion

$$[x_0, x] \rightarrow \mathbb{R} : \xi \mapsto \frac{\partial f}{\partial y}(\xi, y)$$

ist nach Voraussetzung überall differenzierbar. Ihre Ableitung an der Stelle $\xi \in [x_0, x]$ ist die Zahl $\partial^2 f / \partial x \partial y(\xi, y)$. Nach dem Mittelwertsatz existiert daher eine Zahl $x_0 < \xi < x$ mit

$$\frac{\partial^2 f}{\partial x \partial y}(\xi, y) = \frac{\frac{\partial f}{\partial y}(x, y) - \frac{\partial f}{\partial y}(x_0, y)}{x - x_0}.$$

Nun bilden wir die Differenz mit a_0 und nehmen den Betrag. Mit $|\xi - x_0| < \delta$ und $|y - y_0| < \delta$ folgt dann nach (2.5.5) die gewünschte Ungleichung (2.5.7). Im Fall $x < x_0$ argumentiert man genauso mit der gleichen Funktion auf dem Intervall $[x, x_0]$. Damit ist Schritt 1 bewiesen.

Schritt 2. Für alle $x, y \in \mathbb{R}$ mit $0 < |x - x_0| < \delta$ und $0 < |y - y_0| < \delta$ gilt

$$\left| \frac{f(x, y) - f(x_0, y) - f(x, y_0) + f(x_0, y_0)}{(x - x_0)(y - y_0)} - a_0 \right| < \varepsilon. \quad (2.5.8)$$

Wir halten zwei reelle Zahlen x, y fest mit

$$0 < |x - x_0| < \delta, \quad 0 < |y - y_0| < \delta.$$

Wir nehmen zunächst an, dass $y_0 < y < y_0 + \delta$ ist. Dann ist $(x, \eta) \in U$ für alle $\eta \in [y_0, y]$ nach (2.5.5) und wir definieren $g : [y_0, y] \rightarrow \mathbb{R}$ durch

$$g(\eta) := f(x, \eta) - f(x_0, \eta) \quad \text{für } y_0 \leq \eta \leq y.$$

Diese Funktion ist nach Voraussetzung differenzierbar. Nach dem Mittelwertsatz existiert daher eine Zahl $y_0 < \eta < y$ mit

$$g'(\eta) = \frac{g(y) - g(y_0)}{y - y_0}.$$

Nach Definition von g läßt sich diese Gleichung in der Form

$$\frac{\partial f}{\partial y}(x, \eta) - \frac{\partial f}{\partial y}(x_0, \eta) = \frac{f(x, y) - f(x_0, y) - f(x, y_0) + f(x_0, y_0)}{y - y_0}$$

schreiben. Nun teilen wir durch $x - x_0$, bilden die Differenz mit a_0 , und nehmen den Betrag. Mit $0 < |x - x_0| < \delta$ und $|\eta - y_0| < \delta$ folgt dann nach (2.5.7) in Schritt 1 die gewünschte Ungleichung (2.5.8). Im Fall $y < y_0$ argumentiert man genauso mit der gleichen Funktion g auf dem Intervall $[y, y_0]$. Damit ist Schritt 2 bewiesen.

Schritt 3. Jedes $y \in \mathbb{R}$ mit $0 < |y - y_0| < \delta$ erfüllt (2.5.6).

Wir halten eine reelle Zahl y mit $0 < |y - y_0| < \delta$ fest und schreiben die Ungleichung (2.5.8) in Schritt 2 in der Form

$$\left| \frac{f(x, y) - f(x_0, y)}{x - x_0} - \frac{f(x, y_0) - f(x_0, y_0)}{x - x_0} - a_0(y - y_0) \right| < \varepsilon |y - y_0|$$

für alle $x \in \mathbb{R}$ mit $0 < |x - x_0| < \delta$. Mit dem Grenzübergang $x \rightarrow x_0$ ergibt sich daraus die gewünschte Ungleichung

$$\left| \frac{\partial f}{\partial x}(x_0, y) - \frac{\partial f}{\partial x}(x_0, y_0) - a_0(y - y_0) \right| \leq \varepsilon |y - y_0|.$$

Damit sind Schritt 3 und Satz 2.5.3 bewiesen. \square

Definition 2.5.4. Seien $\ell, m, n \in \mathbb{N}$ und sei $U \subset \mathbb{R}^n$ eine offene Teilmenge.

(i) Eine Abbildung $f : U \rightarrow \mathbb{R}^m$ heißt **zweimal stetig differenzierbar** wenn sie stetig differenzierbar ist und ihre partiellen Ableitungen

$$\partial_i f : U \rightarrow \mathbb{R}^m, \quad i = 1, \dots, n,$$

stetig differenzierbar sind.

(ii) Eine Abbildung $f : U \rightarrow \mathbb{R}^m$ heißt **ℓ -mal stetig differenzierbar**, wenn sie stetig differenzierbar ist und ihre partiellen Ableitungen

$$\partial_i f : U \rightarrow \mathbb{R}^m, \quad i = 1, \dots, n,$$

$(\ell - 1)$ -mal stetig differenzierbar sind.

(iii) Die Menge der ℓ -mal stetig differenzierbare Abbildungen von U nach \mathbb{R}^m bezeichnen wir mit

$$C^\ell(U, \mathbb{R}^m) := \{f : U \rightarrow \mathbb{R}^m \mid f \text{ ist } \ell\text{-mal stetig differenzierbar}\}.$$

Die Elemente von $C^\ell(U, \mathbb{R}^m)$ werden auch **C^ℓ -Abbildungen** genannt.

Definition 2.5.5 (Höhere partielle Ableitungen). Seien $\ell, m, n \in \mathbb{N}$, sei $U \subset \mathbb{R}^n$ eine offene Teilmenge, und sei $f : U \rightarrow \mathbb{R}^m$ eine C^ℓ -Abbildung. Für $i_1, \dots, i_\ell \in \{1, \dots, n\}$ wird die **partielle Ableitung**

$$\frac{\partial^\ell f}{\partial x_{i_1} \cdots \partial x_{i_\ell}} : U \rightarrow \mathbb{R}^m$$

der Ordnung ℓ induktiv definiert durch

$$\frac{\partial^\ell f}{\partial x_{i_1} \partial x_{i_2} \cdots \partial x_{i_\ell}} := \frac{\partial}{\partial x_{i_1}} \frac{\partial^{\ell-1} f}{\partial x_{i_2} \cdots \partial x_{i_\ell}}.$$

Bemerkung 2.5.6 (C^ℓ -Abbildungen). Nach Definition 2.5.4 und Definition 2.5.5 ist eine Abbildung $f : U \rightarrow \mathbb{R}^m$ genau dann ℓ -mal stetig differenzierbar wenn alle ihre partiellen Ableitungen bis zur Ordnung ℓ existieren und stetig sind. Dies bedeutet, dass für jede Indexfolge $i_1, \dots, i_\ell \in \{1, \dots, n\}$ (der Länge ℓ) erstens die partielle Ableitung $\partial f / \partial x_{i_\ell}$ überall existiert und stetig ist, zweitens diese wieder nach $x_{i_{\ell-1}}$ partiell differenzierbar ist und die daraus resultierende zweite partielle Ableitung $\partial^2 f / \partial x_{i_{\ell-1}} \partial x_{i_\ell}$ stetig ist, drittens diese dann nach $x_{i_{\ell-2}}$ partiell differenzierbar ist und die daraus resultierende dritte partielle Ableitung $\partial^3 f / \partial x_{i_{\ell-2}} \partial x_{i_{\ell-1}} \partial x_{i_\ell}$ stetig ist, und so weiter bis hin zur partiellen Ableitung $\partial^\ell f / \partial x_{i_1} \cdots \partial x_{i_\ell}$ der Ordnung ℓ , die ebenfalls stetig ist.

Bemerkung 2.5.7 (Multiindizes). Da die Reihenfolge der partiellen Ableitungen einer C^ℓ -Abbildung nach Satz 2.5.3 das Resultat nicht beeinflusst, ist es sinnvoll, für die partiellen Ableitungen eine vereinfachende Schreibweise einzuführen, die Wiederholungen unter den Indizes i_1, \dots, i_ℓ berücksichtigt. Da es nur darauf ankommt, wie oft ein Index $i \in \{1, \dots, n\}$ in einer solchen Indexfolge auftritt, ordnen wir jedem solchen Index eine *Häufigkeit* α_i zu. Diese ist ein Element der Menge

$$\mathbb{N}_0 := \mathbb{N} \cup \{0\} = \{0, 1, 2, 3, \dots\}.$$

Wir erhalten damit einen **Multiindex**

$$\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{N}_0^n.$$

Die **Ordnung** eines solchen Multiindex ist die Zahl

$$|\alpha| := \alpha_1 + \alpha_2 + \dots + \alpha_n \in \mathbb{N}_0.$$

Für $f \in C^{|\alpha|}(U, \mathbb{R}^m)$ wird die partielle Ableitung $\partial^\alpha f : U \rightarrow \mathbb{R}^m$ durch

$$\begin{aligned} \partial^\alpha f &:= \frac{\partial^{|\alpha|} f}{\partial x^\alpha} := \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_n^{\alpha_n}} \\ &:= \frac{\partial^{|\alpha|} f}{\partial x_1 \dots \partial x_1 \partial x_2 \dots \partial x_2 \dots \partial x_n \dots \partial x_n} \end{aligned} \quad (2.5.9)$$

definiert. In dem letzten Ausdruck tritt der Term ∂x_i genau α_i -mal auf, wobei α_i auch Null sein kann. Per Definition ist $\partial^{(0, \dots, 0)} f := f$.

Definition 2.5.8 (Glatte Abbildungen). Seien $m, n \in \mathbb{N}$ und sei $U \subset \mathbb{R}^m$ eine offenen Menge. Eine Abbildung $f : U \rightarrow \mathbb{R}^m$ heißt **glatt** wenn sie beliebig oft stetig differenzierbar ist, das heißt, wenn sie für jedes $\ell \in \mathbb{N}$ zu der Menge $C^\ell(U, \mathbb{R}^m)$ gehört und damit ℓ -mal stetig differenzierbar ist. Eine glatte Abbildung $f : U \rightarrow \mathbb{R}^m$ wird auch **C^∞ -Abbildung** genannt. Die Menge der glatten Abbildungen von U nach \mathbb{R}^m wird mit

$$C^\infty(U, \mathbb{R}^m) := \{f : U \rightarrow \mathbb{R}^m \mid f \text{ ist glatt}\} = \bigcap_{\ell \in \mathbb{N}} C^\ell(U, \mathbb{R}^m)$$

bezeichnet.

Der folgende Satz fasst einige wichtige Eigenschaften von C^ℓ -Abbildungen zusammen, die sich direkt mit vollständiger Induktion aus den Eigenschaften differenzierbarer Abbildungen herleiten lassen.

Satz 2.5.9. *Sei $\ell \in \mathbb{N} \cup \{\infty\}$, seien X, Y, Z endlichdimensionale normierte Vektorräume, und seien $U \subset X$ und $V \subset Y$ offen. Dann gilt folgendes.*

- (i) *Eine Abbildung $f = (f_1, \dots, f_m) : U \rightarrow \mathbb{R}^m$ (mit $m \in \mathbb{N}$) ist genau dann C^ℓ , wenn $f_i : U \rightarrow \mathbb{R}$ für jedes $i \in \{1, \dots, m\}$ eine C^ℓ -Funktion ist.*
- (ii) *Sind $f, g \in C^\ell(U, \mathbb{R})$ so ist $f + g \in C^\ell(U, \mathbb{R})$.*
- (iii) *Sind $f, g \in C^\ell(U, \mathbb{R})$ so ist $fg \in C^\ell(U, \mathbb{R})$.*
- (iv) *Sind $f, g \in C^\ell(U, \mathbb{R})$ und $g(x) \neq 0$ für alle $x \in U$, so ist $f/g \in C^\ell(U, \mathbb{R})$.*
- (v) *Sind $f : U \rightarrow V$ und $g : V \rightarrow Z$ zwei C^ℓ -Abbildungen, so ist auch die Komposition $g \circ f : U \rightarrow Z$ eine C^ℓ -Abbildung.*
- (vi) *Ist $U \subset \mathbb{R}^n$ offen und $f_k : U \rightarrow \mathbb{R}$ eine Folge von C^ℓ -Funktionen, so dass die Funktionenfolge $(\partial^\alpha f_k)_{k \in \mathbb{N}}$ für alle $\alpha \in \mathbb{N}_0^n$ mit $|\alpha| \leq \ell$ auf jeder kompakten Teilmenge von U gleichmäßig konvergiert, so ist $f := \lim_{k \rightarrow \infty} f_k : U \rightarrow \mathbb{R}$ eine C^ℓ -Funktion und es gilt $\partial^\alpha f = \lim_{k \rightarrow \infty} \partial^\alpha f_k$ für alle $\alpha \in \mathbb{N}_0^n$ mit $|\alpha| \leq \ell$.*

Beweis. Wir können annehmen, dass $X = \mathbb{R}^n$ ist, und damit die Ableitung von f als Abbildung $df : U \rightarrow \mathbb{R}^{m \times n}$ betrachten, die jedem Element $x \in U$ die Jacobi-Matrix $df(x) \in \mathbb{R}^{m \times n}$ zuordnet. Nun folgt Teil (i) für $\ell = 1$ direkt aus Lemma 2.1.15. Ist $\ell \geq 2$ und gilt Teil (i) für $C^{\ell-1}$ -Abbildungen, so schließen wir daraus, dass $df : U \rightarrow \mathbb{R}^{m \times n}$ genau dann eine $C^{\ell-1}$ -Abbildung ist, wenn $\partial_i f_j : U \rightarrow \mathbb{R}$ für alle i, j eine $C^{\ell-1}$ Abbildung ist, was wiederum dazu äquivalent ist, dass $df_i : U \rightarrow \mathbb{R}^{1 \times n}$ für jedes i eine $C^{\ell-1}$ -Abbildung ist. Damit folgt Teil (i) durch vollständige Induktion über ℓ .

Teil (ii) folgt aus Teil (i) von Satz 2.3.1 durch vollständige Induktion über ℓ mit $\partial_i(f + g) = \partial_i f + \partial_i g$.

Genauso verhält es sich mit dem Beweis von Teil (iii) unter Verwendung von Teil (ii) und der Leibnizregel $\partial_i(fg) = g\partial_i f + f\partial_i g$ in Satz 2.3.1.

Auch Teil (iv) läßt sich nun mit Hilfe der bereits bewiesenen Teile (ii) und (iii) und der Formel $\partial_i(f/g) = (g\partial_i f - f\partial_i g)/g^2$ in Satz 2.3.1 durch vollständige Induktion über ℓ beweisen.

Mit $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$, $Z = \mathbb{R}^\ell$ verwendet das Induktionsargument für den Beweis von Teil (v) die Kettenregel $\partial_i(g_k \circ f) = \sum_{j=1}^m ((\partial_j g_k) \circ f) \partial_i f_j$ aus Satz 2.3.2 sowie die Teile (i), (ii) und (iii).

Teil (vi) folgt mittels vollständiger Induktion aus Satz 2.2.6 und damit ist Satz 2.5.9 bewiesen. \square

Übung 2.5.10. Was für Formeln ergeben sich aus der Leibnizregel und der Kettenregel für die höheren partiellen Ableitungen von Produkt und Komposition?

Beispiel 2.5.11. Jeder Multiindex $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ bestimmt eine glatte Funktion $f_\alpha : \mathbb{R}^n \rightarrow \mathbb{R}$, die durch

$$f_\alpha(x) := x^\alpha := x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$$

für $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ definiert ist. Für die partiellen Ableitungen von f_α nach x_1 erhalten wir die üblichen Formeln

$$\begin{aligned} \frac{\partial f_\alpha}{\partial x_1}(x) &= \alpha_1 x_1^{\alpha_1-1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}, \\ \frac{\partial^2 f_\alpha}{\partial x_1^2}(x) &= \alpha_1(\alpha_1 - 1) x_1^{\alpha_1-2} x_2^{\alpha_2} \cdots x_n^{\alpha_n}, \\ &\vdots \\ \frac{\partial^{\beta_1} f_\alpha}{\partial x_1^{\beta_1}}(x) &= \alpha_1(\alpha_1 - 1) \cdots (\alpha_1 - \beta_1 + 1) x_1^{\alpha_1 - \beta_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n} \\ &= \frac{\alpha_1!}{(\alpha_1 - \beta_1)!} x_1^{\alpha_1 - \beta_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n} \end{aligned}$$

für $0 \leq \beta_1 \leq \alpha_1$. Für die partiellen Ableitungen nach den anderen Variablen erhält man natürlich die gleichen Formeln. Insgesamt ergibt sich daraus mit

$$\alpha! := \alpha_1! \cdot \alpha_2! \cdots \alpha_n!$$

die Gleichung

$$\partial^\beta f_\alpha(x) = \frac{\alpha!}{(\beta - \alpha)!} x^{\alpha - \beta} \quad (2.5.10)$$

für jeden Multiindex $\beta = (\beta_1, \dots, \beta_n) \in \mathbb{N}_0^n$ mit $\beta_i \leq \alpha_i$ für $i = 1, \dots, n$. Ist andererseits $\beta \in \mathbb{N}_0^n$ ein Multiindex mit $\beta_i > \alpha_i$ für ein i so ist $\partial^\beta f_\alpha \equiv 0$. Insbesondere folgt daraus die Formel

$$\partial^\beta f_\alpha(0) = \begin{cases} \alpha!, & \text{falls } \beta = \alpha, \\ 0, & \text{falls } \beta \neq \alpha, \end{cases} \quad (2.5.11)$$

für die partiellen Ableitungen von f_α an der Stelle $x = 0$.

2.6 Die Taylorreihe

Für Multiindizes $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ und Vektoren $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ hat es sich in Bemerkung 2.5.7 und Beispiel 2.5.11 als nützlich herausgestellt, die Abkürzungen

$$\begin{aligned} |\alpha| &:= \alpha_1 + \alpha_2 + \dots + \alpha_n, \\ \alpha! &:= \alpha_1! \cdot \alpha_2! \cdot \dots \cdot \alpha_n!, \\ x^\alpha &:= x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n} \end{aligned} \tag{2.6.1}$$

einzuführen. Hier verwenden wir im Fall $\alpha_i = 0$ die üblichen Konventionen $0! := 1$ und $s^0 := 1$ für jede reelle Zahl s . Nun sei eine Abbildung

$$\mathbb{N}_0^n \rightarrow \mathbb{R}^m : \alpha \mapsto a_\alpha$$

gegeben, die jedem Multiindex $\alpha \in \mathbb{N}_0^n$ einen Vektor $a_\alpha \in \mathbb{R}^m$ zuordnet derart, dass nur endlich viele a_α von Null verschieden sind. Dann definiert die Formel

$$f(x) := \sum_{\alpha \in \mathbb{N}_0^n} a_\alpha x^\alpha \tag{2.6.2}$$

für $x \in \mathbb{R}^n$ eine glatte Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, da es sich hier um eine endliche Summe handelt und jeder Summand eine glatte Abbildung ergibt. Jede Abbildung dieser Form heißt **Polynom**. Die Gleichung (2.5.10) in Beispiel 2.5.11 zeigt, dass die partiellen Ableitungen des Polynoms (2.6.2) durch

$$\partial^\beta f(x) = \sum_{\alpha_i \geq \beta_i} \frac{\alpha!}{(\alpha - \beta)!} a_\alpha x^{\alpha - \beta} \tag{2.6.3}$$

für $x \in \mathbb{R}^n$ und $\beta \in \mathbb{N}_0^n$ gegeben sind, wobei die Summe über alle Multiindizes $\alpha \in \mathbb{N}_0^n$ mit $\alpha_i \geq \beta_i$ für $i = 1, \dots, n$ zu verstehen ist. Wertet man die Gleichung (2.6.3) an der Stelle $x = 0$ aus so ergibt sich

$$a_\alpha = \frac{\partial^\alpha f(0)}{\alpha!} \tag{2.6.4}$$

für alle $\alpha \in \mathbb{N}_0^n$. Damit lassen sich die Koeffizienten des Polynoms f aus den partiellen Ableitungen an der Stelle $x = 0$ zurückgewinnen. Da die Abbildung $x \mapsto f(x_0 + x)$ ebenfalls ein Polynom ist, erhalten wir die Gleichung

$$f(x) = \sum_{\alpha \in \mathbb{N}_0^n} \frac{\partial^\alpha f(x_0)}{\alpha!} (x - x_0)^\alpha \tag{2.6.5}$$

für jedes Polynom f und alle $x, x_0 \in \mathbb{R}^n$. Dies führt zur folgenden Definition.

Definition 2.6.1 (Taylorpolynom). Sei $U \subset \mathbb{R}^n$ eine offene Teilmenge, sei $x_0 \in U$, und sei $f \in C^\ell(U, \mathbb{R}^m)$ mit $\ell \in \mathbb{N}_0$. Das **Taylorpolynom der Ordnung ℓ von f an der Stelle x_0** ist das durch

$$(T_{x_0}^\ell f)(x) := \sum_{|\alpha| \leq \ell} \frac{\partial^\alpha f(x_0)}{\alpha!} (x - x_0)^\alpha \quad (2.6.6)$$

für $x \in \mathbb{R}^n$ definierte Polynom $T_{x_0}^\ell f : \mathbb{R}^n \rightarrow \mathbb{R}^m$. In (2.6.6) ist die Summe über alle Multiindizes $\alpha \in \mathbb{N}_0^n$ mit $|\alpha| \leq \ell$ zu verstehen.

Ist f ein Polynom, so zeigt die Gleichung (2.6.5), dass f mit all seinen Taylorpolynomen der entsprechenden Ordnung übereinstimmt. Im allgemeinen ist das nicht der Fall. Jedoch ist das Taylorpolynom der Ordnung ℓ einer Funktion f an der Stelle x_0 in einem noch zu beschreibenden Sinne die *beste Näherung* an f in der Nähe von x_0 . Dazu ist es nützlich, für die Differenz einer Funktion zu ihrem Taylorpolynom eine geeignete Formel zu finden.

Satz 2.6.2 (Das Restglied in der Taylorentwicklung). Sei $U \subset \mathbb{R}^n$ eine offene Teilmenge und sei $f \in C^{\ell+1}(U, \mathbb{R}^m)$ mit $\ell \in \mathbb{N}_0$. Sind $x, \xi \in \mathbb{R}^n$ mit $x + t\xi \in U$ für $0 \leq t \leq 1$, so gilt

$$f(x+\xi) - \sum_{|\alpha| \leq \ell} \frac{\partial^\alpha f(x)}{\alpha!} \xi^\alpha = \int_0^1 (\ell+1)(1-t)^\ell \sum_{|\alpha|=\ell+1} \frac{\partial^\alpha f(x+t\xi)}{\alpha!} \xi^\alpha dt. \quad (2.6.7)$$

Proof. Siehe Seite 81. □

Korollar 2.6.3. Sei $f \in C^{\ell+1}(U, \mathbb{R}^m)$ wie in Satz 2.6.2, sei $x_0 \in U$, wähle eine Zahl $r > 0$ mit $\overline{B}_r(x_0) \subset U$, und definiere

$$c_{\ell+1} := \sup_{x \in \overline{B}_r(x_0)} \sum_{|\alpha|=\ell+1} \frac{\|\partial^\alpha f(x)\|}{\alpha!} \quad (2.6.8)$$

Dann erfüllt jedes Element $x \in \overline{B}_r(x_0)$ die Ungleichung

$$\|f(x) - (T_{x_0}^\ell f)(x)\| \leq c_{\ell+1} \|x - x_0\|^{\ell+1}. \quad (2.6.9)$$

Beweis. Nach Satz 2.6.2 und Lemma 2.4.4 gilt

$$\begin{aligned} \|f(x) - (T_{x_0}^\ell f)(x)\| &\leq \int_0^1 (\ell+1)(1-t)^\ell \sum_{|\alpha|=\ell+1} \frac{\|\partial^\alpha f(x_0+t\xi)\|}{\alpha!} |\xi^\alpha| dt \\ &\leq \int_0^1 (\ell+1)(1-t)^\ell dt c_{\ell+1} \|\xi\|^{\ell+1} = c_{\ell+1} \|\xi\|^{\ell+1}. \end{aligned}$$

Hier haben wir die Ungleichung $|\xi^\alpha| \leq \|\xi\|^{|\alpha|}$ verwendet. □

Für die Abschätzung (2.6.9) in Korollar 2.6.3 ist es von Bedeutung, dass die Abbildung f $(\ell + 1)$ -mal stetig differenzierbar ist. Ist f nur ℓ -mal stetig differenzierbar, so erhält man eine schwächere Abschätzung für die Approximationseigenschaft des Taylorpolynoms. Diese genügt, um das Taylorpolynom damit eindeutig zu charakterisieren (Bemerkung 2.6.7).

Korollar 2.6.4. *Seien $f \in C^\ell(U, \mathbb{R}^m)$ und $x_0 \in U$ wie in Definition 2.6.1. Dann gilt*

$$\lim_{x \rightarrow x_0} \frac{\|f(x) - (T_{x_0}^\ell f)(x)\|}{\|x - x_0\|^\ell} = 0. \quad (2.6.10)$$

Beweis. Für $\xi \in \mathbb{R}^n$ mit

$$x_0 + t\xi \in U \quad \text{für } 0 \leq t \leq 1,$$

und $x := x_0 + \xi$ folgt aus Satz 2.6.2 die Gleichung

$$\begin{aligned} f(x) - (T_{x_0}^\ell f)(x) &= f(x) - (T_{x_0}^{\ell-1} f)(x) - \sum_{|\alpha|=\ell} \frac{\partial^\alpha f(x_0)}{\alpha!} \xi^\alpha \\ &= \int_0^1 \ell(1-t)^{\ell-1} \sum_{|\alpha|=\ell} \frac{\partial^\alpha f(x_0 + t\xi) - \partial^\alpha f(x_0)}{\alpha!} \xi^\alpha dt. \end{aligned}$$

Nun sei $\varepsilon > 0$ gegeben. Da alle partiellen Ableitungen von f der Ordnung ℓ stetig sind, existiert ein $\delta > 0$ mit $B_\delta(x_0) \subset U$ und

$$x \in B_\delta(x_0) \quad \implies \quad \sum_{|\alpha|=\ell} \frac{\|\partial^\alpha f(x) - \partial^\alpha f(x_0)\|}{\alpha!} < \varepsilon.$$

Für $x \in B_\delta(x_0)$ und $\xi := x - x_0$ erhalten wir daher unter Verwendung von Lemma 2.4.4 die Ungleichung

$$\begin{aligned} \|f(x) - (T_{x_0}^\ell f)(x)\| &\leq \int_0^1 \ell(1-t)^{\ell-1} \sum_{|\alpha|=\ell} \frac{\|\partial^\alpha f(x_0 + t\xi) - \partial^\alpha f(x_0)\|}{\alpha!} |\xi^\alpha| dt \\ &\leq \int_0^1 \ell(1-t)^{\ell-1} dt \varepsilon \|\xi\|^\ell \\ &= \varepsilon \|x - x_0\|^\ell. \end{aligned}$$

Damit ist Korollar 2.6.3 bewiesen. \square

Bemerkung 2.6.5. Im Fall $\ell = 1$ ist

$$(T_{x_0}^1 f)(x_0 + \xi) = f(x_0) + \sum_{i=1}^n \partial_i f(x_0) \xi_i = f(x_0) + df(x_0)\xi.$$

Daher läßt sich die Gleichung (2.6.10) in der Form

$$\lim_{\xi \rightarrow 0} \frac{\|f(x_0 + \xi) - f(x_0) - df(x_0)\xi\|}{\|\xi\|} = 0$$

schreiben. Dies ist genau die Definition der Differenzierbarkeit.

Bemerkung 2.6.6. Sei $p : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ein Polynom der Form

$$p(x) = \sum_{|\alpha| \leq \ell} a_\alpha x^\alpha$$

für $x \in \mathbb{R}^n$. Erfüllt p die Bedingung

$$\lim_{x \rightarrow 0} \frac{\|p(x)\|}{\|x\|^\ell} = 0, \quad (2.6.11)$$

so sind alle Koeffizienten des Polynoms gleich null. **Beweis:** Im Fall $\ell = 0$ ist das Polynom konstant, und nach (2.6.11) mit $\ell = 0$ kann diese Konstante nur Null sein. Sei nun $\ell \geq 1$ und die Behauptung für $\ell - 1$ erfüllt. Definiere

$$q(x) := \sum_{|\alpha| \leq \ell-1} a_\alpha x^\alpha \quad \text{für } x \in \mathbb{R}^n.$$

Dann gilt $\|q(x) - p(x)\| \leq c \|x\|^\ell$ für alle $x \in \mathbb{R}^n$ und eine geeignete Konstante $c > 0$. Daraus folgt $\lim_{x \rightarrow 0} \|x\|^{1-\ell} \|q(x)\| = 0$ und daher ist $q \equiv 0$ nach Induktionsannahme. Das heißt, p ist ein homogenes Polynom vom Grade ℓ und erfüllt daher die Bedingung $p(tx) = t^\ell p(x)$ für alle $x \in \mathbb{R}^n$ und $t \in \mathbb{R}$. Daher folgt aus (2.6.11) für jedes $x \in \mathbb{R}^n \setminus \{0\}$ die Gleichung

$$0 = \lim_{t \rightarrow 0} \frac{\|p(tx)\|}{\|tx\|^\ell} = \lim_{t \rightarrow 0} \frac{|t|^\ell \|p(x)\|}{|t|^\ell \|x\|^\ell} = \frac{\|p(x)\|}{\|x\|^\ell}.$$

Also ist $p \equiv 0$ und daher $a_\alpha = 0$ für alle α nach (2.6.4).

Bemerkung 2.6.7. Es folgt aus Bemerkung 2.6.6, dass das Taylorpolynom $T_{x_0}^\ell f$ einer C^ℓ -Abbildung $f : U \rightarrow \mathbb{R}^m$ durch die Bedingung (2.6.10) in Korollar 2.6.4 eindeutig bestimmt ist. (Für $\ell = 1$ siehe auch Bemerkung 2.6.5 und Lemma 2.1.3.)

Der Beweis von Satz 2.6.2 beruht auf den beiden folgenden Lemmas. Das erste ist ein Spezialfall der Kettenregel für höhere Ableitungen. Das zweite ist ein Resultat aus Analysis I und liefert die Integralform des Restgliedes für die Taylorentwicklung in einer reellen Variablen.

Lemma 2.6.8. *Seien $f \in C^\ell(U, \mathbb{R}^m)$ und $x_0 \in U$ wie in Definition 2.6.1. Dann gilt*

$$\left. \frac{d^k}{dt^k} \right|_{t=0} f(x_0 + t\xi) = \sum_{|\alpha|=k} \frac{k!}{\alpha!} \partial^\alpha f(x_0) \xi^\alpha \quad (2.6.12)$$

für alle $\xi \in \mathbb{R}^n$ und $k = 0, 1, \dots, \ell$.

Beweis. Für $k = 0$ sind beide Seiten der Gleichung $f(x_0)$ und für $k = 1$ ist dies die Kettenregel

$$\left. \frac{d}{dt} \right|_{t=0} f(x_0 + t\xi) = \sum_{i=1}^n \partial_i f(x_0) \xi_i = df(x_0) \xi.$$

Sei nun die Gleichung (2.6.12) für ein $k \in \{1, \dots, \ell - 1\}$ erfüllt. Dann folgt

$$\begin{aligned} \left. \frac{d^{k+1}}{dt^{k+1}} \right|_{t=0} f(x_0 + t\xi) &= \left. \frac{d}{dt} \right|_{t=0} \left. \frac{d^k}{dt^k} f(x_0 + t\xi) \right. \\ &= \left. \frac{d}{dt} \right|_{t=0} \sum_{|\alpha|=k} \frac{k!}{\alpha!} \partial^\alpha f(x_0 + t\xi) \xi^\alpha \xi_i \\ &= \sum_{|\alpha|=k} \frac{k!}{\alpha!} \sum_{i=1}^n \partial^{\alpha+e_i} f(x_0) \xi^{\alpha+e_i} \\ &= \sum_{|\beta|=k+1} \left(\sum_{\substack{i=1 \\ \beta_i \geq 1}}^n \frac{k!}{(\beta - e_i)!} \right) \partial^\beta f(x_0) \xi^\beta \\ &= \sum_{|\beta|=k+1} \frac{(k+1)!}{\beta!} \partial^\beta f(x_0) \xi^\beta. \end{aligned}$$

Hier folgt die letzte Gleichheit aus der Formel

$$\frac{|\beta|}{\beta_1! \cdots \beta_n!} = \sum_{\substack{i=1 \\ \beta_i \geq 1}}^n \frac{1}{\beta_1! \cdots \beta_{i-1}! (\beta_i - 1)! \beta_{i+1}! \cdots \beta_n!},$$

Damit ist Lemma 2.6.8 bewiesen. \square

Lemma 2.6.9. Sei $u : [0, 1] \rightarrow \mathbb{R}^m$ eine $C^{\ell+1}$ -Funktion. Dann gilt

$$u(1) - \sum_{k=0}^{\ell} \frac{u^{(k)}(0)}{k!} = \int_0^1 \frac{(1-t)^\ell}{\ell!} u^{(\ell+1)}(t) dt. \quad (2.6.13)$$

Beweis. Für $\ell = 0$ ist (2.6.13) der Hauptsatz der Infinitesimalrechnung. Ist $\ell \geq 1$, und gilt die Formel für $\ell - 1$, so erhalten wir mit

$$\phi(t) := -\frac{(1-t)^\ell}{\ell!}, \quad \psi(t) := u^{(\ell)}(t) - u^{(\ell)}(0)$$

die Gleichungskette

$$\begin{aligned} u(1) - \sum_{k=0}^{\ell} \frac{u^{(k)}(0)}{k!} &= \int_0^1 \frac{(1-t)^{\ell-1}}{(\ell-1)!} u^{(\ell)}(t) dt - \frac{u^{(\ell)}(0)}{\ell!} \\ &= \int_0^1 \frac{(1-t)^{\ell-1}}{(\ell-1)!} (u^{(\ell)}(t) - u^{(\ell)}(0)) dt \\ &= \int_0^1 \phi'(t) \psi(t) dt \\ &= - \int_0^1 \phi(t) \psi'(t) dt \\ &= \int_0^1 \frac{(1-t)^\ell}{\ell!} u^{(\ell+1)}(t) dt. \end{aligned}$$

Damit ist Lemma 2.6.9 bewiesen. \square

Beweis von Satz 2.6.2. Sei $u(t) := f(x + t\xi)$ für $0 \leq t \leq 1$. Dann ist

$$u^{(k)}(t) = \sum_{|\alpha|=k} \frac{k!}{\alpha!} \partial^\alpha f(x + t\xi) \xi^\alpha$$

nach Lemma 2.6.8. Daraus folgt (unter Verwendung von Lemma 2.6.9)

$$\begin{aligned} f(x + \xi) - \sum_{|\alpha| \leq \ell} \frac{\partial^\alpha f(x)}{\alpha!} \xi^\alpha &= u(1) - \sum_{k=0}^{\ell} \frac{1}{k!} u^{(k)}(0) \\ &= \int_0^1 \frac{(1-t)^\ell}{\ell!} u^{(\ell+1)}(t) dt \\ &= \int_0^1 (\ell+1)(1-t)^\ell \sum_{|\alpha|=\ell+1} \frac{\partial^\alpha f(x + t\xi)}{\alpha!} \xi^\alpha dt. \end{aligned}$$

Damit ist Satz 2.6.2 bewiesen. \square

Definition 2.6.10 (Taylorreihe).

Seien $m, n \in \mathbb{N}$, sei $U \subset \mathbb{R}^n$ eine offene Teilmenge, sei $f \in C^\infty(U, \mathbb{R}^m)$, und sei $x_0 \in U$. Die **Taylorreihe von f an der Stelle x_0** ist die formale Reihe

$$(T_{x_0}^\infty f)(x) := \sum_{\alpha \in \mathbb{N}_0^n} \frac{\partial^\alpha f(x_0)}{\alpha!} (x - x_0)^\alpha \quad (2.6.14)$$

für $x \in \mathbb{R}^n$.

Bei dieser Definition sind zwei Dinge zu beachten. Erstens handelt es sich bei der rechten Seite der Gleichung (2.6.14) um einen rein formalen Ausdruck, und es ist damit nichts ausgesagt über die Konvergenz. In der Tat kann man schon für $n = 1$ glatte Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$ konstruieren, deren Taylorreihen an der Stelle $x_0 = 0$ den Konvergenzradius Null haben, und damit für jedes $x \neq 0$ divergieren (siehe Teil (ii) von Beispiel 2.6.11). Zweitens ist der Ausdruck auf der rechten Seite von (2.6.14) streng genommen keine „Reihe“, da die Indexmenge \mathbb{N}_0^n nicht die Menge der natürlichen Zahlen ist, sondern eine (möglicherweise nichtsummierbare) Familie von Vektoren im \mathbb{R}^m wie in [4]. Dennoch hat sich der Begriff „Taylorreihe“ in der Literatur durchgesetzt. Man erhält auch tatsächlich eine Reihe, wenn man zunächst alle Summanden der Ordnung k zu einer endlichen Summe zusammengefasst, und dann über $k \in \mathbb{N}_0$ summiert. Der relevante Konvergenzbegriff ist jedoch die absolute Summierbarkeit in (2.6.14).

Daran schließen sich nun zwei Fragen an, nämlich erstens ob sich Bedingungen formulieren lassen, unter denen die Taylorreihe auf einer offenen Umgebung des Punktes x_0 absolut und gleichmäßig summierbar ist, und zweitens, wenn Konvergenz vorliegt, ob die dadurch definierte Funktion $T_{x_0}^\infty f$ in einer Umgebung des Punktes x_0 mit f übereinstimmt. Auch hier gibt es bereits im Fall $n = 1$ wichtige Gegenbeispiele (siehe Teil (i) von Beispiel 2.6.11). Ein Kriterium für eine positive Antwort auf beide Fragen liefert die Abschätzung in Korollar 2.6.3 (siehe Korollar 2.6.13).

Beispiel 2.6.11. (i) Die durch $f(x) := e^{-1/x^2}$ für $x \neq 0$ und $f(0) := 0$ definierte Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ ist glatt und hat die Taylorreihe $(T_0^\infty f)(x) = 0$.

(ii) Die durch $f(x) := \int_0^\infty (1 + x^2 t)^{-1} e^{-t} dt$ definierte Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ ist glatt und hat die Taylorreihe $(T_0^\infty f)(x) = \sum_{k=0}^\infty (-1)^k k! x^k$.

Im folgenden Satz sind $n, m \in \mathbb{N}$ und wir verwenden die Bezeichnungen

$$B_r := \{x \in \mathbb{R}^n \mid \|x\| < r\}, \quad \bar{B}_r := \{x \in \mathbb{R}^n \mid \|x\| \leq r\}$$

für $r > 0$.

Satz 2.6.12 (Analytische Abbildungen).

Sei eine Abbildung $\mathbb{N}_0^n \rightarrow \mathbb{R}^m : \alpha \mapsto a_\alpha$ gegeben mit

$$\rho := \frac{1}{\limsup_{k \rightarrow \infty} c_k^{1/k}} > 0, \quad c_k := \sum_{|\alpha|=k} \|a_\alpha\|. \quad (2.6.15)$$

Dann gilt folgendes.

(i) Die Reihe

$$f(x) := \sum_{\alpha \in \mathbb{N}_0^n} a_\alpha x^\alpha \quad (2.6.16)$$

konvergiert absolut für alle $x \in B_\rho$ und die Konvergenz ist gleichmäßig auf jeder kompakten Teilmenge von B_ρ .

(ii) Die Abbildung $f : B_\rho \rightarrow \mathbb{R}^m$ in (i) ist glatt mit den partiellen Ableitungen

$$\partial^\beta f(x) := \sum_{\alpha_i \geq \beta_i} \frac{\alpha!}{(\alpha - \beta)!} a_\alpha x^{\alpha - \beta} \quad (2.6.17)$$

für $x \in B_\rho$ und $\beta \in \mathbb{N}_0^n$. Auch die Reihe (2.6.17) konvergiert für jedes $\beta \in \mathbb{N}_0^n$ und jedes $x \in B_\rho$ absolut, und für festes β ist die Konvergenz gleichmäßig auf jeder kompakten Teilmenge von B_ρ .

(iii) Die Taylorreihe von f an der Stelle $x_0 = 0$ konvergiert absolut und gleichmäßig auf jeder kompakten Teilmenge von B_ρ und stimmt mit der Funktion f überein, das heißt $(T_0^\infty f)(x) = f(x)$ für alle $x \in B_\rho$.

Beweis. Der Beweis hat vier Schritte.

Schritt 1. Sei $p \in \mathbb{N}_0$ und $0 < r < \rho$. Dann gilt

$$\sum_{k=p}^{\infty} k^{np} c_k r^{k-p} < \infty. \quad (2.6.18)$$

Da $0 < r < \rho$ ist, gilt $r \limsup_{k \rightarrow \infty} c_k^{1/k} < 1$. Da die Folge $k^{np/k} r^{-p/k}$ gegen 1 konvergiert für $k \rightarrow \infty$, folgt daraus $\limsup_{k \rightarrow \infty} k^{np/k} c_k^{1/k} r^{1-p/k} < 1$. Nun wählen wir eine reelle Zahl λ mit

$$\limsup_{k \rightarrow \infty} k^{np/k} c_k^{1/k} r^{1-p/k} < \lambda < 1.$$

Dann existiert ein $k_0 \in \mathbb{N}$, so dass jedes $k \in \mathbb{N}$ mit $k \geq k_0$ die Ungleichung

$$k^{np/k} c_k^{1/k} r^{1-p/k} \leq \lambda$$

erfüllt. Das ergibt $k^{np} c_k r^{k-p} \leq \lambda^k$ für alle $k \in \mathbb{N}$ mit $k \geq k_0$ und daher folgt Schritt 1 aus dem Majorantenkriterium für Reihen reeller Zahlen.

Schritt 2. Wir beweisen Teil (i).

Für $k \in \mathbb{N}_0$ sei $f_k : \mathbb{R}^n \rightarrow \mathbb{R}$ das homogene Polynom

$$f_k(x) := \sum_{|\alpha|=k} a_\alpha x^\alpha \quad (2.6.19)$$

der Ordnung k . Sei $r > 0$. Dann gilt

$$\|f_k(x)\| \leq \sum_{|\alpha|=k} \|a_\alpha\| |x^\alpha| \leq \sum_{|\alpha|=k} \|a_\alpha\| \|x\|^k \leq c_k r^k$$

für alle $x \in \overline{B}_r$ und alle $k \in \mathbb{N}_0$. Ist $0 < r < \rho$, so gilt $\sum_{k=0}^{\infty} c_k r^k < \infty$ nach Schritt 1, und daraus folgt nach Satz 1.3.10 die absolute und gleichmäßige Konvergenz der Reihe $f = \sum_{k=1}^{\infty} f_k$ auf \overline{B}_r . Damit ist Schritt 2 bewiesen.

Schritt 3. Wir beweisen Teil (ii).

Für $k \in \mathbb{N}_0$ sei $f_k : \mathbb{R}^n \rightarrow \mathbb{R}$ das homogene Polynom in (2.6.19). Sei $\beta \in \mathbb{N}_0^n$ und $p := |\beta|$. Dann ist $\partial^\beta f_k \equiv 0$ für $k = 0, 1, \dots, p-1$, und für $k \geq p$ ist

$$\partial^\beta f_k(x) := \sum_{\substack{|\alpha|=k \\ \alpha_i \geq \beta_i}} \frac{\alpha!}{(\beta - \alpha)!} a_\alpha x^{\alpha - \beta} \quad (2.6.20)$$

für alle $x \in \mathbb{R}^n$ nach (2.6.3). Da $\frac{\alpha!}{(\beta - \alpha)!} \leq k^{np}$ ist für alle $\alpha \in \mathbb{N}_0^n$ mit $|\alpha| = k$ und $\alpha_i \geq \beta_i$ für $i = 1, \dots, n$, folgt daraus die Ungleichung

$$\sup_{x \in \overline{B}_r} \|\partial^\beta f_k(x)\| \leq k^{np} \sup_{x \in \overline{B}_r} \sum_{\substack{|\alpha|=k \\ \alpha_i \geq \beta_i}} \|a_\alpha\| \|x\|^{|\alpha - \beta|} \leq k^{np} c_k r^{k-p}$$

für alle $k \in \mathbb{N}_0$ mit $k \geq p$. Daraus folgt nach Schritt 1 und Satz 1.3.10 die absolute und gleichmäßige Konvergenz der Reihe $g_\beta := \sum_{k=p}^{\infty} \partial^\beta f_k$ auf \overline{B}_r für alle $0 < r < \rho$. Nach Satz 2.2.6 folgt daraus durch vollständige Induktion, dass $f : B_\rho \rightarrow \mathbb{R}^m$ eine glatte Abbildung mit den partiellen Ableitungen $\partial^\beta f = g_\beta$ ist. Damit ist Schritt 3 bewiesen.

Schritt 4. Wir beweisen Teil (iii).

Die Gleichung (2.6.17) in Teil (ii) ergibt die Formel $a_\alpha = \partial^\alpha f(0)/\alpha!$ für alle $\alpha \in \mathbb{N}_0^n$. Daraus folgt Schritt 4 und damit ist Satz 2.6.12 bewiesen. \square

Korollar 2.6.13 (Konvergenz der Taylorreihe).

Seien $m, n \in \mathbb{N}$, sei $U \subset \mathbb{R}^n$ eine offene Teilmenge, sei $f \in C^\infty(U, \mathbb{R}^m)$, und seien $x_0 \in U$ und $r > 0$ mit $\overline{B}_r(x_0) \subset U$ und

$$r < \rho := \frac{1}{\limsup_{k \rightarrow \infty} c_k^{1/k}}, \quad c_k := \sup_{x \in \overline{B}_r(x_0)} \sum_{|\alpha|=k} \frac{\|\partial^\alpha f(x)\|}{\alpha!} \quad (2.6.21)$$

Dann konvergiert die Taylorreihe (2.6.14), samt all ihrer partiellen Ableitungen, absolut und gleichmäßig auf $\overline{B}_r(x_0)$ und es gilt $f(x) = (T_{x_0}^\infty f)(x)$ für alle $x \in \overline{B}_r(x_0)$.

Beweis. Nach Korollar 2.6.3 gilt die Ungleichung

$$\sup_{x \in \overline{B}_r(x_0)} \|f(x) - (T_{x_0}^k f)(x)\| \leq c_{k+1} r^{k+1}$$

für alle $k \in \mathbb{N}$. Da $\sum_{k=0}^{\infty} c_k r^k < \infty$ ist nach (2.6.21) und Schritt 1 im Beweis von Satz 2.6.12, haben wir $\lim_{k \rightarrow \infty} c_{k+1} r^{k+1} = 0$, und daraus folgt

$$f(x) = \lim_{k \rightarrow \infty} (T_{x_0}^k f)(x) = (T_{x_0}^\infty f)(x)$$

für alle $x \in \overline{B}_r(x_0)$ und die Konvergenz ist gleichmäßig. Die absolute und gleichmäßige Konvergenz der Taylorpolynome $T_{x_0}^k f$ und aller ihrer partiellen Ableitungen auf dem Ball $\overline{B}_r(x_0)$ folgt aus Satz 2.6.12 mit $a_\alpha := \partial^\alpha f(x_0)/\alpha!$. Damit ist Korollar 2.6.13 bewiesen. \square

Definition 2.6.14 (Analytische Abbildungen).

Seien $m, n \in \mathbb{N}$ und sei $U \subset \mathbb{R}^n$ eine offene Teilmenge. Eine glatte Abbildung $f \in C^\infty(U, \mathbb{R}^m)$ heißt **(reell) analytisch** wenn für jedes $x_0 \in U$ eine reelle Zahl $r > 0$ mit $\overline{B}_r(x_0) \subset U$ existiert, die (2.6.21) erfüllt.

Nach Korollar 2.6.13 konvergieren die Taylorreihen einer analytischen Abbildung f samt aller ihrer partiellen Ableitungen auf hinreichend kleinen Bällen absolut und gleichmäßig gegen f , beziehungsweise deren partielle Ableitungen. Umgekehrt ist jede Funktion wie in Satz 2.6.12 reell analytisch. Das Studium der reell analytischen Funktionen ist ein wichtiges Teilgebiet der Mathematik, das tief hineinführt in Fragen der Geometrie, Algebra, und Analysis. Dieses Thema kann und soll hier nicht weiter vertieft werden. Statt dessen betrachten wir an dieser Stelle nur zwei elementare Beispiele.

Beispiel 2.6.15. Im Fall $m = 1$ wird die Zahl ρ in (2.6.15) der **Konvergenzradius** der Reihe f in (2.6.16) genannt. Im Gegensatz zum Fall $n = 1$ kann es hier jedoch durchaus sein, dass die Taylorreihe f in (2.6.16) auf einem viel größeren Gebiet als nur B_ρ konvergiert. Als Beispiel betrachten wir die Funktion

$$f(x) := \frac{1}{1 - x_1 - x_2 - \cdots - x_n} \quad (2.6.22)$$

auf der offenen Menge

$$U := \{x \in \mathbb{R}^n \mid |x_1 + x_2 + \cdots + x_n| < 1\}.$$

Diese Funktion läßt sich als geometrische Reihe

$$\begin{aligned} f(x) &= \sum_{k=0}^{\infty} (x_1 + x_2 + \cdots + x_n)^k \\ &= \sum_{k=0}^{\infty} \sum_{|\alpha|=k} \frac{k!}{\alpha!} x^\alpha \\ &= \sum_{\alpha \in \mathbb{N}_0^n} \frac{|\alpha|!}{\alpha!} x^\alpha \end{aligned} \quad (2.6.23)$$

darstellen. Hier konvergiert die Reihe in der ersten Formel auf der rechten Seite absolut für alle x in dem unbeschränkten Gebiet U , während der letzte Ausdruck nur für $x \in \mathbb{R}^n$ mit $\|x\| < 1$ eine summierbare Familie darstellt. Dieser letzte Term ist auch die Taylorreihe von f an der Stelle $x_0 = 0$ und ihr Konvergenzradius ist $\rho = 1$ (Übung).

Beispiel 2.6.16. Sei $U := \{(x, y) \in \mathbb{R}^2 \mid x > 0\}$ und sei $f : U \rightarrow \mathbb{R}$ die Funktion

$$f(x, y) := x^y = e^{y \log(x)}. \quad (2.6.24)$$

Die Taylorpolynome der Ordnung 2 und 3 von f an den Stellen $x_0 = (1, 0)$ und $x_0 = (1, 1)$ sind

$$\begin{aligned} (T_{(1,0)}^2 f)(x, y) &= 1 + (x - 1)y, \\ (T_{(1,0)}^3 f)(x, y) &= 1 + (x - 1)y - \frac{1}{2}(x - 1)^2 y, \\ (T_{(1,1)}^2 f)(x, y) &= 1 + (x - 1) + (x - 1)(y - 1), \\ (T_{(1,1)}^3 f)(x, y) &= 1 + (x - 1) + (x - 1)(y - 1) + \frac{1}{2}(x - 1)^2 (y - 1). \end{aligned}$$

(Beweis: Übung.)

2.7 Lokale Extrema

In diesem Abschnitt bezeichnet $\|x\| = \sqrt{x_1^2 + \cdots + x_n^2}$ stets die Euklidische Norm eines Vektors $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, und für $x \in \mathbb{R}^n$ und $r > 0$ bezeichnet $B_r(x) := \{y \in \mathbb{R}^n \mid \|x - y\| < r\}$ den offenen Ball im \mathbb{R}^n bezüglich der Euklidischen Norm.

Definition 2.7.1 (Lokales Extremum). Sei $U \subset \mathbb{R}^n$ eine offene Teilmenge und sei $f : U \rightarrow \mathbb{R}$ eine stetige Funktion. Ein Element $x_0 \in U$ heißt

- **lokales Minimum von f** , wenn ein $\delta > 0$ existiert mit

$$x \in B_\delta(x_0) \quad \Longrightarrow \quad x \in U \text{ und } f(x) \geq f(x_0),$$

- **striktes lokales Minimum von f** , wenn ein $\delta > 0$ existiert mit

$$x \in B_\delta(x_0) \setminus \{x_0\} \quad \Longrightarrow \quad x \in U \text{ und } f(x) > f(x_0),$$

- **lokales Maximum von f** , wenn ein $\delta > 0$ existiert mit

$$x \in B_\delta(x_0) \quad \Longrightarrow \quad x \in U \text{ und } f(x) \leq f(x_0),$$

- **striktes lokales Maximum von f** , wenn ein $\delta > 0$ existiert mit

$$x \in B_\delta(x_0) \setminus \{x_0\} \quad \Longrightarrow \quad x \in U \text{ und } f(x) < f(x_0).$$

Ein Element $x_0 \in U$ heißt **lokales Extremum von f** , wenn es ein lokales Minimum oder ein lokales Maximum von f ist.

Satz 2.7.2. Sei $U \subset \mathbb{R}^n$ eine offene Teilmenge, sei $f : U \rightarrow \mathbb{R}$ eine stetige Funktion, und sei $x_0 \in U$ ein lokales Extremum von f . Ist f an der Stelle x_0 differenzierbar, so gilt

$$df(x_0) = 0.$$

Beweis. Sei x_0 ein lokales Minimum von f , sei $\delta > 0$ wie in Definition 2.7.1, sei $\xi \in \mathbb{R}^n \setminus \{0\}$, und definiere die Funktion $\phi : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$ durch $\varepsilon := \delta / \|\xi\|$ und $\phi(t) := f(x_0 + t\xi)$ für $-\varepsilon < t < \varepsilon$. Dann ist ϕ stetig, besitzt an der Stelle $t = 0$ ein lokales Minimum, und ist an der Stelle $t = 0$ differenzierbar nach Satz 2.3.2. Also ist $\phi'(0) = 0$ nach einem Satz aus Analysis I.

Hier ist nochmals das Argument aus Analysis I. Für $t \geq 0$ erhält man

$$\frac{\phi(-t) - \phi(0)}{-t} \leq 0 \leq \frac{\phi(t) - \phi(0)}{t}$$

und daraus folgt

$$\phi'(0) = \lim_{t \searrow 0} \frac{\phi(-t) - \phi(0)}{-t} \leq 0 \leq \lim_{t \searrow 0} \frac{\phi(t) - \phi(0)}{t} = \phi'(0).$$

Daher gilt nach der Kettenregel $df(x)\xi = \phi'(0) = 0$. Im Falle eines lokalen Maximums argumentiert man genauso oder man erhält das Resultat indem man f durch $-f$ ersetzt. Damit ist Satz 2.7.2 bewiesen. \square

Definition 2.7.3 (Kritischer Punkt). Sei $U \subset \mathbb{R}^n$ eine offene Teilmenge und sei $f : U \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion. Ein Element $x_0 \in U$ heißt **kritischer Punkt von f** wenn $df(x_0) = 0$ ist.

Mit dieser Terminologie sagt Satz 2.7.2, dass jedes lokale Extremum einer stetig differenzierbaren Funktion ein kritischer Punkt ist. Wie wir bereits aus der Analysis I wissen, ist aber nicht jeder kritische Punkt ein lokales Extremum. Zum Beispiel ist $x_0 = 0$ ein kritischer Punkt, aber kein lokales Extremum, der Funktion $f(x) = x^3$ in einer reellen Variablen. Hier sind einige Beispiele in der Dimension zwei.

Beispiel 2.7.4. Die durch $f(x, y) := \sqrt{x^2 + y^2}$ für $(x, y) \in \mathbb{R}^2$ definierte stetige Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ hat im Nullpunkt $(x_0, y_0) = (0, 0)$ ein lokales Minimum, ist an genau dieser Stelle aber nicht differenzierbar. Also ist Satz 2.7.2 auf dieses Beispiel nicht anwendbar.

Beispiel 2.7.5. Die Funktion

$$f(x, y) := x^2 + y^2$$

hat im Nullpunkt ein lokales Minimum und dies ist auch die einzige Stelle, an der die Ableitung von f verschwindet. Die Funktion $f(x, y) := -x^2 - y^2$ hat im Nullpunkt ein lokales Maximum und dies ist ebenfalls der einzige kritische Punkt von f . Die Funktion

$$f(x, y) = xy$$

hat ihren einzigen kritischen Punkt wieder im Nullpunkt, der in diesem Beispiel jedoch kein lokales Extremum ist (siehe Abbildung 2.2). Dies ist ein Beispiel eines sogenannten *Sattelpunktes*.

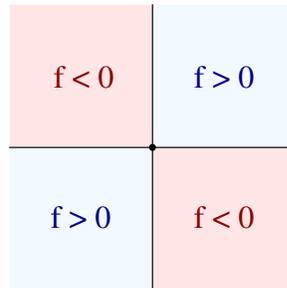


Abbildung 2.2: Die kritische Niveaumenge eines Sattelpunktes.

Beispiel 2.7.6. Die Funktion

$$f(x, y) = xy(x + y)$$

hat ihren einzigen kritischen Punkt ebenfalls im Nullpunkt und dies ist kein lokales Extremum. Die Funktion $f_\lambda(x, y) = (xy - \lambda)(x + y)$ hat für $\lambda > 0$ genau zwei kritische Punkte an den Stellen $\pm(\sqrt{\lambda}, \sqrt{\lambda})$ und auch für $\lambda < 0$ genau zwei kritische Punkte an den Stellen $\pm(\sqrt{|\lambda|}, -\sqrt{|\lambda|})$. Für $\lambda \neq 0$ sind dies Sattelpunkte. Für $\lambda = 0$ handelt es sich um einen sogenannten *degenerierten* kritischen Punkt (siehe Abbildung 2.3).

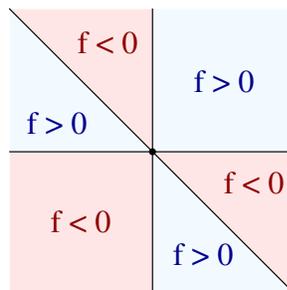


Abbildung 2.3: Ein degenerierter kritischer Punkt.

Beispiel 2.7.7. Die Funktion

$$f_\lambda(x, y) := x^3 - 3\lambda x + y^2$$

hat für $\lambda < 0$ keinen kritischen Punkt, für $\lambda = 0$ genau einen kritischen Punkt an der Stelle $(0, 0)$, und für $\lambda > 0$ genau zwei kritische Punkte an den Stellen $(-\sqrt{\lambda}, 0)$ (Sattelpunkt) und $(\sqrt{\lambda}, 0)$ (lokales Minimum).

Beispiel 2.7.8. Die Funktion $f(x, y) := \sin(2\pi x) + \sin(2\pi y)$ auf dem \mathbb{R}^2 hat unendlich viele kritische Punkte. Betrachtet man sie jedoch als Funktion auf dem *Torus* $\mathbb{T}^2 := \mathbb{R}^2/\mathbb{Z}^2$, indem man zwei Elemente (x, y) und (x', y') im \mathbb{R}^2 miteinander identifiziert wenn ihre Differenz $(x' - x, y' - y) \in \mathbb{Z}^2$ ein Vektor mit ganzzahligen Koordinaten ist, so erhält man eine Funktion mit genau vier kritischen Punkten (einem lokalen Minimum, einem lokalen Maximum, und zwei Sattelpunkten).

Für ein besseres Verständnis dieser und vieler anderer Beispiele ist es nützlich, die *Hessematrix* an einem kritischen Punkt zu betrachten.

Definition 2.7.9 (Hessematrix). Sei $U \subset \mathbb{R}^n$ eine offene Teilmenge und sei $f : U \rightarrow \mathbb{R}$ eine C^2 -Funktion. Für $x \in U$ wird die Matrix

$$d^2f(x) := \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(x) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(x) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(x) & \frac{\partial^2 f}{\partial x_2 \partial x_2}(x) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(x) \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \frac{\partial^2 f}{\partial x_n \partial x_2}(x) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n}(x) \end{pmatrix} \in \mathbb{R}^{n \times n} \quad (2.7.1)$$

der zweiten partiellen Ableitungen die **Hessematrix von f an der Stelle x** genannt. Nach Satz 2.5.3 ist die Hessematrix symmetrisch.

Beispiel 2.7.10. Die Hessematrix der Funktion $f(x, y) = x^2 + y^2$ im Nullpunkt (einem lokalen Minimum) ist die positiv definite Matrix

$$d^2f(x) := \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

Für $f(x, y) = -x^2 - y^2$ erhält man im Nullpunkt (einem lokalen Maximum) eine negativ definite Hessematrix, und für $f(x, y) = xy$ erhält man im Nullpunkt (einem Sattelpunkt) die indefinite Hessematrix

$$d^2f(x) := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Dies ist kein Zufall, wie wir im nächsten Satz sehen werden. Die Hessematrix der Funktion $f(x, y) = xy(x + y)$ im Nullpunkt (einem degenerierten kritischen Punkt) ist die Nullmatrix.

Satz 2.7.11 (Kriterien für lokale Extrema). Sei $U \subset \mathbb{R}^n$ eine offene Teilmenge, sei $f : U \rightarrow \mathbb{R}$ eine zweimal stetig differenzierbare Funktion, und sei $x_0 \in U$ ein kritischer Punkt von f . Dann gilt folgendes.

(i) Ist x_0 ein lokales Minimum, so ist die Hessematrix $d^2f(x_0)$ positiv semi-definit, das heißt, für alle $\xi \in \mathbb{R}^n$ gilt $\xi^T d^2f(x_0)\xi \geq 0$.

(ii) Ist die Hessematrix $d^2f(x_0)$ positiv definit, das heißt, für alle $\xi \in \mathbb{R}^n$ mit $\xi \neq 0$ gilt $\xi^T d^2f(x_0)\xi > 0$, so ist x_0 ein striktes lokales Minimum.

(iii) Ist x_0 ein lokales Maximum, so ist die Hessematrix $d^2f(x_0)$ negativ semi-definit, das heißt, für alle $\xi \in \mathbb{R}^n$ gilt $\xi^T d^2f(x_0)\xi \leq 0$.

(iv) Ist die Hessematrix $d^2f(x_0)$ negativ definit, das heißt, für alle $\xi \in \mathbb{R}^n$ mit $\xi \neq 0$ gilt $\xi^T d^2f(x_0)\xi < 0$, so ist x_0 ein striktes lokales Maximum.

Beweis. Wir beweisen Teil (i). Sei $\xi := (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$ gegeben. Dann existiert eine Zahl $\varepsilon > 0$, so dass $x_0 + t\xi \in U$ ist für alle $t \in \mathbb{R}$ mit $|t| < \varepsilon$. Wir definieren die Funktion $\phi : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$ durch

$$\phi(t) := f(x_0 + t\xi) \quad \text{für } -\varepsilon < t < \varepsilon.$$

Nach Satz 2.3.2 ist diese Funktion zweimal stetig differenzierbar mit

$$\begin{aligned} \phi'(t) &= \sum_{j=1}^n \partial_j f(x_0 + t\xi) \xi_j, \\ \phi''(t) &= \sum_{j=1}^n \sum_{i=1}^n \partial_i \partial_j f(x_0 + t\xi) \xi_i \xi_j = \xi^T d^2f(x_0 + t\xi) \xi. \end{aligned} \tag{2.7.2}$$

Da ϕ an der Stelle $t = 0$ ein lokales Minimum hat, gilt $\phi''(0) \geq 0$ nach einem Satz aus Analysis I. Hier ist nochmals das Argument. Sei $0 < \delta < \varepsilon$ so gewählt, dass $\phi(0) \leq \phi(t)$ ist für $-\delta \leq t \leq \delta$. Dann existiert nach dem Mittelwertsatz für jedes $k \in \mathbb{N}$ eine Zahl $0 < t_k \leq \delta/k$ mit

$$\phi'(t_k) = \frac{\phi(\delta/k) - \phi(0)}{\delta/k} \geq 0.$$

(Man kann etwa das grösste solche t_k in dem Intervall $0 < t \leq \delta/k$ wählen, sollte einem daran gelegen sein, das abzählbare Auswahlaxiom zu vermeiden.) Da $\phi'(0) = 0$ ist, folgt daraus unter Verwendung von (2.7.2) die Ungleichung

$$\xi^T d^2f(x_0)\xi = \phi''(0) = \lim_{t \searrow 0} \frac{\phi'(t)}{t} = \lim_{k \rightarrow \infty} \frac{\phi'(t_k)}{t_k} \geq 0.$$

Damit ist Teil (i) bewiesen. Man beachte, dass dieses Argument die Stetigkeit der zweiten partiellen Ableitungen von f nicht verwendet.

Wir beweisen Teil (ii) in zwei Schritten.

Schritt 1. Sei $A = A^T \in \mathbb{R}^{n \times n}$ eine symmetrische positiv definite Matrix, das heißt, für alle $\xi \in \mathbb{R}^n \setminus \{0\}$ gilt $\xi^T A \xi > 0$. Dann existiert eine Konstante $\varepsilon > 0$, so dass jeder Vektor $\xi \in \mathbb{R}^n$ die Ungleichung $\xi^T A \xi \geq \varepsilon \|\xi\|^2$ erfüllt.

Die Einheitssphäre $S^{n-1} := \{\xi \in \mathbb{R}^n \mid \|\xi\| = 1\}$ ist nach dem Satz von Heine–Borel eine kompakte Teilmenge des \mathbb{R}^n . Da die durch $g(\xi) := \xi^T A \xi$ definierte Funktion $g : S^{n-1} \rightarrow (0, \infty)$ stetig ist, existiert nach einem Satz aus Analysis I ein Element $\xi_0 \in \mathbb{R}^n$ mit $g(\xi_0) \leq g(\xi)$ für alle $\xi \in S^{n-1}$. Daraus folgt

$$\|\xi\|^{-2} \xi^T A \xi = (\|\xi\|^{-1} \xi)^T A (\|\xi\|^{-1} \xi) \geq \xi_0^T A \xi_0 =: \varepsilon > 0$$

für alle $\xi \in \mathbb{R}^n \setminus \{0\}$ und damit ist Schritt 1 bewiesen.

Schritt 2. Wir beweisen Teil (ii).

Da die Hessematrix von f an der Stelle x_0 positiv definit ist, existiert nach Schritt 1 eine Konstante $\varepsilon > 0$ mit

$$\xi^T d^2f(x_0)\xi \geq \varepsilon \|\xi\|^2 \quad \text{für alle } \xi \in \mathbb{R}^2. \quad (2.7.3)$$

Nach Korollar 2.6.4 existiert nun eine Zahl $\delta > 0$ mit $B_\delta(x_0) \subset U$, so dass jeder Vektor $\xi \in \mathbb{R}^n$ mit $0 < \|\xi\| < \delta$ die Ungleichung

$$\frac{|f(x_0 + \xi) - (T_{x_0}^2 f)(x_0 + \xi)|}{\|\xi\|^2} < \frac{\varepsilon}{4} \quad (2.7.4)$$

erfüllt. Mit $(T_{x_0}^2 f)(x_0 + \xi) = f(x_0) + df(x_0)\xi + \frac{1}{2}\xi^T d^2f(x_0)\xi$ und $df(x_0) = 0$ folgt daraus für jeden Vektor $\xi \in \mathbb{R}^n$ mit $0 < \|\xi\| < \delta$ die Ungleichung

$$\begin{aligned} f(x_0 + \xi) - f(x_0) &= \frac{1}{2}\xi^T d^2f(x_0)\xi + f(x_0 + \xi) - (T_{x_0}^2 f)(x_0 + \xi) \\ &\geq \frac{1}{2}\xi^T d^2f(x_0)\xi - |f(x_0 + \xi) - (T_{x_0}^2 f)(x_0 + \xi)| \\ &\geq \frac{\varepsilon}{4} \|\xi\|^2. \end{aligned}$$

Im letzten Schritt haben wir die Ungleichungen (2.7.3) und (2.7.4) verwendet. Damit ist Teil (ii) bewiesen.

Die Aussagen in (iii) und (iv) folgen direkt aus (i) und (ii) indem man f durch $-f$ ersetzt. Damit ist Satz 2.7.11 bewiesen. \square

Beispiel 2.7.12. Die Funktion $f(x, y) := 3x^4 + y^2$ auf dem \mathbb{R}^2 besitzt an der Stelle $(x_0, y_0) = (0, 0)$ ein striktes lokales Minimum, jedoch ist ihre Hessematrix an dieser Stelle nicht positiv definit.

Beispiel 2.7.13. Die Funktion

$$f(x, y) := 3x^4 + y^2 - 4x^2y = (y - x^2)(y - 3x^2)$$

besitzt an der Stelle $(x_0, y_0) = (0, 0)$ einen kritischen Punkt mit der folgenden Eigenschaft. Schränkt man die Funktion auf eine beliebige Gerade durch den Nullpunkt ein, so besitzt die so eingeschränkte Funktion im Nullpunkt ein striktes lokales Minimum. Genauer, definieren wir für einen beliebigen Vektor $(x, y) \in \mathbb{R}^2 \setminus \{0, 0\}$ die Funktion $\phi_{x,y} : \mathbb{R} \rightarrow \mathbb{R}$ durch

$$\phi_{x,y}(t) := f(tx, ty) = 3x^4t^4 - 4x^2yt^3 + t^2y^2$$

für $t \in \mathbb{R}$, so ist $t_0 = 0$ ein striktes lokales Minimum von $\phi_{x,y}$ (sowohl im Fall $y = 0$ als auch im Fall $y \neq 0$). Jedoch ist der Nullpunkt kein lokales Minimum von f , denn es gilt $f(\varepsilon, 2\varepsilon^2) < 0 < f(\varepsilon, 0)$ für alle $\varepsilon > 0$.

Übung 2.7.14. Sei $f : U \rightarrow \mathbb{R}$ eine zweimal stetig differenzierbare Funktion auf einer offenen Teilmenge $U \subset \mathbb{R}^n$ und sei $x_0 \in U$ ein kritischer Punkt von f . Existieren zwei Vektoren $\xi, \eta \in \mathbb{R}^n$ mit $\xi^T d^2f(x_0)\xi < 0 < \eta^T d^2f(x_0)\eta$, so ist x_0 kein lokales Extremum.

Übung 2.7.15. Sei $f : U \rightarrow \mathbb{R}$ stetig differenzierbar und sei $\nabla f : U \rightarrow \mathbb{R}^n$ an der Stelle $x_0 \in U$ differenzierbar. Dann gilt

$$\lim_{\xi \rightarrow 0} \frac{|f(x_0 + \xi) - f(x_0) - df(x_0)\xi - \frac{1}{2}\xi^T d^2f(x_0)\xi|}{\|\xi\|^2} = 0. \quad (2.7.5)$$

Daraus folgt, dass alle Aussagen von Satz 2.7.11 unter diesen schwächeren Voraussetzungen gültig bleiben. (Vergleiche (2.7.5) mit (2.6.10).)

Übung 2.7.16. Sei $A = A^T \in \mathbb{R}^{n \times n}$ eine positiv semidefinite Matrix.

(i) Für alle $\xi \in \mathbb{R}^n$ gilt $\langle \xi, A\xi \rangle = 0 \iff A\xi = 0$.

(ii) Die Matrix A ist genau dann positiv definit wenn $\det(A) \neq 0$ ist.

Hinweis: Ist $\langle \xi, A\xi \rangle = 0$ so besitzt die Funktion $\mathbb{R}^n \rightarrow \mathbb{R} : x \mapsto \langle x, Ax \rangle$ ein lokales Minimum an der Stelle $x = \xi$. Für Teil (ii) siehe Satz D.3.7.

Übung 2.7.17. Sei $U \subset \mathbb{R}^n$ eine offene Menge und sei $f : U \rightarrow \mathbb{R}$ eine C^2 -Funktion. Ein kritischer Punkt $x_0 \in U$ von f heißt **nichtdegeneriert**, wenn die Hessematrix $d^2f(x_0)$ eine von Null verschiedene Determinante besitzt. Sei $x_0 \in U$ ein lokales Minimum von f und ein nichtdegenerierter kritischer Punkt von f . Dann ist x_0 ein striktes lokales Minimum von f .

Konvexe Funktionen

Es sei zunächst daran erinnert, dass eine Teilmenge $U \subset \mathbb{R}^n$ **konvex** genannt wird, wenn für je zwei Elemente $x_0, x_1 \in U$ auch alle ihre *Konvexkombinationen* $(1-t)x_0 + tx_1$ mit $0 \leq t \leq 1$ in U enthalten sind. Wir verwenden im folgenden den Begriff der lokalen Lipschitz-Stetigkeit (Definition 4.1.1), der erst in Kapitel 4 ausführlicher diskutiert wird (siehe Lemma 4.1.2).

Definition 2.7.18 (Konvexe Funktionen). Sei $U \subset \mathbb{R}^n$ eine konvexe Teilmenge. Eine Funktion $f : U \rightarrow \mathbb{R}$ heißt **konvex**, wenn sie für alle $x_0, x_1 \in U$ und alle $0 \leq t \leq 1$ die Ungleichung

$$f((1-t)x_0 + tx_1) \leq (1-t)f(x_0) + tf(x_1)$$

erfüllt. Sie heißt **strikt konvex**, wenn sie für alle $x_0, x_1 \in U$ mit $x_0 \neq x_1$ und alle $0 < t < 1$ die strikte Ungleichung

$$f((1-t)x_0 + tx_1) < (1-t)f(x_0) + tf(x_1)$$

erfüllt.

Satz 2.7.19. Sei $U \subset \mathbb{R}^n$ eine konvexe offene Teilmenge und sei $f : U \rightarrow \mathbb{R}$ eine konvexe Funktion. Dann gilt folgendes.

(i) f ist lokal Lipschitz-stetig.

(ii) Ist f an der Stelle $x_0 \in U$ differenzierbar mit $df(x_0) = 0$ so gilt

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

(iii) Die Menge $\{x \in U \mid f(x) = \inf_U f\}$ ist konvex. (Sie kann auch leer sein.)

Beweis. Im Beweis von Teil (i) verwenden wir die Norm $\|x\|_\infty := \max_i |x_i|$ ebenso wie die Euklidische Norm $\|x\| := \sqrt{\sum_i x_i^2}$ für $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, und bezeichnen die Standardbasis des \mathbb{R}^n mit e_1, \dots, e_n . Sei nun $x_0 \in U$ und wähle $r > 0$ mit

$$Q := \{x \in \mathbb{R}^n \mid \|x - x_0\|_\infty \leq 3r\} \subset U.$$

Wir zeigen in drei Schritten, dass f auf dem abgeschlossenen Ball

$$B := \{x \in \mathbb{R}^n \mid \|x - x_0\| \leq r\} \subset Q \subset U$$

Lipschitz-stetig ist.

Schritt 1. Es gilt $M := \sup_{x \in Q} f(x) < \infty$.

Sei $\mathcal{E} \subset \mathbb{R}^n$ die endliche Indexmenge

$$\mathcal{E} := \{ \varepsilon = (\varepsilon_1, \dots, \varepsilon_n) \in \mathbb{R}^n \mid \varepsilon_i \in \{-1, +1\} \text{ für } i = 1, \dots, n \},$$

und für $\varepsilon \in \mathcal{E}$ sei

$$v_\varepsilon := x_0 + 3r \sum_{i=1}^n \varepsilon_i e_i \in Q.$$

Dies sind die *Extremalpunkte* des Würfels Q und das Maximum von f über diese Punkte ist die Zahl $M := \max\{f(v_\varepsilon) \mid \varepsilon \in \mathcal{E}\}$. Sei nun $x \in Q$. Dann existiert eine Abbildung $\lambda : \mathcal{E} \rightarrow [0, 1]$ mit $x = \sum_{\varepsilon \in \mathcal{E}} \lambda(\varepsilon) v_\varepsilon$ und $\sum_{\varepsilon \in \mathcal{E}} \lambda(\varepsilon) = 1$. Die Existenz einer solchen Abbildung kann mit vollständiger Induktion über die Dimension n gezeigt werden. Da f konvex ist, gilt

$$f(x) \leq \sum_{\varepsilon \in \mathcal{E}} \lambda(\varepsilon) f(v_\varepsilon) \leq \sum_{\varepsilon \in \mathcal{E}} \lambda(\varepsilon) M = M$$

und damit ist Schritt 1 bewiesen.

Schritt 2. Für alle $x \in Q$ gilt $f(x) \geq m := 2f(x_0) - M$.

Ist $x \in Q$, so ist auch $2x_0 - x \in Q$. Da $x_0 = \frac{1}{2}x + \frac{1}{2}(2x_0 - x)$ ist, folgt aus der Konvexität von f die Ungleichung $2f(x_0) \leq f(x) + f(2x_0 - x) \leq f(x) + M$ und daher $f(x) \geq 2f(x_0) - M = m$. Damit ist Schritt 2 bewiesen.

Schritt 3. Für alle $x, y \in B$ gilt $|f(x) - f(y)| \leq \frac{M-m}{2r} \|x - y\|$.

Seien $x, y \in B$ mit $x \neq y$. Dann gilt $0 < \|y - x\| \leq 2r$. Sei

$$t := \frac{\|y - x\|}{2r}, \quad \xi := \frac{y - x}{t}.$$

Dann ist $\|\xi\| = 2r$, daher $\|x + \xi - x_0\| \leq 3r$ und daraus folgt $x + \xi \in Q$.

Weiterhin gilt $y = x + t\xi = (1-t)x + t(x + \xi)$ und $0 \leq t \leq 1$ und daher

$$\begin{aligned} f(y) &\leq (1-t)f(x) + tf(x + \xi) \\ &= f(x) + \frac{f(x + \xi) - f(x)}{2r} \|y - x\| \\ &\leq f(x) + \frac{M - m}{2r} \|y - x\|. \end{aligned}$$

Die Ungleichung $f(x) \leq f(y) + \frac{M-m}{2r} \|y - x\|$ folgt nun durch Vertauschen der Rollen von x und y . Damit sind Schritt 3 und Teil (i) bewiesen.

Wir beweisen Teil (ii). Sei f an der Stelle x_0 differenzierbar. Dann gilt

$$f(x) \geq f(x_0) + df(x_0)(x - x_0) \quad (2.7.6)$$

für alle $x \in U$. Zum Beweis dieser Ungleichung halten wir ein Element $x \in U$ fest und definieren $\xi := x_0 - x$. Dann gilt $x_0 + t\xi \in U$ für jedes hinreichend kleine $t > 0$. Mit $x_0 - \xi = x$ und

$$x_0 = \frac{t}{1+t}(x_0 - \xi) + \frac{1}{1+t}(x_0 + t\xi).$$

ergibt sich daraus die Ungleichung

$$f(x_0) \leq \frac{t}{1+t}f(x) + \frac{1}{1+t}f(x_0 + t\xi).$$

Daraus folgt

$$f(x_0) + tf(x_0) \leq tf(x) + f(x_0 + t\xi)$$

und daher

$$f(x_0) - f(x) \leq \frac{f(x_0 + t\xi) - f(x_0)}{t}.$$

Mit dem Grenzübergang $t \rightarrow 0$ folgt die behauptete Ungleichung (2.7.6). Ist $df(x_0) = 0$ so gilt $f(x_0) \leq f(x)$ für alle $x \in U$ nach (2.7.6). Damit ist Teil (ii) bewiesen. Teil (iii) folgt direkt aus der Definition der Konvexität und damit ist Satz 2.7.19 bewiesen. \square

Beispiel 2.7.20. Die durch $f(x) := 0$ für $0 < x < 1$ und $f(0) := f(1) := 1$ definierte Funktion $f : [0, 1] \rightarrow \mathbb{R}$ ist konvex aber unstetig am Rand. Auf die Voraussetzung, dass U offen ist, kann also in Teil (i) von Satz 2.7.19 nicht verzichtet werden.

Teil (ii) von Satz 2.7.19 sagt dass jeder kritische Punkt einer stetig differenzierbaren konvexen Funktion $f : U \rightarrow \mathbb{R}$ auf einer konvexen offenen Teilmenge $U \subset \mathbb{R}^n$ ein absolutes Minimum von f ist. Umgekehrt sagt Satz 2.7.2 das jedes absolute Minimum von f ein kritischer Punkt ist. Daher folgt aus Teil (iii) von Satz 2.7.19, dass die Menge der kritischen Punkte jeder stetig differenzierbaren konvexen Funktion $f : U \rightarrow \mathbb{R}$ eine konvexe Teilmenge von U ist. Der folgende Satz charakterisiert zweimal stetig differenzierbare konvexe Funktionen über die Eigenschaften ihrer Hessematrizen.

Satz 2.7.21. Sei $U \subset \mathbb{R}^n$ eine konvexe offene Teilmenge und sei $f : U \rightarrow \mathbb{R}$ eine zweimal stetig differenzierbare Funktion. Dann gilt folgendes.

(i) f ist genau dann konvex, wenn die Hessematrix von f an jeder Stelle positiv semidefinit ist, das heißt wenn die Ungleichung $\xi^T d^2f(x)\xi \geq 0$ für alle $x \in U$ und alle $\xi \in \mathbb{R}^n$ erfüllt ist.

(ii) Wenn die Hessematrix von f an jeder Stelle positiv definit ist, das heißt wenn die Ungleichung $\xi^T d^2f(x)\xi > 0$ für alle $x \in U$ und alle $\xi \in \mathbb{R}^n \setminus \{0\}$ erfüllt ist, dann ist f strikt konvex.

Beweis. Ist f konvex und $x_0 \in U$, so ist auch die durch

$$g(x) := f(x) - f(x_0) - df(x_0)(x - x_0)$$

definierte Funktion $g : U \rightarrow \mathbb{R}$ konvex. Diese Funktion g hat einen kritischen Punkt an der Stelle x_0 . Nach Teil (ii) von Satz 2.7.19 ist x_0 daher ein lokales Minimum von g , und nach Teil (i) von Satz 2.7.11 ist daher die Hessematrix $d^2g(x_0) = d^2f(x_0)$ positiv semidefinit.

Umgekehrt nehmen wir an, dass die Hessematrix $d^2f(x)$ für jedes $x \in U$ positiv semidefinit ist. Seien $x_0, x_1 \in U$ und sei $\phi : [0, 1] \rightarrow \mathbb{R}$ die Funktion

$$\phi(t) := f((1-t)x_0 + tx_1), \quad 0 \leq t \leq 1.$$

Diese Funktion ist nach Satz 2.3.2 zweimal stetig differenzierbar und hat nach (2.7.2) die zweite Ableitung

$$\phi''(t) = (x_1 - x_0)^T d^2f((1-t)x_0 + tx_1)(x_1 - x_0) \geq 0.$$

Nach dem Mittelwertsatz ist daher die erste Ableitung $\phi' : [0, 1] \rightarrow \mathbb{R}$ monoton wachsend und daraus folgt für jedes $t \in [0, 1]$ die Ungleichung

$$\phi(t) - \phi(0) = t \int_0^1 \phi'(ts) ds \leq t \int_0^1 \phi'(s) ds = t(\phi(1) - \phi(0)). \quad (2.7.7)$$

Daher gilt $\phi(t) \leq (1-t)\phi(0) + t\phi(1)$ und damit ist Teil (i) bewiesen.

Ist die Hessematrix $d^2f(x)$ für jedes $x \in U$ positiv definit und ist $x_0 \neq x_1$, so gilt $\phi''(t) > 0$ in dieser Herleitung, damit ist ϕ' strikt monoton wachsend, und daher ist die Ungleichung in (2.7.7) für $0 < t < 1$ strikt. Damit ist auch Teil (ii) von Satz 2.7.21 bewiesen. \square

Beispiel 2.7.22. (i) Die Funktion $f(x) := 1/x$ für $x > 0$ ist strikt konvex, hat aber keine kritischen Punkte. Sie ist nicht global Lipschitz-stetig.

(ii) Die Funktion $f(x) := x^4$ ist strikt konvex aber ihre "Hessematrix" (in diesem Fall einfach die reelle Zahl $f''(x)$) verschwindet an der Stelle $x = 0$.

(iii) Die Funktion $f(x, y) := x^2$ ist konvex, aber nicht strikt konvex, und die Menge ihrer kritischen Punkte ist $C := \{(x, y) \in \mathbb{R}^2 \mid x = 0\}$.

(iv) Die Funktion $f(x, y) := |x|$ ist konvex, aber nicht strikt konvex, und ist global Lipschitz-stetig. Sie hat die gleiche Menge C der globalen Minima wie in (iii), ist aber an genau diesen Stellen nicht differenzierbar.

Übung 2.7.23. Sei $U \subset \mathbb{R}^n$ eine konvexe offene Menge und sei $f : U \rightarrow \mathbb{R}$ eine zweimal stetig differenzierbare konvexe Funktion. Besitzt f einen nicht-degenerierten kritischen Punkt $x_0 \in U$, so gilt

$$f(x_0) < f(x) \quad \text{für alle } x \in U \setminus \{x_0\}$$

und f besitzt keine anderen kritischen Punkte. **Hinweis:** Übung 2.7.17 und Satz 2.7.19.

Kapitel 3

Implizite Funktionen

Dieses Kapitel befasst sich mit den Lösungen einer Gleichung der Form

$$f(x) = y,$$

wobei $f : U \rightarrow Y$ eine stetig differenzierbare Abbildung von einer offenen Teilmenge $U \subset X$ eines endlichdimensionalen normierten Vektorraumes X in einen anderen endlichdimensionalen normierten Vektorraum Y ist. Der Abschnitt 3.1 behandelt den Fall, dass beide Vektorräume dieselbe Dimension haben und liefert ein Kriterium für die lokale Existenz und Eindeutigkeit der Lösung $x \in U$ sowie deren differenzierbare Abhängigkeit von y . Dies ist der *Satz von der Umkehrabbildung*, der wichtigste Satz dieses Kapitels. Der Abschnitt 3.2 behandelt den Fall $\dim(X) > \dim(Y)$ und liefert Bedingungen, unter denen die Lösungsmenge sich lokal als Graph einer differenzierbaren Abbildung darstellen läßt. Dies ist der *Satz von der impliziten Funktion*, der sich leicht auf den Satz von der Umkehrabbildung zurückführen läßt. Dieser Satz führt zu dem Konzept einer *Untermannigfaltigkeit des \mathbb{R}^n* , das im Abschnitt 3.3 behandelt wird. Der Abschnitt 3.4 beschäftigt sich mit der Frage, wie man die Extrema einer differenzierbaren Funktion auf einer solchen Untermannigfaltigkeit findet, und das führt zu dem Begriff von *Lagrange-Multiplikatoren*.

3.1 Der Satz von der Umkehrabbildung

Es sei zunächst vermerkt, dass wir den Begriff einer C^ℓ -Abbildung über die partiellen Ableitungen eingeführt haben, und dies erfordert natürlich, dass

unsere Abbildung auf einer offenen Teilmenge $U \subset \mathbb{R}^n$ definiert ist. Es gibt jedoch diverse Beispiele von glatten Abbildungen auf offenen Teilmengen von endlichdimensionalen normierten Vektorräumen, die nicht mit einer kanonischen Basis ausgestattet sind, und es ist in diesem Fall nützlich, von C^ℓ -Abbildungen Abbildungen zu sprechen, ohne zuvor eine Basis zu wählen. Dies läßt sich induktiv wie folgt formulieren. Seien X, Y endlichdimensionale normierte Vektorräume, sei $U \subset X$ eine offene Teilmenge, und sei $\ell \in \mathbb{N}$. Wir nennen eine Abbildung $f : U \rightarrow Y$ **ℓ -mal stetig differenzierbar**, wenn sie an jeder Stelle $x \in U$ differenzierbar und ihre Ableitung $df : U \rightarrow \mathcal{L}(X, Y)$ $(\ell-1)$ -mal stetig differenzierbar ist. (Siehe Definition 2.2.2 für den Fall $\ell = 1$.) Wir nennen sie **glatt** oder C^∞ , wenn sie für jedes $\ell \in \mathbb{N}$ ℓ -mal stetig differenzierbar ist. Diese Definitionen funktionieren ebensogut für unendlichdimensionale Banachräume.

Definition 3.1.1 (Diffeomorphismus). *Seien X und Y endlichdimensionale normierte Vektorräume und seien $U \subset X$ und $V \subset Y$ offene Teilmengen. Sei $\ell \in \mathbb{N} \cup \{\infty\}$. Eine Abbildung $f : U \rightarrow V$ heißt **C^ℓ -Diffeomorphismus**, wenn f bijektiv ist und sowohl $f : U \rightarrow V$ als auch die Umkehrabbildung $f^{-1} : V \rightarrow U$ eine C^ℓ -Abbildung ist. Ein C^∞ -Diffeomorphismus wird auch **glatter Diffeomorphismus** genannt. Die Mengen U und V heißen **diffeomorph** wenn ein glatter Diffeomorphismus $f : U \rightarrow V$ existiert.*

Lemma 3.1.2. *Seien $U \subset X$ und $V \subset Y$ offene Teilmengen von endlichdimensionalen normierten Vektorräumen, und sei $f : U \rightarrow V$ ein C^1 -Diffeomorphismus. Dann ist $\dim(X) = \dim(Y)$, für jedes $x \in U$ ist $df(x) : X \rightarrow Y$ eine invertierbare lineare Abbildung, und für jedes $y \in V$ gilt*

$$df^{-1}(y) = df(f^{-1}(y))^{-1}. \quad (3.1.1)$$

Beweis. Sei $g := f^{-1} : V \rightarrow U$ die Umkehrabbildung von f , sei $x \in U$, und seien $y \in V$, $A \in \mathcal{L}(X, Y)$, und $B \in \mathcal{L}(Y, X)$ durch

$$y := f(x), \quad A := df(x), \quad B := dg(y)$$

gegeben. Da $g \circ f = \text{id}_U$ und $f \circ g = \text{id}_V$ ist, erhalten wir aus der Kettenregel in Satz 2.3.2 die Gleichungen

$$\begin{aligned} B \circ A &= dg(f(x)) \circ df(x) = d(g \circ f)(x) = \mathbb{1}_X, \\ A \circ B &= df(g(y)) \circ dg(y) = d(f \circ g)(y) = \mathbb{1}_Y. \end{aligned}$$

Daraus folgt, dass A invertierbar und $A^{-1} = B$ ist. Insbesondere haben X und Y die gleiche Dimension und damit ist Lemma 3.1.2 bewiesen. \square

Beispiel 3.1.3. Die Exponentialfunktion $\exp : \mathbb{R} \rightarrow (0, \infty)$ ist ein glatter Diffeomorphismus, deren Umkehrabbildung per Definition die Logarithmusfunktion $\log := \exp^{-1} : (0, \infty) \rightarrow \mathbb{R}$ ist. Dass beide Funktionen glatt sind, wurde in Analysis I bewiesen. Ihre Potenzreihendarstellung zeigt, dass sie sogar analytisch sind, wie sich auch aus Korollar 2.6.13 herleiten lässt.

Beispiel 3.1.4. Die Funktion $\mathbb{R} \rightarrow \mathbb{R} : x \mapsto x^3$ ist glatt und bijektiv, jedoch verschwindet ihre Ableitung an der Stelle $x = 0$, und ihre Umkehrabbildung $\mathbb{R} \rightarrow \mathbb{R} : y \mapsto \text{sign}(y)|y|^{1/3}$ ist an der Stelle $y = 0$ nicht differenzierbar.

Beispiel 3.1.5. Sei $I \subset \mathbb{R}$ ein offenes Intervall und $f : I \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion mit $f(x) > 0$ für alle $x \in I$. Dann ist $J := f(I)$ ein Intervall nach dem Zwischenwertsatz, und ist offen, weil f keine lokalen Extrema haben kann. Darüber hinaus ist $f : I \rightarrow J$ strikt monoton wachsend nach dem Mittelwertsatz, und ist damit bijektiv. Die Umkehrabbildung $f^{-1} : J \rightarrow I$ ist differenzierbar nach der eindimensionalen Version des Satzes von der Umkehrabbildung in Analysis I.

Beispiel 3.1.6. Sei $\|\cdot\|$ die Euklidische Norm auf dem \mathbb{R}^n . Der offene Einheitsball $U := \{x \in \mathbb{R}^n \mid \|x\| < 1\}$ ist diffeomorph zum \mathbb{R}^n . Ein glatter Diffeomorphismus $f : U \rightarrow \mathbb{R}^n$ und seine Umkehrabbildung sind durch

$$f(x) = \frac{x}{\sqrt{1 - \|x\|^2}}, \quad f^{-1}(y) = \frac{y}{\sqrt{1 + \|y\|^2}}$$

für $x \in U$ und $y \in \mathbb{R}^n$ gegeben.

Beispiel 3.1.7. Die obere Halbebene $\mathbb{H} := \{z \in \mathbb{C} \mid \text{Im}(z) > 0\}$ ist diffeomorph zur offenen Einheitskreisscheibe $\mathbb{D} := \{\zeta \in \mathbb{C} \mid |\zeta| < 1\}$. Ein glatter Diffeomorphismus $f : \mathbb{H} \rightarrow \mathbb{D}$ und seine Umkehrabbildung sind durch

$$f(z) = \frac{z - \mathbf{i}}{z + \mathbf{i}}, \quad f^{-1}(\zeta) = \mathbf{i} \frac{1 + \zeta}{1 - \zeta}$$

für $z \in \mathbb{H}$ und $\zeta \in \mathbb{D}$ gegeben.

Beispiel 3.1.8. Die Abbildung

$$\text{Inv} : \text{GL}(n, \mathbb{R}) \rightarrow \text{GL}(n, \mathbb{R}), \quad \text{Inv}(A) := A^{-1},$$

ist nach Satz 1.4.1 stetig und nach Beispiel 2.3.14 stetig differenzierbar. Da die Abbildung Inv bijektiv ist, und mit ihrer Umkehrabbildung übereinstimmt, ist Inv also ein C^1 -Diffeomorphismus. Dass Inv sogar ein glatter Diffeomorphismus ist, wird im folgenden Lemma bewiesen.

Lemma 3.1.9. *Die Abbildung $\text{Inv} : \text{GL}(n, \mathbb{R}) \rightarrow \text{GL}(n, \mathbb{R})$ ist glatt.*

Beweis. Nach Beispiel 2.3.14 ist die Abbildung Inv stetig differenzierbar und ihre Ableitung $d\text{Inv}(A) \in \mathcal{L}(\mathbb{R}^{n \times n}, \mathbb{R}^{n \times n})$ an der Stelle $A \in \text{GL}(n, \mathbb{R})$ ist durch

$$d\text{Inv}(A)\widehat{A} = -A^{-1}\widehat{A}A^{-1} = -\text{Inv}(A)\widehat{A}\text{Inv}(A) \quad (3.1.2)$$

für $\widehat{A} \in \mathbb{R}^{n \times n}$ gegeben. Es ist an dieser Stelle nützlich, die folgende Bezeichnung einzuführen. Je zwei Matrizen $B_0, B_1 \in \mathbb{R}^{n \times n}$ bestimmen eine lineare Abbildung $\Lambda(B_0, B_1) : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$, die durch die Formel

$$(\Lambda(B_0, B_1))(\widehat{A}) := -B_0\widehat{A}B_1$$

für $\widehat{A} \in \mathbb{R}^{n \times n}$ gegeben ist. Wir erhalten damit eine Abbildung

$$\Lambda : \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n} \rightarrow \mathcal{L}(\mathbb{R}^{n \times n}, \mathbb{R}^{n \times n}).$$

Mit Hilfe von Lemma 2.1.15 und der Beispiele 2.1.6, 2.1.7, und 2.1.8 läßt sich leicht verifizieren, dass diese Abbildung glatt ist. Genauer, die Abbildung Λ von dem $2n^2$ -dimensionalen Vektorraum $X := \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n}$ mit Werten in dem n^4 -dimensionalen Vektorraum $Y := \mathcal{L}(\mathbb{R}^{n \times n}, \mathbb{R}^{n \times n})$ ist quadratisch, daher nach Beispiel 2.1.8 koordinatenweise differenzierbar, und ist daher nach Lemma 2.1.15 differenzierbar. Außerdem zeigt die Formel in Beispiel 2.1.8, dass die Ableitung von Λ eine lineare Abbildung $d\Lambda : X \rightarrow \mathcal{L}(X, Y)$ ist. Diese ist nach Beispiel 2.1.7 wieder differenzierbar und hat eine konstante Ableitung (mit Werten in $\mathcal{L}(X, \mathcal{L}(X, Y))$), und jede konstante Abbildung ist glatt nach Beispiel 2.1.6. Nun läßt sich die Gleichung (3.1.2) in der Form

$$d\text{Inv}(A) = \Lambda(\text{Inv}(A), \text{Inv}(A)) \quad (3.1.3)$$

schreiben. Das heißt, dass die Ableitung von $\text{Inv} : \text{GL}(n, \mathbb{R}) \rightarrow \text{GL}(n, \mathbb{R})$ sich als die Komposition

$$d\text{Inv} = \Lambda \circ \iota \circ \text{Inv} : \text{GL}(n, \mathbb{R}) \rightarrow \mathcal{L}(\mathbb{R}^{n \times n}, \mathbb{R}^{n \times n}) \quad (3.1.4)$$

schreiben läßt, wobei die Abbildung

$$\iota : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n}$$

durch $\iota(B) := (B, B)$ für $B \in \mathbb{R}^{n \times n}$ definiert ist. Da ι eine lineare Abbildung ist, ist sie ebenfalls glatt. Ist nun Inv eine C^ℓ -Abbildung, so folgt aus der Gleichung (3.1.4) und Satz 2.5.9, dass $d\text{Inv}$ ebenfalls eine C^ℓ -Abbildung und Inv damit eine $C^{\ell+1}$ -Abbildung ist. Da wir bereits wissen, dass Inv eine C^1 -Abbildung ist, haben wir also durch vollständige Induktion gezeigt, dass Inv glatt ist. Damit ist Lemma 3.1.9 bewiesen. \square

Lemma 3.1.10. *Seien $U \subset X$ und $V \subset Y$ offene Teilmengen von endlichdimensionalen normierten Vektorräumen, und sei $f : U \rightarrow V$ ein C^1 -Diffeomorphismus. Ist f eine C^ℓ -Abbildung für ein $\ell \in \mathbb{N} \cup \{\infty\}$, so ist auch die Umkehrabbildung $f^{-1} : V \rightarrow U$ eine C^ℓ -Abbildung und f damit ein C^ℓ -Diffeomorphismus.*

Beweis. Nach Lemma 3.1.2 haben X und Y die gleiche Dimension und wir dürfen daher annehmen, dass $X = Y = \mathbb{R}^n$ ist. Dann zeigt die Formel

$$df^{-1}(y) = df(f^{-1}(y))^{-1} = \text{Inv}(df(f^{-1}(y))) \quad \text{für } y \in V$$

dass sich die Ableitung der Umkehrabbildung $f^{-1} : V \rightarrow U$ als Komposition

$$V \xrightarrow{f^{-1}} U \xrightarrow{df} \text{GL}(n, \mathbb{R}) \xrightarrow{\text{Inv}} \text{GL}(n, \mathbb{R})$$

schreiben läßt. Nun argumentieren wir mit vollständiger Induktion über ℓ . Für $\ell = 1$ ist nichts zu beweisen. Nehmen wir also an, dass $\ell \geq 2$ ist, dass die Aussage des Lemmas für $\ell - 1$ gilt, und dass $f : U \rightarrow V$ eine bijektive C^ℓ -Abbildung mit einer stetig differenzierbaren Umkehrabbildung ist. Dann ist $f^{-1} : V \rightarrow U$ nach Induktionsannahme $C^{\ell-1}$ und $df : U \rightarrow \text{GL}(n, \mathbb{R})$ ist nach Voraussetzung $C^{\ell-1}$. Da $\text{Inv} : \text{GL}(n, \mathbb{R}) \rightarrow \text{GL}(n, \mathbb{R})$ nach Lemma 3.1.9 eine glatte Abbildung ist, folgt aus Teil (v) von Satz 2.5.9 dass die Komposition $df^{-1} = \text{Inv} \circ df \circ f^{-1} : V \rightarrow \text{GL}(n, \mathbb{R})$ eine $C^{\ell-1}$ -Abbildung und f^{-1} daher eine C^ℓ -Abbildung ist. Damit ist Lemma 3.1.10 bewiesen. \square

Mit diesen Vorbereitungen sind wir bereit für den Hauptsatz dieses Kapitels. Wir formulieren den Satz zwar für Abbildungen zwischen endlichdimensionalen normierten Vektorräumen, jedoch läßt sich der Satz ebenso wie sein Beweis Wort für Wort auf Abbildungen zwischen unendlichdimensionalen Banachräumen übertragen.

Satz 3.1.11 (Satz von der Umkehrabbildung). *Seien X und Y endlichdimensionale normierte Vektorräume, sei $\Omega \subset X$ eine offene Teilmenge, sei $f : \Omega \rightarrow Y$ eine C^ℓ -Abbildung mit $\ell \in \mathbb{N} \cup \{\infty\}$, und sei $x_0 \in \Omega$. Wir nehmen an, dass die Ableitung $df(x_0) \in \mathcal{L}(X, Y)$ eine invertierbare lineare Abbildung ist. Dann existiert eine offene Teilmenge $U \subset \Omega$ mit $x_0 \in U$, so dass $V := f(U)$ eine offene Teilmenge von Y und*

$$f|_U : U \rightarrow V$$

ein C^ℓ -Diffeomorphismus ist.

Beweis. Siehe Seite 107. \square

Der Beweis von Satz 3.1.11 beruht auf dem folgenden Lemma. In diesem Lemma verwenden wir die Operatornorm

$$\|A\| := \sup_{0 \neq \xi \in X} \frac{\|A\xi\|}{\|\xi\|}$$

für die Endomorphismen (lineare Selbstabbildungen) $A : X \rightarrow X$ eines endlichdimensionalen normierten Vektorraumes $(X, \|\cdot\|)$. Außerdem erinnern wir an die Bezeichnung $B_r(x_0) := \{x \in X \mid \|x - x_0\| < r\}$ für den offenen Ball in X mit Mittelpunkt x_0 und Radius $r > 0$.

Lemma 3.1.12. *Sei $(X, \|\cdot\|)$ ein endlichdimensionaler normierter Vektorraum und seien $x_0 \in X$ und zwei Konstanten $r > 0$ und $0 < \gamma < 1$ gegeben. Weiter sei $\psi : B_r(x_0) \rightarrow X$ eine stetig differenzierbare Abbildung mit*

$$\|\mathbb{1} - d\psi(x)\| \leq \gamma \quad \text{für alle } x \in B_r(x_0). \quad (3.1.5)$$

Dann ist die Abbildung $\psi : B_r(x_0) \rightarrow X$ injektiv, ihre Bildmenge $\psi(B_r(x_0))$ ist offen, es gilt

$$B_{r(1-\gamma)}(\psi(x_0)) \subset \psi(B_r(x_0)) \subset B_{r(1+\gamma)}(\psi(x_0)), \quad (3.1.6)$$

und die Umkehrabbildung $\psi^{-1} : \psi(B_r(x_0)) \rightarrow B_r(x_0)$ ist stetig differenzierbar.

Beweis. Nach Voraussetzung und Satz 2.3.1 ist die Abbildung

$$\phi : B_r(x_0) \rightarrow X, \quad \phi(x) := x - \psi(x),$$

stetig differenzierbar mit der Ableitung $d\phi(x) = \mathbb{1} - d\psi(x)$ für $x \in B_r(x_0)$. Nach (3.1.5) erfüllt die Ableitung von ϕ daher die Ungleichung $\|d\phi(x)\| \leq \gamma$ für alle $x \in B_r(x_0)$, und nach Satz 2.4.1 folgt daraus

$$\|\phi(x) - \phi(x')\| \leq \gamma \|x - x'\| \quad (3.1.7)$$

für alle $x, x' \in B_r(x_0)$. Daraus wiederum folgen die Ungleichungen

$$\begin{aligned} \|\psi(x) - \psi(x')\| &= \|x - x' - \phi(x) + \phi(x')\| \\ &\leq \|x - x'\| + \|\phi(x) - \phi(x')\| \\ &\leq (1 + \gamma) \|x - x'\|, \end{aligned} \quad (3.1.8)$$

$$\begin{aligned} \|\psi(x) - \psi(x')\| &= \|x - x' - \phi(x) + \phi(x')\| \\ &\geq \|x - x'\| - \|\phi(x) - \phi(x')\| \\ &\geq (1 - \gamma) \|x - x'\| \end{aligned} \quad (3.1.9)$$

für alle $x, x' \in B_r(x_0)$. Nach (3.1.9) ist die Abbildung $\psi : B_r(x_0) \rightarrow X$ injektiv. Die weiteren Behauptungen beweisen wir in drei Schritten.

Schritt 1. *Wir beweisen (3.1.6).*

Die Inklusion $\psi(B_r(x_0)) \subset B_{r(1+\gamma)}(\psi(x_0))$ folgt direkt aus (3.1.8). Zu zeigen ist die Inklusion $B_{r(1-\gamma)}(\psi(x_0)) \subset \psi(B_r(x_0))$. Sei $y \in B_{r(1-\gamma)}(\psi(x_0))$ gegeben. Dann ist $\|y - \psi(x_0)\| < r(1 - \gamma)$ und daher

$$0 \leq s := \frac{\|y - \psi(x_0)\|}{1 - \gamma} < r. \quad (3.1.10)$$

Damit ist die Menge

$$K := \{x \in X \mid \|x - x_0\| \leq s\} \subset B_r(x_0)$$

nichtleer und sie ist ein vollständiger metrischer Raum mit der Abstandsfunktion $d(x, x') := \|x - x'\|$ für $x, x' \in K$. Nun sei $f_y : K \rightarrow X$ die durch

$$f_y(x) := y + \phi(x) \quad \text{für } x \in K$$

definierte Abbildung. Dann erfüllt jedes $x \in K$ die Ungleichung

$$\begin{aligned} \|f_y(x) - x_0\| &= \|y + \phi(x) - x_0\| \\ &= \|\phi(x) - \phi(x_0) + y - \psi(x_0)\| \\ &\leq \|\phi(x) - \phi(x_0)\| + \|y - \psi(x_0)\| \\ &\leq \gamma \|x - x_0\| + (1 - \gamma)s \\ &\leq \gamma s + (1 - \gamma)s \\ &= s. \end{aligned}$$

Hier folgt der zweite Schritt aus der Gleichung $x_0 = \phi(x_0) + \psi(x_0)$, der dritte aus der Dreiecksungleichung, der vierte aus (3.1.7) und (3.1.10), und der fünfte aus der Definition der Menge K . Damit haben wir gezeigt, dass $f_y(K) \subset K$ ist. Außerdem gilt

$$\|f_y(x) - f_y(x')\| = \|\phi(x) - \phi(x')\| \leq \gamma \|x - x'\|$$

für alle $x, x' \in K$ nach (3.1.7). Also ist $f_y : K \rightarrow K$ eine Kontraktion auf einem nichtleeren vollständigen metrischen Raum, und besitzt daher nach dem Banachschen Fixpunktsatz (Satz E.0.1) einen eindeutigen Fixpunkt

$$x = f_y(x) = y + \phi(x) \in K \subset B_r(x_0).$$

Dieser Fixpunkt erfüllt die Gleichung $\psi(x) = x - \phi(x) = y$ und damit ist Schritt 1 bewiesen.

Schritt 2. Die Bildmenge $\psi(B_r(x_0)) = \{\psi(x) \mid x \in B_r(x_0)\}$ ist offen.

Sei $y_1 \in \psi(B_r(x_0))$ gegeben und wähle $x_1 \in B_r(x_0)$ mit $\psi(x_1) = y_1$. Dann ist $\varepsilon := r - \|x_1 - x_0\| > 0$ und es gilt $B_\varepsilon(x_1) \subset B_r(x_0)$ nach der Dreiecksungleichung. Nach Schritt 1 folgt daraus $B_{\varepsilon(1-\gamma)}(y_1) \subset \psi(B_\varepsilon(x_1)) \subset \psi(B_r(x_0))$ und damit ist Schritt 2 bewiesen.

Schritt 3. Die Umkehrabbildung ψ^{-1} ist stetig differenzierbar.

Die Abbildung $\psi : B_r(x_0) \rightarrow \psi(B_r(x_0))$ ist bijektiv nach (3.1.9), und dieselbe Ungleichung zeigt, dass ψ^{-1} Lipschitz-stetig ist. Wir zeigen, dass ψ^{-1} differenzierbar ist. Nach Lemma 3.1.2 ist ihre Ableitung dann die Komposition

$$\psi(B_r(x_0)) \xrightarrow{\psi^{-1}} B_r(x_0) \xrightarrow{d\psi} \text{GL}(X) \xrightarrow{\text{Inv}} \text{GL}(X)$$

und ist daher stetig. (Hier bezeichnet $\text{GL}(X) \subset \mathcal{L}(X, X)$ die Gruppe der bijektiven linearen Abbildungen $A : X \rightarrow X$.)

Sei $y_1 \in \psi(B_r(x_0))$, sei $x_1 \in B_r(x_0)$ mit $\psi(x_1) = y_1$, und sei $\Psi := d\psi(x_1)$. Nach (3.1.5) gilt $\|\mathbb{1} - \Psi\| \leq \gamma < 1$, und daher ist Ψ nach Satz 1.4.1 invertierbar mit $\Psi^{-1} = \sum_{k=0}^{\infty} (\mathbb{1} - \Psi)^k$. Daraus folgt

$$\|\Psi^{-1}\| \leq \sum_{k=0}^{\infty} \|(\mathbb{1} - \Psi)^k\| \leq \sum_{k=0}^{\infty} \|\mathbb{1} - \Psi\|^k \leq \sum_{k=0}^{\infty} \gamma^k = \frac{1}{1 - \gamma}. \quad (3.1.11)$$

Sei nun $\varepsilon > 0$ gegeben. Nach Definition der Ableitung existiert ein $\rho > 0$ mit $\rho \leq r - \|x_0 - x_1\|$, so dass jedes $x \in B_\rho(x_1) \subset B_r(x_0)$ die Ungleichung

$$\|\psi(x) - \psi(x_1) - \Psi(x - x_1)\| \leq \varepsilon(1 - \gamma)^2 \|x - x_1\| \quad (3.1.12)$$

erfüllt. Mit $\delta := \rho(1 - \gamma)$ gilt $B_\delta(y_1) \subset \psi(B_\rho(x_1))$ nach Schritt 1.

Sei nun $y \in X$ gegeben mit $\|y - y_1\| < \delta$. Dann existiert ein $x \in B_\rho(x_1)$ mit $\psi(x) = y$. Unter Verwendung von (3.1.12), (3.1.11), und (3.1.9) (in genau dieser Reihenfolge) erhalten wir damit die Ungleichungskette

$$\begin{aligned} \|\psi^{-1}(y) - \psi^{-1}(y_1) - \Psi^{-1}(y - y_1)\| &= \|\Psi^{-1}(y - y_1 - \Psi(x - x_1))\| \\ &\leq \|\Psi^{-1}\| \|\psi(x) - \psi(x_1) - \Psi(x - x_1)\| \\ &\leq \|\Psi^{-1}\| \varepsilon(1 - \gamma)^2 \|x - x_1\| \\ &\leq \varepsilon(1 - \gamma) \|x - x_1\| \\ &\leq \varepsilon \|y - y_1\|. \end{aligned}$$

Daraus folgt, dass ψ^{-1} an der Stelle y_1 differenzierbar und $d\psi^{-1}(y_1) = \Psi^{-1}$ ist. Damit sind Schritt 3 und Lemma 3.1.12 bewiesen. \square

Beweis von Satz 3.1.11. Nach Voraussetzung ist die lineare Abbildung

$$A := df(x_0) : X \rightarrow Y$$

bijektiv und wir definieren die Abbildung $\psi : \Omega \rightarrow X$ durch

$$\psi(x) := A^{-1}f(x) \quad \text{für } x \in \Omega.$$

Dann ist ψ stetig differenzierbar nach Satz 2.3.2 und es gilt

$$d\psi(x) - \mathbb{1} = A^{-1}df(x) - \mathbb{1} = A^{-1}(df(x) - df(x_0))$$

für alle $x \in \Omega$. Nun sei

$$\varepsilon := \frac{1}{2\|A^{-1}\|} > 0, \quad \|A^{-1}\| = \sup_{0 \neq \eta \in Y} \frac{\|A^{-1}\eta\|_X}{\|\eta\|_Y}.$$

Da die Abbildung $df : \Omega \rightarrow \mathcal{L}(X, Y)$ nach Voraussetzung stetig ist existiert eine Konstante $\delta > 0$ mit $B_\delta(x_0) \subset \Omega$, so dass jedes $x \in B_\delta(x_0)$ die Ungleichung $\|df(x) - df(x_0)\| < \varepsilon$ erfüllt. Daraus folgt die Ungleichung

$$\begin{aligned} \|d\psi(x) - \mathbb{1}\| &= \|A^{-1}(df(x) - df(x_0))\| \\ &\leq \|A^{-1}\| \|df(x) - df(x_0)\| \\ &< \|A^{-1}\| \varepsilon \\ &= \frac{1}{2} \end{aligned}$$

für alle $x \in B_\delta(x_0)$. Nach Lemma 3.1.12 mit $\gamma = 1/2$ ist daher $\psi(B_\delta(x_0)) \subset X$ eine offene Teilmenge und $\psi : B_\delta(x_0) \rightarrow \psi(B_\delta(x_0))$ ein C^1 -Diffeomorphismus. Wir definieren nun

$$U := B_r(x_0) \subset \Omega, \quad V := A(\psi(B_\delta(x_0))) \subset Y.$$

Dann ist $V = \{A\psi(x) \mid x \in U\} = f(U)$ eine offene Teilmenge von Y , da die lineare Abbildung $A^{-1} : Y \rightarrow X$ stetig und V das Urbild der offenen Teilmenge $\psi(B_\delta(x_0)) \subset X$ unter A^{-1} ist. Nun ist die Abbildung

$$f|_U = A\psi|_U : U \rightarrow V$$

bijektiv mit der Umkehrabbildung

$$(f|_U)^{-1}(y) = (\psi|_U)^{-1}(A^{-1}y) \quad \text{für } y \in V.$$

Also ist $(f|_U)^{-1} : V \rightarrow U$ stetig differenzierbar nach Lemma 3.1.12 und Satz 2.3.2. Da f nach Voraussetzung eine C^ℓ -Abbildung ist, folgt nun aus Lemma 3.1.10, dass $(f|_U)^{-1} : V \rightarrow U$ ebenfalls eine C^ℓ -Abbildung ist. Damit ist Satz 3.1.11 bewiesen. \square

Korollar 3.1.13 (Offenheitssatz). *Seien X, Y endlichdimensionale normierte Vektorräume, sei $U \subset X$ eine offene Menge, und sei $f : U \rightarrow Y$ eine stetig differenzierbare Abbildung, so dass die Ableitung $df(x) \in \mathcal{L}(X, Y)$ für jedes $x \in U$ eine invertierbare lineare Abbildung ist. Dann ist $V := f(U)$ eine offene Teilmenge von Y .*

Beweis. Sei $y_0 \in V$ und wähle $x_0 \in U$ mit $f(x_0) = y_0$. Dann existiert nach Satz 3.1.11 eine offene Teilmenge $U_0 \subset U$ mit $x_0 \in U_0$, so dass $V_0 := f(U_0)$ eine offene Teilmenge von Y ist. Da $y_0 = f(x_0) \in f(U_0) = V_0$ ist, existiert also ein $\delta > 0$ mit $B_\delta(y_0) \subset V_0 \subset V$, und damit ist Korollar 3.1.13 bewiesen. \square

Korollar 3.1.14 (Diffeomorphiesatz). *Seien X, Y endlichdimensionale normierte Vektorräume, sei $U \subset X$ eine offene Menge, und sei $f : U \rightarrow Y$ ein injektive C^ℓ -Abbildung mit $\ell \in \mathbb{N} \cup \{\infty\}$, so dass $df(x) \in \mathcal{L}(X, Y)$ für jedes $x \in U$ eine invertierbare lineare Abbildung ist. Dann ist $V := f(U)$ eine offene Teilmenge von Y und $f : U \rightarrow V$ ein C^ℓ -Diffeomorphismus.*

Beweis. Die Teilmenge $V \subset Y$ ist offen nach Korollar 3.1.13 und die Abbildung $f : U \rightarrow V$ ist bijektiv nach Voraussetzung. Sei nun $y_0 \in V$ gegeben. Nach Satz 3.1.11 existiert dann eine offene Menge $V_0 \subset V$ mit $y_0 \in V_0$, so dass $f^{-1}|_{V_0} : V_0 \rightarrow X$ eine C^ℓ -Abbildung ist. Da dies für jedes $y_0 \in V$ gilt, ist f^{-1} eine C^ℓ -Abbildung und damit ist Korollar 3.1.14 bewiesen. \square

Beispiel 3.1.15. Die Sinusfunktion ist glatt mit $\sin'(0) = 1$. Nach Satz 3.1.11 ist daher ihre Einschränkung auf ein geeignetes offenes Intervall $U_0 \subset \mathbb{R}$ mit $0 \in U_0$ ein glatter Diffeomorphismus auf die entsprechende Bildmenge. Das größtmögliche solche Intervall ist $U_0 = (-\pi/2, \pi/2)$. Da die Sinusfunktion auf diesem Intervall injektiv ist und eine positive Ableitung hat, ist $\sin : (-\pi/2, \pi/2) \rightarrow (-1, 1)$ ein Diffeomorphismus nach Korollar 3.1.14.

Beispiel 3.1.16. Wir betrachten den Raum $\mathcal{A} := \mathbb{R}^{n \times n}$ mit einer beliebigen Matrixnorm $\|\cdot\|$ (Beispiel 1.3.6). Die Exponentialabbildung $\exp : \mathcal{A} \rightarrow \mathcal{A}$ ist nach Beispiel 2.3.13 stetig differenzierbar und hat die Ableitung $d \exp(0) = \text{id}$ an der Stelle $A_0 = 0$. In Korollar 4.4.4 werden wir sehen, dass die Exponentialabbildung sogar glatt ist. Daher folgt aus Satz 3.1.11, dass offene Mengen $U, V \subset \mathcal{A}$ existieren mit $0 \in U$, so dass $\exp : U \rightarrow V$ ein glatter Diffeomorphismus ist. Die Logarithmusreihe in Beispiel 1.6.2 zeigt, dass U, V so gewählt werden können, dass $V = \{B \in \mathcal{A} \mid \|B - \mathbb{1}\| < 1\}$ ist. Die Umkehrabbildung ist dann $\exp^{-1}(B) = \log(B) = -\sum_{k=1}^{\infty} (\mathbb{1} - B)^k / k$ für $B \in V$.

Übung 3.1.17. Nichtdegenerierte kritische Punkte sind isoliert. (Für die Definition siehe Übung 2.7.17.)

3.2 Der Satz von der impliziten Funktion

In diesem Abschnitt geht es um die Frage, unter welchen Voraussetzungen man eine Gleichung der Form

$$f(x, y) = 0 \quad (3.2.1)$$

lokal nach y auflösen kann, das heißt, dass eine stetig differenzierbare Funktion $y = g(x)$ existiert, so dass die Lösungsmenge lokal der Graph der Funktion g ist. Wir sagen dann, dass die Funktion g durch die Gleichung (3.2.1) implizit definiert ist. Hier ist zunächst ein einfaches Beispiel.

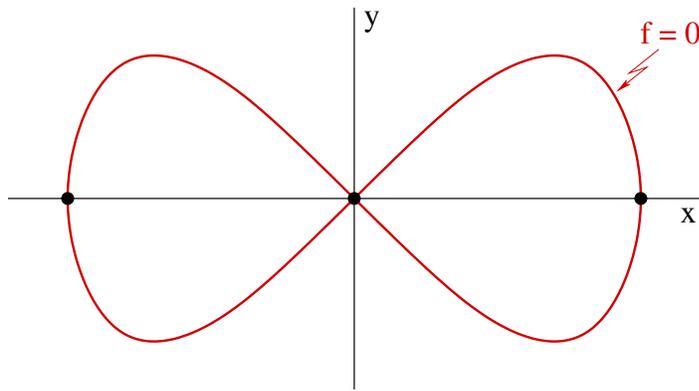


Abbildung 3.1: Eine implizit definiert Funktion.

Beispiel 3.2.1. Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ die Funktion

$$f(x, y) := x^2 - x^4 - y^2 \quad (3.2.2)$$

für $(x, y) \in \mathbb{R}^2$. Dann erfüllt jede Lösung der Gleichung (3.2.1) die Ungleichungen $|x| \leq 1$ und $|y| \leq 1/2$. Außerdem läßt sich die Gleichung explizit nach y auflösen, und man erhält die Formel

$$y = \pm x\sqrt{1 - x^2} \quad \text{für } -1 \leq x \leq 1. \quad (3.2.3)$$

Diese Formel stellt mit beiden Vorzeichen für $-1 < x < 0$ und für $0 < x < 1$ glatte Funktionen $y = g(x)$ dar. Die drei Ausnahmepunkte in der Lösungsmenge sind $(x, y) = (\pm 1, 0)$ und $(x, y) = (0, 0)$, und dies sind genau die Elemente des \mathbb{R}^2 , die die Bedingungen $f(x, y) = 0$ und $\frac{\partial f}{\partial y}(x, y) = 0$ erfüllen (siehe Abbildung 3.1).

Die allgemeine Situation in diesem Abschnitt betrifft stetig differenzierbare Abbildungen $f = (f_1, \dots, f_m) : \Omega \rightarrow \mathbb{R}^m$, die auf einer offenen Teilmenge $\Omega \subset \mathbb{R}^n \times \mathbb{R}^m$ definiert sind. Die Elemente von $\mathbb{R}^n \times \mathbb{R}^m$ bezeichnen wir mit $(x, y) = (x_1, \dots, x_n, y_1, \dots, y_m)$. In dieser Situation ist es nützlich, die Jacobi-Matrix von f an einer Stelle $(x, y) \in \Omega$ als Blockmatrix

$$df(x, y) = \left(\frac{\partial f}{\partial x}(x, y) \mid \frac{\partial f}{\partial y}(x, y) \right) \in \mathbb{R}^{m \times (n+m)}$$

zu schreiben mit

$$\frac{\partial f}{\partial x}(x, y) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x, y) & \cdots & \frac{\partial f_1}{\partial x_n}(x, y) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x, y) & \cdots & \frac{\partial f_m}{\partial x_n}(x, y) \end{pmatrix} \in \mathbb{R}^{m \times n},$$

$$\frac{\partial f}{\partial y}(x, y) := \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(x, y) & \cdots & \frac{\partial f_1}{\partial y_m}(x, y) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial y_1}(x, y) & \cdots & \frac{\partial f_m}{\partial y_m}(x, y) \end{pmatrix} \in \mathbb{R}^{m \times m}.$$

Damit wir das Gleichungssystem $f(x, y) = 0$ lokal nach y auflösen können, ist die entscheidende Bedingung, dass die quadratische Matrix $\frac{\partial f}{\partial y}(x, y)$ eine von Null verschiedene Determinante hat. Nun hat die Richtungsableitung von f an der Stelle $(x, y) \in \Omega$ in Richtung eines Vektors $(\xi, \eta) \in \mathbb{R}^n \times \mathbb{R}^m$ die Form

$$\begin{aligned} df(x, y) \begin{pmatrix} \xi \\ \eta \end{pmatrix} &= \frac{\partial f}{\partial x}(x, y)\xi + \frac{\partial f}{\partial y}(x, y)\eta \\ &= \frac{d}{dt} \Big|_{t=0} f(x + t\xi, y) + \frac{d}{dt} \Big|_{t=0} f(x, y + t\eta). \end{aligned} \tag{3.2.4}$$

Die gleiche Formel erhält man, wenn X, Y, Z endlichdimensionale normierte Vektorräume sind und $f : \Omega \rightarrow Z$ eine C^1 -Abbildung auf einer offenen Teilmenge $\Omega \subset X \times Y$ ist. In dem Fall definiert man die linearen Abbildungen

$$d^X f(x, y) := \frac{\partial f}{\partial x}(x, y) \in \mathcal{L}(X, Z), \quad d^Y f(x, y) := \frac{\partial f}{\partial y}(x, y) \in \mathcal{L}(Y, Z)$$

für $(x, y) \in \Omega$ durch die rechte Seite der Gleichung (3.2.4).

Satz 3.2.2 (Satz von der impliziten Funktion). *Seien X, Y, Z endlich-dimensionale normierte Vektorräume, sei $\Omega \subset X \times Y$ eine offene Teilmenge, sei $f : \Omega \rightarrow Z$ eine C^ℓ Abbildung mit $\ell \in \mathbb{N} \cup \{\infty\}$, und sei $(x_0, y_0) \in \Omega$ gegeben, so dass $f(x_0, y_0) = 0$ und die lineare Abbildung $d^Y f(x_0, y_0) \in \mathcal{L}(Y, Z)$ invertierbar ist. Dann existieren offene Teilmengen $U \subset X$ und $V \subset Y$ und eine C^ℓ -Abbildung $g : U \rightarrow V$ mit folgenden Eigenschaften.*

(i) *Es gilt $(x_0, y_0) \in U \times V \subset \Omega$ und*

$$g(x_0) = y_0. \quad (3.2.5)$$

(ii) *Für alle $x \in U$ und alle $y \in V$ gilt*

$$f(x, y) = 0 \quad \iff \quad g(x) = y. \quad (3.2.6)$$

(iii) *Für jedes $x \in X$ ist die lineare Abbildung $d^Y f(x, g(x)) \in \mathcal{L}(Y, Z)$ invertierbar und es gilt*

$$dg(x) = -d^Y f(x, g(x))^{-1} d^X f(x, g(x)) \quad (3.2.7)$$

Beweis. Wir verwenden die Abkürzungen

$$A := d^X f(x_0, y_0) \in \mathcal{L}(X, Z), \quad B := d^Y f(x_0, y_0) \in \mathcal{L}(Y, Z),$$

und definieren die Abbildung $F : \Omega \rightarrow X \times Z$ durch

$$F(x, y) := (x, f(x, y))$$

für $(x, y) \in \Omega$. Dann ist F eine C^ℓ -Abbildung und ihre Ableitung an der Stelle (x_0, y_0) ist durch

$$\begin{aligned} dF(x_0, y_0)(\xi, \eta) &= \left. \frac{d}{dt} \right|_{t=0} F(x_0 + t\xi, y_0 + t\eta) \\ &= \left. \frac{d}{dt} \right|_{t=0} (x_0 + t\xi, f(x_0 + t\xi, y_0 + t\eta)) \\ &= (\xi, d^X f(x_0, y_0)\xi + d^Y f(x_0, y_0)\eta) \\ &= (\xi, A\xi + B\eta) \end{aligned}$$

für $(\xi, \eta) \in X \times Y$ gegeben. Da $B : Y \rightarrow Z$ bijektiv ist, ist auch die Ableitung $dF(x_0, y_0) : X \times Y \rightarrow X \times Z$ bijektiv mit der Umkehrabbildung

$$dF(x_0, y_0)^{-1}(\xi, \zeta) = (\xi, B^{-1}(\zeta - A\xi))$$

für $\xi \in X$ und $\zeta \in Z$.

Nach Satz 3.1.11 existiert eine offene Menge $\Omega_0 \subset \Omega$ mit $(x_0, y_0) \in \Omega_0$, so dass $F(\Omega_0) \subset X \times Z$ offen und $F_0 := F|_{\Omega_0} : \Omega_0 \rightarrow F(\Omega_0)$ ein C^ℓ -Diffeomorphismus ist. Dann ist also $F_0^{-1} : F(\Omega_0) \rightarrow \Omega_0$ eine C^ℓ -Abbildung mit

$$(x_0, 0) \in F(\Omega_0), \quad F_0^{-1}(x_0, 0) = (x_0, y_0). \quad (3.2.8)$$

Da Ω_0 eine offene Teilmenge von $X \times Y$ ist, existiert ein $\varepsilon > 0$ mit

$$B_\varepsilon(x_0; X) \times B_\varepsilon(y_0; Y) \subset \Omega_0. \quad (3.2.9)$$

Da die Abbildung $F_0^{-1} : F(\Omega_0) \rightarrow \Omega_0$ stetig ist und $F_0^{-1}(x_0, 0) = (x_0, y_0)$ ist, ist $F(B_\varepsilon(x_0; X) \times B_\varepsilon(y_0; Y)) = (F_0^{-1})^{-1}(B_\varepsilon(x_0; X) \times B_\varepsilon(y_0; Y)) \subset X \times Z$ eine offene Umgebung von $(x_0, 0)$. Daher existiert eine Konstante $0 < \delta \leq \varepsilon$ mit

$$B_\delta(x_0; X) \times B_\delta(0; Z) \subset F(B_\varepsilon(x_0; X) \times B_\varepsilon(y_0; Y)) \subset F(\Omega_0). \quad (3.2.10)$$

Nun definieren wir die Abbildungen $\iota_X : X \rightarrow X \times Z$ und $\pi_Y : X \times Y \rightarrow Y$ durch $\iota_X(x) := (x, 0)$ für $x \in X$ und $\pi_Y(x, y) := y$ für $(x, y) \in X \times Y$. Diese Abbildungen sind beide linear und daher glatt. Ist nun $x \in B_\delta(x_0; X)$, so ist $\iota_X(x) = (x, 0) \in F(B_\varepsilon(x_0; X) \times B_\varepsilon(y_0; Y))$ nach (3.2.10) und damit gilt $\pi_Y \circ F_0^{-1} \circ \iota_X(x) \in B_\varepsilon(y_0; Y)$. Also erhalten wir eine wohldefinierte Abbildung $g : U \rightarrow V$ durch

$$U := B_\delta(x_0; X), \quad V := B_\varepsilon(y_0; Y), \quad g(x) := \pi_Y \circ F_0^{-1} \circ \iota_X(x) \quad (3.2.11)$$

für $x \in U$. Dies ist eine C^ℓ -Abbildung nach Satz 2.5.9. Weiterhin gilt

$$(x_0, y_0) \in U \times V \subset \Omega_0 \subset \Omega$$

nach (3.2.9) und $g(x_0) = \pi_Y(F_0^{-1}(x_0, 0)) = \pi_Y(x_0, y_0) = y_0$ nach (3.2.8). Damit erfüllt g die Bedingung (i). Für den Beweis von (ii) halten wir zwei Elemente $x \in U$ und $y \in V$ fest. Dann sind $(x, y), (x, g(x)) \in \Omega_0$ und es gelten die Gleichungen $F(x, y) = (x, f(x, y))$ und $F(x, g(x)) = (x, 0)$. Da $F|_{\Omega_0}$ injektiv ist, gilt also $g(x) = y$ genau dann, wenn $f(x, y) = 0$ ist. Damit erfüllt g die Bedingung (ii). Für den Beweis von (iii) halten wir ein Element $x \in U$ fest. Dann ist $(x, g(x)) \in \Omega_0$ und nach Lemma 3.1.2 ist daher die lineare Abbildung $dF(x, g(x)) : X \times Y \rightarrow X \times Z$ invertierbar. Mit

$$dF(x, g(x))(\xi, \eta) = (\xi, d^X f(x, g(x))\xi + d^Y f(x, g(x))\eta)$$

für $(\xi, \eta) \in X \times Y$ folgt daraus, dass $d^Y f(x, g(x)) : Y \rightarrow Z$ invertierbar ist. Differenzieren wir nun die Gleichung $f(x, g(x)) = 0$, so erhalten wir

$$0 = d^X f(x, g(x))\xi + d^Y f(x, g(x))dg(x)\xi$$

für alle $\xi \in X$ und daraus folgt (3.2.7). Damit ist Satz 3.2.2 bewiesen. \square

Beispiel 3.2.3. Sei $\Omega \subset \mathbb{R}^n$ eine offene Menge mit $x_0 \in \Omega$, sei $\Lambda \subset \mathbb{R}^N$ eine offene Menge mit $0 \in \Lambda$, und sei $f : \Lambda \times \Omega \rightarrow \mathbb{R}$ eine C^2 -Funktion. Für $\lambda \in \Lambda$ definieren wir die C^2 -Funktion $f_\lambda : \Omega \rightarrow \mathbb{R}$ durch

$$f_\lambda(x) := f(\lambda, x) \quad \text{für } x \in \Omega.$$

Sei $x_0 \in \Omega_0$ ein **nichtdegenerierter kritischer Punkt** von f_0 , das heißt

$$df_0(x_0) = 0, \quad \det(d^2f_0(x_0)) \neq 0. \quad (3.2.12)$$

(Siehe auch Übung 2.7.17 und Übung 3.1.17.) Dann existieren offene Umgebungen $\Lambda_0 \subset \Lambda$ von $0 \in \mathbb{R}^N$ und $U_0 \subset \Omega$ von $x_0 \in \mathbb{R}^n$, sowie eine C^1 -Abbildung $g : \Lambda_0 \rightarrow U_0$ mit den folgenden Eigenschaften.

(i) $g(0) = x_0$.

(ii) Für alle $\lambda \in \Lambda_0$ und alle $x \in U_0$ gilt $x = g(\lambda) \iff df_\lambda(x) = 0$.

(iii) Für $\lambda \in \Lambda_0$ ist $g(\lambda)$ ein nichtdegenerierter kritischer Punkt von f_λ und

$$\frac{\partial g}{\partial \lambda_k}(\lambda) = - (d^2f_\lambda(g(\lambda)))^{-1} \frac{\partial \nabla f}{\partial \lambda_k}(\lambda, g(\lambda)), \quad k = 1, \dots, N.$$

(iv) Ist f glatt so ist auch g glatt.

Die Aussagen (i-iii) folgen aus Satz 3.2.2, wobei $(n, m, x, y, \Omega, X, Y, Z, f)$ durch das Tupel $(N, n, \lambda, x, \Lambda \times \Omega, \mathbb{R}^N, \mathbb{R}^n, \mathbb{R}^n, \nabla f)$ zu ersetzen ist. Teil (iv) ist eine Übung. Zusammenfassend zeigt dieses Beispiel, dass nichtdegenerierte kritische Punkte unter kleinen Störungen der Funktion f_0 erhalten bleiben.

Beispiel 3.2.4 (Legendre-Transformation). Sei $\Omega \subset \mathbb{R}^n \times \mathbb{R}^n$ eine offene Menge. Wir bezeichnen die Elemente von Ω mit

$$(x, v) = (x_1, \dots, x_n, v_1, \dots, v_n).$$

Sei $L : \Omega \rightarrow \mathbb{R}$ eine glatte Funktion, welche die Bedingung

$$\det \left(\left(\frac{\partial^2 L}{\partial v_i \partial v_j}(x, v) \right)_{i,j=1}^n \right) \neq 0 \quad (3.2.13)$$

für alle $(x, v) \in \Omega$ erfüllt. Dies ist die **Legendre-Bedingung**. Im folgenden Verwenden wir für $(x, v) \in \Omega$ die Notation

$$\begin{aligned} \frac{\partial L}{\partial x}(x, v) &:= \left(\frac{\partial L}{\partial x_1}(x, v), \dots, \frac{\partial L}{\partial x_n}(x, v) \right), \\ \frac{\partial L}{\partial v}(x, v) &:= \left(\frac{\partial L}{\partial v_1}(x, v), \dots, \frac{\partial L}{\partial v_n}(x, v) \right). \end{aligned}$$

Sei nun $(x_0, v_0) \in \Omega$ und $y_0 := \frac{\partial L}{\partial v}(x_0, v_0)$. Dann folgt aus der Legendre-Bedingung (3.2.13) und Satz 3.2.2, dass sich das Gleichungssystem

$$y = \frac{\partial L}{\partial v}(x, v) \quad (3.2.14)$$

lokal nach v auflösen läßt. Es existieren also offene Mengen $U, V, W \subset \mathbb{R}^n$ mit

$$x_0 \in U, \quad v_0 \in V, \quad y_0 \in W, \quad U \times V \subset \Omega$$

sowie eine glatte Abbildung $g : U \times W \rightarrow V$, so dass $g(x_0, y_0) = v_0$ ist und die folgende Aussage für alle $x \in U$, $v \in V$ und $y \in W$ erfüllt ist:

$$v = g(x, y) \quad \iff \quad \frac{\partial L}{\partial v}(x, v) = y. \quad (3.2.15)$$

Nun definieren wir die **Hamilton-Funktion** $H : U \times W \rightarrow \mathbb{R}$ durch

$$H(x, y) := \sum_{j=1}^n y_j g_j(x, y) - L(x, g(x, y)) \quad (3.2.16)$$

für $x \in U$ und $y \in W$. Mit Hilfe der Gleichung $\frac{\partial L}{\partial v_j}(x, g(x, y)) = y_j$ ergeben sich dann für die partiellen Ableitungen der Hamilton-Funktion die Formeln

$$\begin{aligned} \frac{\partial H}{\partial x_i}(x, y) &= \frac{\partial}{\partial x_i} \left(\sum_{j=1}^n y_j g_j(x, y) - L(x, g(x, y)) \right) \\ &= \sum_{j=1}^n y_j \frac{\partial g_j}{\partial x_i}(x, y) - \frac{\partial L}{\partial x_i}(x, g(x, y)) - \sum_{j=1}^n \frac{\partial L}{\partial v_j}(x, g(x, y)) \frac{\partial g_j}{\partial x_i}(x, y) \\ &= -\frac{\partial L}{\partial x_i}(x, g(x, y)), \\ \frac{\partial H}{\partial y_i}(x, y) &= \frac{\partial}{\partial y_i} \left(\sum_{j=1}^n y_j g_j(x, y) - L(x, g(x, y)) \right) \\ &= g_i(x, y) + \sum_{j=1}^n y_j \frac{\partial g_j}{\partial y_i}(x, y) - \sum_{j=1}^n \frac{\partial L}{\partial v_j}(x, g(x, y)) \frac{\partial g_j}{\partial y_i}(x, y) \\ &= g_i(x, y). \end{aligned}$$

Zusammenfassend erhalten wir also für $x \in U$ und $y \in W$ die Gleichungen

$$\frac{\partial H}{\partial x_i}(x, y) = -\frac{\partial L}{\partial x_i}(x, g(x, y)), \quad \frac{\partial H}{\partial y_i}(x, y) = g_i(x, y). \quad (3.2.17)$$

Sei nun $I \subset \mathbb{R}$ ein offenes Intervall und $x : I \rightarrow U$ eine (glatte) Lösung der **Euler–Lagrange-Gleichungen**

$$\frac{d}{dt} \frac{\partial L}{\partial v_i}(x(t), \dot{x}(t)) = \frac{\partial L}{\partial x_i}(x(t), \dot{x}(t)), \quad i = 1, \dots, n \quad (3.2.18)$$

mit $\dot{x}(t) \in V$ und $y(t) \in W$ für alle $t \in I$, wobei $y(t)$ durch

$$y_i(t) := \frac{\partial L}{\partial v_i}(x(t), \dot{x}(t)), \quad i = 1, \dots, n \quad (3.2.19)$$

definiert ist. Dann gilt $g_i(x(t), y(t)) = \dot{x}_i(t)$ für $i = 1, \dots, n$ nach (3.2.15), und daraus folgt nach (3.2.17), dass die Abbildung $(x, y) : I \rightarrow U \times W$ die **Hamiltonschen Gleichungen**

$$\dot{x}_i(t) = \frac{\partial H}{\partial y_i}(x(t), y(t)), \quad \dot{y}_i(t) = -\frac{\partial H}{\partial x_i}(x(t), y(t)) \quad (3.2.20)$$

für $i = 1, \dots, n$ und $t \in I$ löst. Zusammenfassend heißt das, dass die Legendre-Transformation (3.2.14) die Euler–Lagrange-Gleichungen (3.2.18) (eine gewöhnliche Differentialgleichung zweiter Ordnung in n Variablen) in die Hamiltonschen Gleichungen (3.2.20) (eine gewöhnliche Differentialgleichung erster Ordnung in $2n$ Variablen) mit der Hamilton-Funktion (3.2.16) überführt. Darüber hinaus hat die Hamilton-Funktion H die physikalische Bedeutung der *Energie*, die entlang der Lösungen der Gleichung (3.2.20), und damit auch entlang der Lösungen von (3.2.18), konstant ist (Übung).

Beispiel 3.2.5. Sei $\Omega \subset \mathbb{R}^n$ eine offene Menge und sei $V : \Omega \rightarrow \mathbb{R}$ eine glatte Funktion. Dann entsprechen die **Newtonschen Gleichungen**

$$\ddot{x}(t) = -\nabla V(x(t)) \quad (3.2.21)$$

dem Spezialfall von Gleichung (3.2.18), wo $L : \Omega \times \mathbb{R}^n \rightarrow \mathbb{R}$ durch

$$L(x, v) = \frac{1}{2} \|v\|^2 - V(x) \quad (3.2.22)$$

(*kinetische* minus *potentielle* Energie) gegeben ist. In diesem Fall hat die Gleichung (3.2.14) die Form $y = \frac{\partial L}{\partial v}(x, v) = v$, die Hamilton-Funktion ist

$$H(x, y) = \frac{1}{2} \|y\|^2 + V(x) \quad (3.2.23)$$

(*kinetische* plus *potentielle* Energie), und die Hamiltonschen Gleichungen liefern nichts Neues im Vergleich zu (3.2.21). Jedoch ergibt sich aus (3.2.20) unmittelbar der *Energieerhaltungssatz*. (Siehe auch Beispiel 2.3.8.)

3.3 Untermannigfaltigkeiten

Der Begriff einer Untermannigfaltigkeit

In diesem Abschnitt wird der Begriff einer Untermannigfaltigkeit (eines endlichdimensionalen normierten Vektorraumes) eingeführt und es werden einige Beispiele diskutiert. Dieses Konzept spielt eine grundlegende Rolle für viele Gebiete der Mathematik und Physik. In diesem Manuskript wird es in Abschnitt 3.4 für die Untersuchung von Extrema mit Nebenbedingungen verwendet, ebenso wie in Kapitel 6 über die Integration von Funktionen auf Mannigfaltigkeiten und den Divergenzsatz von Gauß und in Kapitel 7 über die Integration von Differentialformen und den Satz von Stokes.

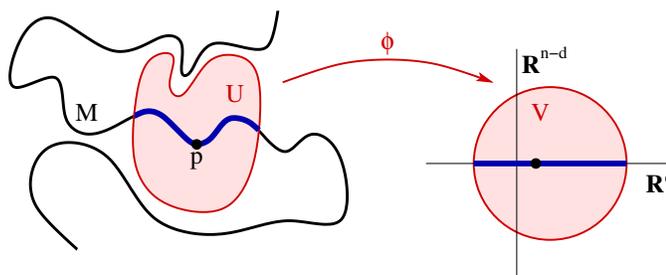


Abbildung 3.2: Eine Untermannigfaltigkeit.

Definition 3.3.1 (Untermannigfaltigkeit). Seien d und n ganze Zahlen mit $0 \leq d \leq n$, sei $\ell \in \mathbb{N} \cup \{\infty\}$, und sei X ein n -dimensionaler normierter Vektorraum. Eine Teilmenge $M \subset X$ heißt **d -dimensionale C^ℓ -Untermannigfaltigkeit** von X , wenn sie folgende Bedingung erfüllt. Für jeden Punkt $p \in M$ existiert eine offene Teilmenge $U \subset X$ mit $p \in U$, eine offene Teilmenge $V \subset \mathbb{R}^n$, und ein C^ℓ -Diffeomorphismus $\phi : U \rightarrow V$ mit

$$\phi(U \cap M) = V \cap (\mathbb{R}^d \times \{0\}). \quad (3.3.1)$$

(Siehe Abbildung 3.2. Im Fall $\ell = \infty$ sprechen wir auch von einer **glatten d -dimensionalen Untermannigfaltigkeit** von X .) In dieser Situation wird der C^ℓ -Diffeomorphismus ϕ eine **Karte von M** und die Menge $U \cap M$ ein **Kartengebiet von M** genannt. Ist I eine Indexmenge und $\phi_i : U_i \rightarrow V_i$ eine Karte von M für jedes $i \in I$, so dass die Kartengebiete ganz M überdecken, das heißt $M \subset \bigcup_{i \in I} U_i$, so wird die Kollektion $\{\phi_i\}_{i \in I}$ dieser Karten ein **Atlas von M** genannt.

Beispiel 3.3.2. Im Fall $d = 0$ ist $M \subset X$ genau dann eine 0-dimensionale C^ℓ -Untermannigfaltigkeit von X , wenn M eine **diskrete Teilmenge** von X ist, das heißt, wenn für jeden Punkt $p \in M$ eine offene Teilmenge $U \subset X$ existiert mit $U \cap M = \{p\}$.

Beispiel 3.3.3. Im Fall $d = n = \dim(X)$ ist $M \subset X$ genau dann eine n -dimensionale C^ℓ -Untermannigfaltigkeit von X , wenn M eine offene Teilmenge von X ist.

Beispiel 3.3.4. Die Einheitssphäre

$$S^{n-1} = \left\{ x \in \mathbb{R}^n \mid \sum_{i=1}^n x_i^2 = 1 \right\}$$

ist eine glatte $(n-1)$ -dimensionale Untermannigfaltigkeit des \mathbb{R}^n . Ein Atlas mit $2n$ Karten $\phi_i^\pm : U_i^\pm \rightarrow V^\pm$ ist durch

$$U_i^\pm := \left\{ x \in \mathbb{R}^n \mid \pm x_i > 0, \sum_{j \neq i} x_j^2 < 1 \right\},$$

$$V^\pm := \left\{ y \in \mathbb{R}^n \mid \sum_{i=1}^{n-1} y_i^2 < 1, \sqrt{1 - \sum_{i=1}^{n-1} y_i^2} \pm y_n > 0 \right\},$$

$$\phi_i^\pm(x) := \left(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n, x_i \mp \sqrt{1 - \sum_{j \neq i} x_j^2} \right)$$

für $i = 1, \dots, n$ und $x \in U_i^\pm$ gegeben.

Beispiel 3.3.5. Seien X und Y normierte Vektorräume mit $\dim(X) = d$ und $\dim(Y) = m$, sei $\Omega \subset X$ eine offene Teilmenge, und sei $g : \Omega \rightarrow Y$ eine C^ℓ -Abbildung. Dann ist die Menge

$$M := \text{graph}(g) = \{(x, y) \in X \times Y \mid x \in \Omega, y = g(x)\}$$

eine d -dimensionale C^ℓ -Untermannigfaltigkeit von $X \times Y$. Sie besitzt einen Atlas, der aus einer einzigen Karte $\phi : U \rightarrow V$ besteht. Dazu wählen wir Vektorraumisomorphismen $i : X \rightarrow \mathbb{R}^d$ und $j : Y \rightarrow \mathbb{R}^m$, und definieren

$$U := \Omega \times Y, \quad V := i(\Omega) \times \mathbb{R}^m, \quad \phi(x, y) := (i(x), j(y - g(x)))$$

für $x \in \Omega$ und $y \in Y$.

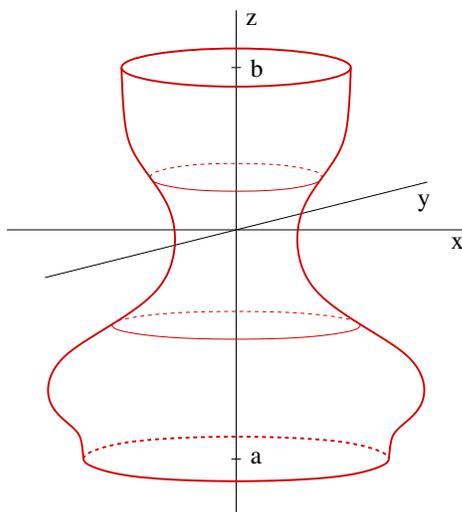


Abbildung 3.3: Eine Rotationsfläche.

Beispiel 3.3.6. Seien $a < b$ reelle Zahlen und sei $r : (a, b) \rightarrow (0, \infty)$ eine glatte Funktion. Dann ist die Menge

$$M := \{(x, y, z) \in \mathbb{R}^3 \mid a < z < b, x^2 + y^2 = r(z)^2\} \quad (3.3.2)$$

eine glatte 2-dimensionale Untermannigfaltigkeit des \mathbb{R}^3 . Jede Untermannigfaltigkeit dieser Form wird **Rotationsfläche** genannt (siehe Abbildung 3.3).

Übung: Konstruieren Sie einen Atlas von M .

Beispiel 3.3.7. Die Menge

$$S := \{(x, y) \in \mathbb{R}^2 \mid y \neq 0 \implies 1/y \in \mathbb{N}\}$$

ist keine Untermannigfaltigkeit des \mathbb{R}^2 . Zum Beweis kann man den Zusammenhangsbegriff in Anhang B verwenden. Ist $M \subset X$ eine Untermannigfaltigkeit und $p \in M$, so kann man eine Karte $\phi : U \rightarrow V$ von M mit $p \in U$ so wählen, dass $V = B_\varepsilon(\phi(p))$ ein Ball im \mathbb{R}^n mit Mittelpunkt $\phi(p)$ ist. Dann ist $V \cap (\mathbb{R}^d \times \{0\})$ eine zusammenhängende Teilmenge des \mathbb{R}^n und nach (3.3.1) ist damit auch das Kartengebiet $U \cap M$ zusammenhängend. In unserem Beispiel ist aber die Menge $U \cap S$ für jede offene Umgebung $U \subset \mathbb{R}^2$ des Punktes $(0, 0) \in S$ unzusammenhängend (Übung), und daher kann S keine Untermannigfaltigkeit des \mathbb{R}^2 sein.

Beispiel 3.3.8. Die Menge $S := \{(x, y) \in \mathbb{R}^2 \mid xy = 0\}$ ist keine Untermannigfaltigkeit des \mathbb{R}^2 . Zum Beweis kann man den folgenden Satz verwenden.

Reguläre Werte

Satz 3.3.9. *Sei X ein n -dimensionaler normierter Vektorraum und $M \subset X$ eine nichtleere Teilmenge. Sei $\ell \in \mathbb{N} \cup \{\infty\}$ und sei d eine ganze Zahl mit $0 \leq d < n$. Dann sind folgende Aussagen äquivalent.*

- (i) *M ist eine d -dimensionale C^ℓ -Untermannigfaltigkeit von X .*
- (ii) *Für jeden Punkt $p \in M$ existiert eine offene Teilmenge $U \subset X$ mit $p \in U$ und C^ℓ -Funktionen $f_1, \dots, f_{n-d} : U \rightarrow \mathbb{R}$, so dass*

$$U \cap M = \{x \in U \mid f_1(x) = f_2(x) = \dots = f_{n-d}(x) = 0\} \quad (3.3.3)$$

ist und die linearen Funktionale $df_1(x), df_2(x), \dots, df_{n-d}(x) \in \mathcal{L}(X, \mathbb{R})$ für jeden Punkt $x \in U \cap M$ linear unabhängig sind.

Der Beweis verwendet das folgende Lemma aus der Linearen Algebra.

Lemma 3.3.10. *Sei X ein Vektorraum und seien $\Lambda_1, \dots, \Lambda_m : X \rightarrow \mathbb{R}$ lineare Funktionale. Dann sind folgende Aussagen äquivalent.*

- (i) *Die linearen Funktionale $\Lambda_1, \dots, \Lambda_m$ sind linear unabhängig.*
- (ii) *Es gibt Vektoren $x_1, \dots, x_m \in X$ mit $\Lambda_i(x_j) = \delta_{ij}$ für $i, j = 1, \dots, m$.*
- (iii) *Die Abbildung $X \rightarrow \mathbb{R}^m : x \mapsto (\Lambda_1(x), \dots, \Lambda_m(x))$ ist surjektiv.*

Beweis. Wir beweisen (i) \implies (ii) durch vollständige Induktion über m . Für $m = 1$ bedeutet lineare Unabhängigkeit, dass $\Lambda_1 \neq 0$ ist, und daher existiert ein Vektor $x_1 \in X$ mit $\Lambda_1(x_1) = 1$. Nun sei $m \geq 2$ und wir nehmen an, dass die Aussage für $m - 1$ gültig ist. Dann zeigen wir

$$\bigcap_{j \neq i} \ker \Lambda_j \not\subset \ker \Lambda_i \quad \text{für } i = 1, \dots, m. \quad (3.3.4)$$

Nach Induktionsannahme gibt es Vektoren $\xi_1, \dots, \xi_{m-1} \in X$ mit $\Lambda_i(\xi_j) = \delta_{ij}$ für $i, j = 1, \dots, m - 1$. Gilt die Aussage (3.3.4) nicht für $i = m$, so ergibt sich für jedes $x \in X$, dass $x - \sum_{i=1}^{m-1} \Lambda_i(x)\xi_i \in \bigcap_{j=1}^{m-1} \ker \Lambda_j \subset \ker \Lambda_m$ ist und daher $\Lambda_m(x) = \sum_{i=1}^m \Lambda_i(x)\Lambda_m(\xi_i)$ gilt, woraus $\Lambda_m = \sum_{i=1}^{m-1} \Lambda_m(\xi_i)\Lambda_i$ folgt, im Widerspruch zu (i). Damit ist (3.3.4) für $i = m$ bewiesen. Durch Vertauschen von i und m erhalten wir (3.3.4) für alle i , und daraus folgt sofort Teil (ii).

Wir beweisen (ii) \implies (iii). Seien $x_1, \dots, x_m \in X$ wie in Teil (ii) und sei $\lambda = (\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$. Dann erfüllt der Vektor $x := \sum_{j=1}^m \lambda_j x_j$ die Bedingung $\Lambda_i(x) = \sum_{j=1}^m \lambda_j \Lambda_i(x_j) = \sum_{j=1}^m \lambda_j \delta_{ij} = \lambda_i$ für $i = 1, \dots, m$.

Wir beweisen (iii) \implies (i). Seien $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ mit $\sum_{i=1}^m \lambda_i \Lambda_i = 0$. Nach (iii) existiert dann ein Vektor $x \in X$ mit $\Lambda_i(x) = \lambda_i$ für $i = 1, \dots, m$. Daraus folgt $0 = \sum_{i=1}^m \lambda_i \Lambda_i(x) = \sum_{i=1}^m \lambda_i^2$ und daher $\lambda_i = 0$ für alle i . \square

Beweis von Satz 3.3.9. Wir beweisen (i) \implies (ii). Sei $p \in M$. Dann existiert nach (i) eine offene Menge $U \subset X$ mit $p \in U$, eine offene Menge $V \subset \mathbb{R}^n$, und ein C^ℓ -Diffeomorphismus $\phi = (\phi_1, \dots, \phi_n) : U \rightarrow V$ der (3.3.1) erfüllt. Da die lineare Abbildung $d\phi(x) : X \rightarrow \mathbb{R}^n$ nach Lemma 3.1.2 bijektiv ist, sind die linearen Funktionale $d\phi_1(x), \dots, d\phi_n(x) : X \rightarrow \mathbb{R}$ nach Lemma 3.3.10 für jedes $x \in U$ linear unabhängig. Ausserdem gilt für jedes $x \in U$ nach (3.3.1)

$$x \in M \cap U \iff \phi(x) \in \mathbb{R}^d \times \{0\} \iff \phi_{d+1}(x) = \dots = \phi_n(x) = 0.$$

Damit erfüllen die Funktionen $f_i := \phi_{d+i}$ für $i = 1, \dots, n-d$ Teil (ii).

Wir beweisen (ii) \implies (i). Sei $p \in M$ und seien $f_1, \dots, f_{n-d} : U \rightarrow \mathbb{R}$ wie in (ii). Dann existieren lineare Funktionale $\lambda_1, \dots, \lambda_d : X \rightarrow \mathbb{R}$ so dass

$$\lambda_1, \dots, \lambda_d, df_1(p), \dots, df_{n-d}(p)$$

eine Basis des n -dimensionalen Vektorraumes $\mathcal{L}(X, \mathbb{R})$ ist. Nun definieren wir die Abbildung $\phi : U \rightarrow \mathbb{R}^n$ durch

$$\phi(x) := (\lambda_1(x), \dots, \lambda_d(x), f_1(x), \dots, f_{n-d}(x))$$

für $x \in U$. Diese ist C^ℓ nach Satz 2.5.9 und ihre Ableitung $d\phi(p) : X \rightarrow \mathbb{R}^n$ an der Stelle p ist

$$d\phi(p)\xi = (\lambda_1(\xi), \dots, \lambda_d(\xi), df_1(p)\xi, \dots, df_{n-d}(p)\xi) \quad \text{für } x \in U.$$

Da die linearen Funktionale $\lambda_1, \dots, \lambda_d, df_1(p), \dots, df_{n-d}(p)$ linear unabhängig sind, ist $d\phi(p) : X \rightarrow \mathbb{R}^n$ bijektiv nach Lemma 3.3.10. Nach Satz 3.1.11 existiert daher eine offene Menge $U_0 \subset X$ mit $p \in U_0$, so dass $V_0 := \phi(U_0) \subset \mathbb{R}^n$ offen und $\phi_0 := \phi|_{U_0} : U_0 \rightarrow V_0$ ein C^ℓ -Diffeomorphismus ist. Dann gilt

$$\begin{aligned} \phi_0(U_0 \cap M) &= \{\phi(x) \mid x \in U_0, f_1(x) = \dots = f_{n-d}(x) = 0\} \\ &= V_0 \cap (\mathbb{R}^d \times \{0\}) \end{aligned}$$

und damit ist ϕ_0 eine Karte von M . Da $p \in M$ beliebig gewählt war, heißt das, dass M eine d -dimensionale C^ℓ -Untermannigfaltigkeit von X ist. Damit ist Satz 3.3.9 bewiesen. \square

Definition 3.3.11 (Regulärer Wert). Seien X, Y endlichdimensionale normierte Vektorräume, sei $\Omega \subset X$ eine offene Menge, und sei $f : \Omega \rightarrow Y$ eine stetig differenzierbare Abbildung. Ein Vektor $y \in Y$ heißt **regulärer Wert von f** wenn die Ableitung $df(x) : X \rightarrow Y$ für jedes Element $x \in \Omega$ mit $f(x) = y$ eine surjektive lineare Abbildung ist. Ein Vektor $y \in Y$ heißt **singulärer Wert von f** , wenn y kein regulärer Wert von f ist.

Korollar 3.3.12 (Reguläre Werte). Seien $\ell \in \mathbb{N} \cup \{\infty\}$ und $d, m, n \in \mathbb{Z}$ mit $0 \leq d \leq n$ und $m := n - d$. Seien X und Y normierte Vektorräume der Dimensionen $\dim(X) = n$ und $\dim(Y) = m$, sei $\Omega \subset X$ eine offene Menge, sei $f : \Omega \rightarrow Y$ eine C^ℓ -Abbildung, und sei $y \in Y$ ein regulärer Wert von f . Dann ist

$$M := f^{-1}(y) = \{x \in \Omega \mid f(x) = y\}$$

eine d -dimensionale C^ℓ -Untermannigfaltigkeit von X .

Beweis. Im Fall $d = n$ ist $Y = \{0\}$ und dann ist $M = \Omega$ eine offene Teilmenge von X wie behauptet. Sei also $d < n$ und damit $m = n - d > 0$. Sei e_1, \dots, e_m eine Basis von Y und seien die Funktionen $f_1, \dots, f_m : \Omega \rightarrow \mathbb{R}$ durch

$$f_1(x)e_1 + \dots + f_m(x)e_m := f(x) - y \quad \text{für } x \in \Omega$$

definiert. Diese Funktionen sind C^ℓ nach Satz 2.3.2. Ausserdem gilt

$$M = \{x \in \Omega \mid f(x) = y\} = \{x \in \Omega \mid f_1(x) = \dots = f_m(x) = 0\}.$$

Nun ist die lineare Abbildung $X \rightarrow \mathbb{R}^m : \xi \mapsto (df_1(x)\xi, \dots, df_m(x)\xi)$ für jedes $x \in M$ surjektiv, da y ein regulärer Wert von f ist. Nach Lemma 3.3.10 sind daher die linearen Funktionale $df_1(x), \dots, df_m(x)$ für jedes $x \in M$ linear unabhängig. Damit folgt Korollar 3.3.12 aus Satz 3.3.9. \square

Bemerkung 3.3.13. (i) Korollar 3.3.12 behält auch dann seine Gültigkeit, wenn $y \notin f(\Omega)$ und M daher die leere Menge ist. Denn die leere Menge $M = \emptyset \subset X$ ist eine d -dimensionale Untermannigfaltigkeit für jedes d .

(ii) Seien X, Y normierte Vektorräume, sei $\Omega \subset X$ eine offene Teilmenge, und sei $f : \Omega \rightarrow Y$ eine C^ℓ -Abbildung mit $\ell \geq 1$ und $\ell \geq \dim(X) - \dim(Y)$. Dann sagt der **Satz von Sard**, dass *fast alle* $y \in Y$ reguläre Werte von f sind. Das bedeutet, dass die Menge der singulären Werte von f eine Lebesgue-Nullmenge ist, das heißt, dass sie sich durch abzählbar viele Quader von beliebig kleinem Gesamtvolumen überdecken läßt. Insbesondere ist die Menge der regulären Werte von f dicht in Y .

(iii) Der Satz von Sard gilt auch im Fall $\dim(Y) > \dim(X)$. Dann ist die Menge der regulären Werte das Komplement der Bildmenge von f , da die Ableitung $df(x) : X \rightarrow Y$ aus Dimensionsgründen niemals surjektiv sein kann. Im Fall $\dim(Y) > \dim(X)$ sagt der Satz von Sard also, dass die Bildmenge $f(\Omega)$ jeder C^1 -Abbildung $f : \Omega \rightarrow Y$ eine Lebesgue-Nullmenge in Y ist. Dies folgt aus Teil (iii) von Satz 5.6.7, da Ω eine abzählbare Vereinigung kompakter Mengen ist. Der Beweis im hier relevanten Fall $\dim(Y) \leq \dim(X)$ ist erheblich schwieriger. Für glatte Funktionen findet sich ein Beweis in [3].

Beispiele

Beispiel 3.3.14. Dass die Rotationsfläche $M \subset \mathbb{R}^3$ eine glatte 2-dimensionale Untermannigfaltigkeit des \mathbb{R}^3 ist, folgt sofort aus Korollar 3.3.12 mit

$$\Omega := \{(x, y, z) \in \mathbb{R}^3 \mid a < z < b\}$$

und der durch

$$f(x, y, z) := x^2 + y^2 - r(z)^2$$

für $(x, y, z) \in \Omega$ definierten glatten Funktion $f : \Omega \rightarrow \mathbb{R}$. Da $r(z) > 0$ ist für alle $(x, y, z) \in \Omega$, ist 0 ein regulärer Wert von f , und daher ist $M = f^{-1}(0)$ eine glatte 2-dimensionale Untermannigfaltigkeit von \mathbb{R}^3 .

Beispiel 3.3.15. Dass die Menge $S = \{(x, y) \in \mathbb{R}^2 \mid xy = 0\}$ in Beispiel 3.3.8 keine Untermannigfaltigkeit ist, läßt sich mit Satz 3.3.9 beweisen. Da S weder offen noch diskret ist, käme nur die Zahl $d = 1$ für die Dimension in Frage. Ist aber $f : U \rightarrow \mathbb{R}$ eine C^1 Funktion auf einer offenen Menge $U \subset \mathbb{R}^2$ mit $(0, 0) \in U \cap M = \{(x, y) \in U \mid f(x, y) = 0\}$ so ist $\frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial y}(0, 0) = 0$. Der Punkt $p = (0, 0) \in S$ verletzt daher die Bedingung (ii) in Satz 3.3.9, und somit ist S keine Untermannigfaltigkeit von \mathbb{R}^2 .

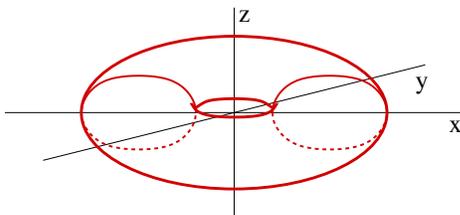


Abbildung 3.4: Ein 2-Torus im \mathbb{R}^3 .

Beispiel 3.3.16. Sei $0 < r < 1$ und sei $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ die glatte Funktion

$$f(x, y, z) := (x^2 + y^2 - 1 + r^2 - z^2)^2 - 4(x^2 + y^2)(r^2 - z^2).$$

Dann ist 0 ein regulärer Wert von f und daher ist $M := f^{-1}(0)$ eine glatte Untermannigfaltigkeit von \mathbb{R}^3 (siehe Abbildung 3.4). Die Abbildung

$$S^1 \times S^1 \rightarrow M : (e^{is}, e^{it}) \mapsto \begin{pmatrix} (1 + r \cos(s)) \cos(t) \\ (1 + r \cos(s)) \sin(t) \\ r \sin(s) \end{pmatrix}$$

identifiziert den Standard 2-Torus $S^1 \times S^1 \subset \mathbb{C} \times \mathbb{C}$ mit M .

Beispiel 3.3.17. Sind $M \subset \mathbb{R}^n$ und $M' \subset \mathbb{R}^{n'}$ glatte Untermannigfaltigkeiten so ist $M \times M'$ eine glatte Untermannigfaltigkeit von $\mathbb{R}^n \times \mathbb{R}^{n'} \cong \mathbb{R}^{n+n'}$. Insbesondere ist das n -fache Produkt

$$\mathbb{T}^n := S^1 \times \dots \times S^1$$

eine glatte Untermannigfaltigkeit von \mathbb{C}^n . Dies ist der **standard n -Torus**.

Beispiel 3.3.18. Die orthogonale Gruppe

$$\mathrm{O}(n) := \{A \in \mathbb{R}^{n \times n} \mid A^T A = \mathbb{1}\}$$

ist eine glatte Untermannigfaltigkeit von $\mathbb{R}^{n \times n}$ und

$$\dim(\mathrm{O}(n)) = \frac{n(n-1)}{2}.$$

Zum Beweis betrachten wir die normierten Vektorräume

$$X := \mathbb{R}^{n \times n}, \quad Y := \mathcal{S}_n := \{S \in \mathbb{R}^{n \times n} \mid S^T = S\}.$$

und die Abbildung $f : X \rightarrow Y$ mit

$$f(A) := A^T A \quad \text{für } A \in X$$

Wir zeigen, dass die Einheitsmatrix $\mathbb{1} \in Y$ ein regulärer Wert von f ist. Die Ableitung von f an der Stelle $A \in X$ ist die durch

$$df(A)\widehat{A} = A^T \widehat{A} + \widehat{A}^T A$$

gegebene lineare Abbildung $df(A) : X \rightarrow Y$. Sei $A \in \mathrm{O}(n)$ und $S = S^T \in \mathcal{S}_n$, und definiere

$$\widehat{A} := \frac{1}{2}AS \in \mathbb{R}^{n \times n}.$$

Dann gilt

$$df(A)\widehat{A} = A^T \widehat{A} + \widehat{A}^T A = \frac{1}{2}A^T AS + \frac{1}{2}(AS)^T A = \frac{1}{2}(S + S^T) = S.$$

Dies zeigt, dass die lineare Abbildung $df(A) : X \rightarrow Y$ für alle $A \in \mathrm{O}(n)$ surjektiv ist. Damit ist $\mathbb{1} \in Y$ ein regulärer Wert von f , wie behauptet, und nach Korollar 3.3.12 ist daher $\mathrm{O}(n) = f^{-1}(\mathbb{1})$ eine glatte Untermannigfaltigkeit von X der Dimension

$$\dim(\mathrm{O}(n)) = \dim(X) - \dim(Y) = n^2 - \frac{n(n+1)}{2} = \frac{n(n-1)}{2}.$$

Beispiel 3.3.19. Sei $0 \neq A = A^T \in \mathbb{R}^{n \times n}$ eine von Null verschiedene Symmetrische Matrix und sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ die glatte Funktion

$$f(x) := x^T A x \quad \text{für } x \in \mathbb{R}^n.$$

Dann ist jede Zahl $c \in \mathbb{R} \setminus \{0\}$ ein regulärer Wert von f (Übung). Daher ist

$$Q := f^{-1}(c) = \{x \in \mathbb{R}^n \mid x^T A x = c\}$$

für jede reelle Zahl $c \neq 0$ eine glatte $(n - 1)$ -dimensionale Untermannigfaltigkeit von \mathbb{R}^n ist. Hier sind einige Spezialfälle dieser Konstruktion.

(a) Mit $A = \text{diag}(1/r_1^2, \dots, 1/r_n^2)$ und $c = 1$ erhält man den Ellipsoid

$$E(r_1, \dots, r_n) := \left\{ x \in \mathbb{R}^n \mid \sum_{i=1}^n \frac{x_i^2}{r_i^2} = 1 \right\}.$$

(b) In (a) ergibt sich mit $r_i = 1$ die Einheitskugel $Q = S^{n-1}$ in Beispiel 3.3.4.

(c) Die Hyperbel $Q = \{(x, y) \in \mathbb{R}^2 \mid xy = c\}$ ist eine glatte 1-dimensionale Untermannigfaltigkeit von \mathbb{R}^2 für $c \neq 0$, aber nicht für $c = 0$ (Beispiel 3.3.14).

(d) Mit $f(x, y) := \|x - y\|^2$ für $x, y \in \mathbb{R}^3$ und $c = r^2$ mit $r > 0$ ergibt sich, dass die Menge $Q := \{(x, y) \in \mathbb{R}^3 \times \mathbb{R}^3 \mid \|x - y\| = r\}$ für jedes $r > 0$ eine glatte 5-dimensionale Untermannigfaltigkeit von \mathbb{R}^6 ist.

(e) In Modifikation des Beispiels in (d) kann man beweisen, dass die Menge

$$M := \{(x, y) \in \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \mid \|x\| = \|x - y\| = \|y - z\| = 1\}$$

eine glatte 6-dimensionale Untermannigfaltigkeit von \mathbb{R}^9 ist. Dies ist der Konfigurationsraum eines 3-Fach-Pendels.

Beispiel 3.3.20. Die Menge

$$S := \{(x^2, y^2, z^2, yz, zx, xy) \mid x, y, z \in \mathbb{R}, x^2 + y^2 + z^2 = 1\} \quad (3.3.5)$$

ist eine glatte 2-dimensionale Untermannigfaltigkeit von \mathbb{R}^6 (Übung). Dieses Beispiel geht auf Jakob Steiner zurück. Es handelt sich hier um eine *Einbettung* des 2-dimensionalen *reell projektiven Raumes* (der Menge aller Geraden im \mathbb{R}^3 durch den Ursprung) in den \mathbb{R}^6 . Projiziert man diese Untermannigfaltigkeit auf die letzten drei Koordinaten, so erhält man die sogenannte **Römerfläche**

$$\begin{aligned} R &= \{(yz, zx, xy) \mid x, y, z \in \mathbb{R}, x^2 + y^2 + z^2 = 1\} \\ &= \{(\xi, \eta, \zeta) \in \mathbb{R}^3 \mid \eta^2 \zeta^2 + \zeta^2 \xi^2 + \xi^2 \eta^2 = \xi \eta \zeta\}. \end{aligned}$$

Dies ist keine Untermannigfaltigkeit des \mathbb{R}^3 (Übung).

Der Tangentialraum

Definition 3.3.21 (Tangentialvektor).

Sei X ein endlichdimensionaler normierter Vektorraum und sei $M \subset X$ eine d -dimensionale C^1 -Untermannigfaltigkeit von X . Sei $p \in M$. Ein Vektor $v \in X$ heißt **Tangentialvektor von M im Punkte p** , wenn ein $\varepsilon > 0$ und eine C^1 -Abbildung $\gamma : (-\varepsilon, \varepsilon) \rightarrow M$ existiert mit

$$\gamma(0) = p, \quad \dot{\gamma}(0) = v. \quad (3.3.6)$$

(Siehe Abbildung 3.5.) Die Menge

$$T_p M := \{v \in X \mid v \text{ ist ein Tangentialvektor von } M \text{ im Punkte } p\} \quad (3.3.7)$$

wird der **Tangentialraum von M im Punkte p** genannt.

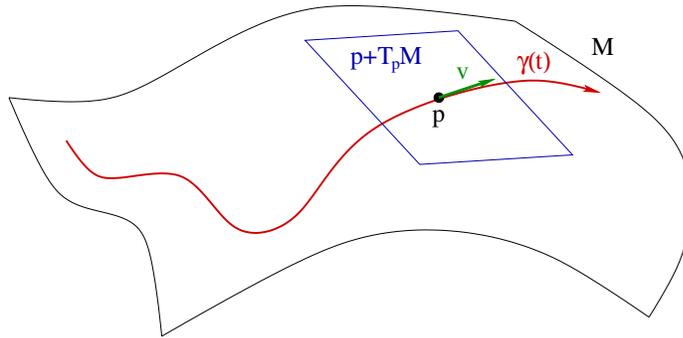


Abbildung 3.5: Der Tangentialraum.

Satz 3.3.22 (Der Tangentialraum). Sei X ein n -dimensionaler normierter Vektorraum und sei $M \subset X$ eine d -dimensionale C^ℓ -Untermannigfaltigkeit von X mit $\ell \in \mathbb{N} \cup \{\infty\}$. Sei $p \in M$. Dann gilt folgendes.

(i) Sind $U \subset X$ und $V \subset \mathbb{R}^n$ offene Mengen mit $p \in U$ und ist $\phi : U \rightarrow V$ ein C^ℓ -Diffeomorphismus mit $\phi(U \cap M) = V \cap (\mathbb{R}^d \times \{0\})$, so gilt

$$T_p M = d\phi(p)^{-1}(\mathbb{R}^d \times \{0\}). \quad (3.3.8)$$

(ii) Ist $U \subset X$ offen mit $p \in U$ und ist $f : U \rightarrow \mathbb{R}^{n-d}$ eine C^ℓ -Abbildung mit $U \cap M = f^{-1}(0)$, so dass $df(p) : X \rightarrow \mathbb{R}^{n-d}$ surjektiv ist, so gilt

$$T_p M = \ker df(p). \quad (3.3.9)$$

(iii) $T_p M$ ist ein d -dimensionaler linearer Unterraum von X .

(iv) Ist $v \in T_p M$, so existiert eine C^ℓ -Kurve $\gamma : \mathbb{R} \rightarrow M$, die (3.3.6) erfüllt.

Beweis. In der Situation von Teil (i) zeigen wir zunächst die Inklusion

$$d\phi(p)^{-1}(\mathbb{R}^d \times \{0\}) \subset T_p M \quad (3.3.10)$$

Sei also $v \in d\phi(p)^{-1}(\mathbb{R}^d \times \{0\})$ gegeben und sei

$$x := \phi(p) \in V \cap (\mathbb{R}^d \times \{0\}), \quad \xi := d\phi(p)v \in \mathbb{R}^d \times \{0\}. \quad (3.3.11)$$

Da $V \subset \mathbb{R}^n$ offen ist, existiert eine Zahl $\varepsilon > 0$ mit $x + t\xi \in V \cap (\mathbb{R}^d \times \{0\})$ für alle $t \in (-\varepsilon, \varepsilon)$. Nun definieren wir $\gamma : (-\varepsilon, \varepsilon) \rightarrow U \cap M$ durch

$$\gamma(t) := \phi^{-1}(x + t\xi) \quad \text{für } -\varepsilon < t < \varepsilon. \quad (3.3.12)$$

Dies ist eine C^ℓ -Abbildung mit

$$\gamma(0) = \phi^{-1}(x) = p, \quad \dot{\gamma}(0) = d\phi^{-1}(x)\xi = d\phi(p)^{-1}\xi = v.$$

Hier folgt die vorletzte Gleichung aus Lemma 3.1.2 und die letzte aus (3.3.11). Also ist $v \in T_p M$ und damit ist (3.3.10) bewiesen.

In der Situation von Teil (ii) zeigen wir die Inklusion

$$T_p M \subset \ker df(p). \quad (3.3.13)$$

Sei also $v \in T_p M$ und sei $\gamma : (-\varepsilon, \varepsilon) \rightarrow M$ eine C^1 -Kurve, die (3.3.6) erfüllt. Nun sei $f : U \rightarrow \mathbb{R}^{n-d}$ wie in Teil (ii). Da $\gamma(0) = p \in U$ und U eine offene Teilmenge von X ist, existiert eine Zahl $0 < \delta \leq \varepsilon$, so dass $\gamma(t) \in U \cap M$ ist für $-\delta < t < \delta$. Daraus folgt $f(\gamma(t)) = 0$ für $-\delta < t < \delta$ und daher gilt

$$df(p)v = \left. \frac{d}{dt} \right|_{t=0} f(\gamma(t)) = 0.$$

Also ist $v \in \ker df(p)$ und damit ist (3.3.13) bewiesen.

Nach (3.3.10) und (3.3.13) gilt $d\phi(p)^{-1}(\mathbb{R}^d \times \{0\}) \subset T_p M \subset \ker df(p)$. Nach Voraussetzung sind $d\phi(p)^{-1}(\mathbb{R}^d \times \{0\})$ und $\ker df(p)$ beides d -dimensionale lineare Unterräume von X und müssen daher übereinstimmen. Damit sind Teil (i), Teil (ii) und Teil (iii) bewiesen.

Wir beweisen Teil (iv). Sei also $v \in T_p M$. Dann ist $d\phi(p)v \in \mathbb{R}^d \times \{0\}$ nach Teil (i), und daher definieren die Gleichungen (3.3.11) und (3.3.12) eine C^ℓ -Kurve $\gamma : (-\varepsilon, \varepsilon) \rightarrow M$, die (3.3.6) erfüllt. Daraus folgt, dass die Formel

$$\beta(t) := \gamma \left(\frac{\varepsilon t}{\sqrt{\varepsilon^2 + t^2}} \right) \quad \text{für } t \in \mathbb{R} \quad (3.3.14)$$

eine C^ℓ -Kurve $\beta : \mathbb{R} \rightarrow M$ definiert mit $\beta(0) = \gamma(0) = p$, $\dot{\beta}(0) = \dot{\gamma}(0) = v$. Damit sind Teil (iv) und Satz 3.3.22 bewiesen. \square

Beispiel 3.3.23. Im Fall $d = 0$ (Beispiel 3.3.2) ist der Tangentialraum einer diskreten Menge $M \subset X$ an jeder Stelle $p \in M$ der Unterraum $T_p M = \{0\}$.

Beispiel 3.3.24. Im Fall $d = n$ (Beispiel 3.3.3) ist der Tangentialraum einer offenen Menge $M \subset X$ an jeder Stelle $p \in M$ der Unterraum $T_p M = X$.

Beispiel 3.3.25. Der Tangentialraum der orthogonalen Gruppe $O(n)$ in Beispiel 3.3.18 an der Stelle $A \in O(n)$ ist

$$T_A O(n) = \left\{ \widehat{A} \in \mathbb{R}^{n \times n} \mid A^T \widehat{A} + \widehat{A}^T A = 0 \right\}$$

nach Teil (ii) von Satz 3.3.22. Für $A = \mathbb{1}$ erhält man den Raum

$$\mathfrak{o}(n) := T_{\mathbb{1}} O(n) = \left\{ \widehat{A} \in \mathbb{R}^{n \times n} \mid \widehat{A} + \widehat{A}^T = 0 \right\}$$

der schiefsymmetrischen Matrizen. Dieser Tangentialraum hat die Eigenschaft, dass der Kommutator $[\widehat{A}, \widehat{B}] := \widehat{A}\widehat{B} - \widehat{B}\widehat{A}$ zweier Matrizen $\widehat{A}, \widehat{B} \in \mathfrak{o}(n)$ wieder eine Element von $\mathfrak{o}(n)$ ist. Darüber hinaus gilt $\exp(t\widehat{A}) \in O(n)$ für alle $\widehat{A} \in \mathfrak{o}(n)$ und alle $t \in \mathbb{R}$ (Übung).

Beispiel 3.3.26. Der Tangentialraum der $(n-1)$ -dimensionalen Untermannigfaltigkeit $Q = \{x \in \mathbb{R}^n \mid x^T A x\}$ in Beispiel 3.3.19 an der Stelle $x \in Q$ ist

$$T_x Q = \{\xi \in \mathbb{R}^n \mid \langle Ax, \xi \rangle = 0\} = (Ax)^\perp.$$

nach Teil (ii) von Satz 3.3.22. Insbesondere gilt $T_x S^{n-1} = x^\perp$ für $x \in S^{n-1}$.

Beispiel 3.3.27. Der Tangentialraum der Untermannigfaltigkeit $S \subset \mathbb{R}^6$ in Beispiel 3.3.20 im Punkte $p = (x^2, y^2, z^2, yz, zx, xy) \in S$ mit $x^2 + y^2 + z^2 = 1$ ist der 2-dimensionale Unterraum

$$T_p S := \left\{ \left(\begin{array}{c} 2x\xi \\ 2y\eta \\ 2z\zeta \\ y\zeta + z\eta \\ z\xi + x\zeta \\ x\eta + y\xi \end{array} \right) \mid \xi, \eta, \zeta \in \mathbb{R}, x\xi + y\eta + z\zeta = 0 \right\}. \quad (3.3.15)$$

Hier zeigt man die Inklusion “ \subset ” durch das Differenzieren von Kurven in S . Die Gleichheit folgt dann aus der Tatsache, dass nach Teil (iii) von Satz 3.3.22 beide Seiten der Gleichung (3.3.15) 2-dimensionale lineare Unterräume des \mathbb{R}^6 sind. Mit dieser Methode kann man in vielen Beispielen die Tangentialräume auf einfache Weise bestimmen.

3.4 Extrema mit Nebenbedingungen

In diesem Abschnitt untersuchen wir die Frage, wie man Elemente einer Menge U finden kann, die eine Funktion $f : U \rightarrow \mathbb{R}$ unter Nebenbedingungen der Form $h_i(x) = 0$ minimieren. Hier ist $U \subset \mathbb{R}^n$ eine offene Menge, und wir nehmen an, dass $f : U \rightarrow \mathbb{R}$ und $h = (h_1, \dots, h_m) : U \rightarrow \mathbb{R}^m$ differenzierbar sind. Weiterhin erinnern wir an die Bezeichnung

$$\nabla f(x) := \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix} \in \mathbb{R}^n$$

für den Gradienten von f an der Stelle $x \in U$. Mit $\|x\|$ bezeichnen wir hier stets die Euklidische Norm eines Vektors $x \in \mathbb{R}^n$.

Satz 3.4.1 (Lagrange). *Sei $U \subset \mathbb{R}^n$ eine offene Menge, seien $f : U \rightarrow \mathbb{R}$ und $h = (h_1, \dots, h_m) : U \rightarrow \mathbb{R}^m$ C^1 -Abbildungen, sei $0 \in \mathbb{R}^m$ ein regulärer Wert von h , und sei*

$$M := h^{-1}(0) = \{x \in U \mid h(x) = 0\}. \quad (3.4.1)$$

Erfüllt ein Element $x_0 \in M$ die Bedingung

$$f(x_0) \leq f(x) \quad \text{für alle } x \in M, \quad (3.4.2)$$

so existieren reelle Zahlen $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ mit

$$\nabla f(x_0) = \sum_{i=1}^m \lambda_i \nabla h_i(x_0). \quad (3.4.3)$$

*Diese Zahlen λ_i werden **Lagrange-Multiplikatoren** genannt.*

Beweis. Wir zeigen zunächst, dass

$$\nabla f(x_0) \in (T_{x_0}M)^\perp \quad (3.4.4)$$

ist. Sei also $\xi \in T_{x_0}M$. Dann existiert nach Teil (iv) von Satz 3.3.22 eine C^1 -Kurve $\gamma : \mathbb{R} \rightarrow M \subset \mathbb{R}^n$ mit $\gamma(0) = x_0$ und $\dot{\gamma}(0) = \xi$. Daraus folgt die Ungleichung $f(\gamma(0)) \leq f(\gamma(t))$ für alle $t \in \mathbb{R}$ nach (3.4.2), und daher gilt

$$\langle \nabla f(x_0), \xi \rangle = df(x_0)\xi = df(\gamma(0))\dot{\gamma}(0) = \left. \frac{d}{dt} \right|_{t=0} f(\gamma(t)) = 0.$$

Hier folgt der letzte Schritt aus Satz 2.7.2. Damit ist (3.4.4) bewiesen.

Nun sei $d := n - m$. Nach Korollar 3.3.12 ist $M \subset \mathbb{R}^n$ eine d -dimensionale C^1 -Untermannigfaltigkeit von \mathbb{R}^n und nach Teil (ii) von Satz 3.3.22 ist ihr Tangentialraum an der Stelle $x_0 \in M$ der d -dimensionale lineare Unterraum

$$\begin{aligned} T_{x_0}M &= \ker dh(x_0) \\ &= \{\xi \in \mathbb{R}^n \mid dh(x_0)\xi = 0\} \\ &= \{\xi \in \mathbb{R}^n \mid \langle \nabla h_i(x_0), \xi \rangle = 0 \text{ für } i = 1, \dots, m\} \\ &= \bigcap_{i=1}^m \nabla h_i(x_0)^\perp. \end{aligned} \tag{3.4.5}$$

Dass dieser Unterraum die Dimension $d = n - m$ hat folgt aus den genannten Sätzen, oder auch aus der Tatsache, dass die Vektoren $\nabla h_1(x_0), \dots, \nabla h_m(x_0)$ linear unabhängig sind (Lemma 3.3.10). Da jeder lineare Unterraum $V \subset \mathbb{R}^n$ die Gleichung $V = V^{\perp\perp}$ erfüllt, erhalten wir nach (3.4.5) die Gleichung

$$\begin{aligned} (T_{x_0}M)^\perp &= \left(\bigcap_{i=1}^m \nabla h_i(x_0)^\perp \right)^\perp \\ &= \left\{ \sum_{i=1}^m \lambda_i \nabla h_i(x_0) \mid \lambda_1, \dots, \lambda_n \in \mathbb{R} \right\}^{\perp\perp} \\ &= \left\{ \sum_{i=1}^m \lambda_i \nabla h_i(x_0) \mid \lambda_1, \dots, \lambda_n \in \mathbb{R} \right\}. \end{aligned} \tag{3.4.6}$$

Da $\nabla f(x_0)$ nach (3.4.4) orthogonal zu $T_{x_0}M$ ist, folgt daraus die Behauptung von Satz 3.4.1. \square

Beispiel 3.4.2. Dieses Beispiel zeigt, dass auf die Annahme, dass 0 ein regulärer Wert von h ist, in Satz 3.4.1 nicht verzichtet werden kann. Seien $f, h : \mathbb{R} \rightarrow \mathbb{R}$ die Funktionen

$$f(x) := x, \quad h(x) := x^2$$

für $x \in \mathbb{R}$. Dann ist 0 kein regulärer Wert von h . Alle anderen Voraussetzungen von Satz 3.4.1 sind mit $x_0 = 0 \in M = \{0\}$ erfüllt, jedoch ist $\nabla f(x_0) = 1$ in diesem Beispiel kein reelles Vielfaches von $\nabla h(x_0) = 0$.

Beispiel 3.4.3 (Geometrisches und Arithmetisches Mittel).

Sei $U \subset \mathbb{R}^n$ der *positive Quadrant*

$$U := \{x = (x_1, \dots, x_n) \in \mathbb{R}^n \mid x_i > 0 \text{ für } i = 1, \dots, n\}$$

und seien $f, h : \mathbb{R}^n \rightarrow \mathbb{R}$ die Funktionen

$$f(x) := x_1 x_2 \cdots x_n, \quad h(x) := x_1 + x_2 + \cdots + x_n - 1 \quad (3.4.7)$$

für $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$. Dann ist $M := U \cap h^{-1}(0)$ die Menge

$$M = \left\{ x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n \mid \begin{array}{l} x_i > 0 \text{ für } i = 1, 2, \dots, n, \\ x_1 + x_2 + \cdots + x_n = 1 \end{array} \right\}. \quad (3.4.8)$$

Die Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ist stetig und der Abschluss \overline{M} von M ist kompakt. Daher existiert ein Vektor $a = (a_1, a_2, \dots, a_n) \in \overline{M}$ mit

$$f(x) \leq f(a) \quad \text{für alle } x \in M. \quad (3.4.9)$$

Da f auf M positiv ist und auf $\overline{M} \setminus M$ verschwindet, ist $a \in M$. Nach Satz 3.4.1 existiert daher eine Zahl $\lambda \in \mathbb{R}$ mit $\nabla f(a) = \lambda \nabla h(a)$. Dann gilt

$$\frac{a_1 a_2 \cdots a_n}{a_i} = \frac{\partial f}{\partial x_i}(a) = \lambda \frac{\partial h}{\partial x_i}(a) = \lambda \quad (3.4.10)$$

für $i = 1, \dots, n$ und daraus folgt $a_1 = a_2 = \cdots = a_n = 1/n$. Es gibt also nur einen Vektor $a \in M$, der (3.4.10) erfüllt, und dieser muss also auch (3.4.9) erfüllen. Damit haben wir die Ungleichung

$$x_1 x_2 \cdots x_n \leq \left(\frac{1}{n}\right)^n \quad (3.4.11)$$

für alle $x = (x_1, x_2, \dots, x_n) \in M$ bewiesen. Seien nun x_1, x_2, \dots, x_n positive reelle Zahlen. Dann ist $\frac{1}{x_1 + x_2 + \cdots + x_n}(x_1, x_2, \dots, x_n) \in M$ und nach (3.4.11) gilt

$$\frac{x_1 x_2 \cdots x_n}{(x_1 + x_2 + \cdots + x_n)^n} \leq \left(\frac{1}{n}\right)^n.$$

Daraus wiederum folgt $x_1 x_2 \cdots x_n \leq \left(\frac{x_1 + x_2 + \cdots + x_n}{n}\right)^n$, und daher gilt

$$\left(\prod_{i=1}^n x_i\right)^{1/n} \leq \frac{1}{n} \sum_{i=1}^n x_i \quad (3.4.12)$$

für jedes n -Tupel von positiven reellen Zahlen x_i . Damit haben wir gezeigt, dass das *geometrische Mittel* kleiner oder gleich dem *arithmetischen Mittel* ist (wie es natürlich bereits aus der Analysis I bekannt ist.)

Beispiel 3.4.4 (Eigenvektoren). Sei $A = A^T \in \mathbb{R}^{n \times n}$ eine symmetrische Matrix und seien $f, h : \mathbb{R}^n \rightarrow \mathbb{R}$ die glatten Funktionen

$$f(x) := \frac{1}{2} x^T A x, \quad h(x) := \frac{1}{2} (\|x\|^2 - 1)$$

für $x \in \mathbb{R}^n$. Dann ist 0 ein regulärer Wert von h und $M := h^{-1}(0) = S^{n-1}$ ist die Einheitskugel. Dann existiert ein Vektor $v_1 \in S^{n-1}$ mit

$$f(v_1) \leq f(x) \quad \text{für alle } x \in S^{n-1}.$$

Nach Satz 3.4.1 existiert nun eine reelle Zahl $\lambda_1 \in \mathbb{R}$ mit $\nabla f(v_1) = \lambda_1 \nabla h(v_1)$. Mit $\nabla f(x) = Ax$ und $\nabla h(x) = x$ ist diese Gleichung äquivalent zu

$$Av_1 = \lambda_1 v_1. \quad (3.4.13)$$

Wir haben damit die Existenz eines reellen Eigenwertes von A bewiesen.

Nun betrachten wir ein neues Extremalproblem mit der gleichen Funktion f und zwei Nebenbedingungen, die durch $h, h_1 : \mathbb{R}^n \rightarrow \mathbb{R}$ mit

$$h_1(x) = \langle v_1, x \rangle \quad \text{für } x \in \mathbb{R}^n$$

gegeben sind. Dann ist $(0, 0)$ ein regulärer Wert von $(h, h_1) : \mathbb{R}^n \rightarrow \mathbb{R}^2$ und

$$\begin{aligned} M_1 &:= h^{-1}(0) \cap h_1^{-1}(0) \\ &= S^{n-1} \cap v_1^\perp \\ &= \{x \in S^{n-1} \mid \langle v_1, x \rangle = 0\} \end{aligned}$$

Da M_1 kompakt ist, existiert ein Vektor $v_2 \in M_1$ mit

$$f(v_2) \leq f(x) \quad \text{für alle } x \in M_1.$$

Nach Satz 3.4.1 existieren nun zwei Lagrange-Multiplikatoren $\lambda_2, \mu \in \mathbb{R}$ mit

$$\nabla f(v_2) = \lambda_2 \nabla h(v_2) + \mu \nabla h_1(v_2).$$

Mit $\nabla f(x) = Ax$, $\nabla h(x) = x$, $\nabla h_1(x) = v_1$ folgt daraus $Av_2 = \lambda_2 v_2 + \mu v_1$. Da $\langle \lambda_2 v_2 - Av_2, v_1 \rangle = 0$ ist folgt daraus wiederum $\|\lambda_2 v_2 - Av_2\|^2 + \mu^2 = 0$. Damit haben wir gezeigt, dass v_2 und λ_2 die Bedingungen

$$Av_2 = \lambda_2 v_2, \quad \langle v_1, v_2 \rangle = 0 \quad (3.4.14)$$

erfüllen. Setzt man dieses Argument mit vollständiger Induktion fort, so erhält man eine Orthonormalbasis v_1, v_2, \dots, v_n des \mathbb{R}^n , die aus Eigenvektoren der Matrix A besteht. Die zugehörigen Eigenwerte $\lambda_1, \lambda_2, \dots, \lambda_n$ sind reell und erscheinen bei dieser Methode in der Reihenfolge $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$.

Beispiel 3.4.5 (Abstand zu einer Untermannigfaltigkeit). Sei $M \subset \mathbb{R}^n$ eine glatte kompakte d -dimensionale Untermannigfaltigkeit mit $0 \leq d < n$. Sei $y_0 \in \mathbb{R}^n$ gegeben. Dann existiert ein Element $x_0 \in M$ mit

$$\|y_0 - x_0\| \leq \|y_0 - x\| \quad \text{für alle } x \in M \quad (3.4.15)$$

und dieses Element x_0 erfüllt die Bedingung

$$y_0 - x_0 \perp T_{x_0}M. \quad (3.4.16)$$

Die Existenz eines solchen Elementes x_0 folgt aus der Tatsache, dass die durch $f(x) := \frac{1}{2} \|y_0 - x\|^2$ für $x \in \mathbb{R}^n$ definiert stetige Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ auf der kompakten Teilmenge $M \subset \mathbb{R}^n$ ihr Minimum annimmt. Die Bedingung (3.4.16) hat dann die Form $\nabla f(x_0) \in (T_{x_0}M)^\perp$ und dies folgt aus der Gleichung (3.4.4) im Beweis von Satz 3.4.1. In konkreten Beispielen kann der gesuchte Vektor $x_0 \in M$, der die Bedingung (3.4.15) erfüllt, oft dadurch gefunden werden, dass man alle Vektoren untersucht, die die notwendige Bedingung (3.4.16) erfüllen.

Übung 3.4.6 (Tubenumgebungen). Sei $M \subset \mathbb{R}^n$ wie in Beispiel 3.4.5. Für $\varepsilon > 0$ definieren wir die Mengen

$$\begin{aligned} W_\varepsilon &:= \{(x, \eta) \in \mathbb{R}^n \times \mathbb{R}^n \mid x \in M, \eta \perp T_x M, \|\eta\| < \varepsilon\}, \\ U_\varepsilon &:= \left\{ y \in \mathbb{R}^n \mid \inf_{x \in M} \|y - x\| < \varepsilon \right\}, \end{aligned} \quad (3.4.17)$$

und die Abbildung $\phi_\varepsilon : W_\varepsilon \rightarrow \mathbb{R}^n$ durch $\phi_\varepsilon(x, \eta) := x + \eta$ für $(x, \eta) \in W_\varepsilon$. Dann gilt folgendes.

- (i) W_ε ist eine glatte n -dimensionale Untermannigfaltigkeit von $\mathbb{R}^n \times \mathbb{R}^n$.
- (ii) U_ε ist eine offene Teilmenge von \mathbb{R}^n und es gilt $\phi_\varepsilon(W_\varepsilon) = U_\varepsilon$.
- (iii) Ist $\varepsilon > 0$ hinreichend klein, so ist $\phi_\varepsilon : W_\varepsilon \rightarrow U_\varepsilon$ bijektiv und besitzt eine glatte Umkehrabbildung. Genauer, wenn $\varepsilon > 0$ hinreichend klein ist, dann existiert für jedes Element $y \in U_\varepsilon$ genau ein Element $\pi(y) \in M$ mit

$$\|y - \pi(y)\| = \inf_{x \in M} \|y - x\|,$$

die Abbildung $\pi : U_\varepsilon \rightarrow M$ ist glatt, und es gilt $\phi_\varepsilon^{-1}(y) = (\pi(y), y - \pi(y))$ für alle $y \in U_\varepsilon$.

- (iv) Für $\varepsilon > 0$ hinreichend klein wird die offene Menge $U_\varepsilon \subset \mathbb{R}^n$ in (3.4.17) eine **Tubenumgebung** von M genannt und die Aussage (iii) ist der *Satz über Tubenumgebungen*. Ein Beweis findet sich in [3].

Kapitel 4

Vektorfelder und Flüsse

Dieses Kapitel dient der Einführung in die gewöhnlichen Differentialgleichungen. Der erste Abschnitt wiederholt einige Inhalte aus Analysis I, welche die Existenz und Eindeutigkeit von Lösungen eines Anfangswertproblems betreffen. Der Abschnitt 4.2 führt den Fluss eines Vektorfeldes ein und zeigt, dass eine Lösung, die nur auf einem endlichen Zeitintervall existiert, jede kompakte Teilmenge des Definitionsgebietes des Vektorfeldes verlassen muss. Abschnitt 4.3 zeigt die stetige Abhängigkeit der Lösungen vom Anfangswert und Abschnitt 4.4 beweist die Differenzierbarkeit des Flusses.

4.1 Existenz und Eindeutigkeit

Sei $U \subset \mathbb{R}^n$ eine offene Menge und sei $f : U \rightarrow \mathbb{R}^n$ eine lokal Lipschitz-stetige Abbildung. Diese Abbildung ist als **Vektorfeld** zu verstehen, das jedem Punkt $x \in U$ einen Geschwindigkeitsvektor $f(x) \in \mathbb{R}^n$ zuordnet. Es soll darum gehen, die Lösungen der Differentialgleichung

$$\dot{x}(t) = f(x(t)), \quad x(0) = x_0 \quad (4.1.1)$$

zu verstehen. Hier ist $x_0 \in U$ der **Anfangswert** und die Gleichungen in (4.1.1) werden **Anfangswertproblem** genannt. Eine **Lösung** des Anfangswertproblems (4.1.1) ist eine stetig differenzierbare Abbildung $x : I \rightarrow U$, definiert auf einem offenen Intervall $I \subset \mathbb{R}$ mit $0 \in I$, die zum Zeitpunkt $t = 0$ den Wert x_0 annimmt und dessen Ableitung durch $\dot{x}(t) = f(x(t))$ für $t \in I$ gegeben ist. Zunächst sei daran erinnert, was aus der Analysis I Vorlesung über lokal Lipschitz-stetige Abbildungen und die Lösungen von (4.1.1) bekannt ist.

Im folgenden bezeichnen wir stets mit $\|x\| := \sqrt{x_1^2 + \dots + x_n^2}$ die Euklidische Norm eines Vektors $x = (x_1, \dots, x_n) \in \mathbb{R}^n$.

Definition 4.1.1. Sei $U \subset \mathbb{R}^n$ offen. Eine Abbildung $f : U \rightarrow \mathbb{R}^n$ heißt **lokal Lipschitz-stetig** wenn für jedes $x_0 \in U$ zwei Zahlen $\varepsilon > 0$ und $c > 0$ existieren, so dass

$$B_\varepsilon(x_0) = \{x \in \mathbb{R}^n \mid \|x - x_0\| < \varepsilon\} \subset U$$

ist und alle $x, y \in B_\varepsilon(x_0)$ die Ungleichung

$$\|f(x) - f(y)\| \leq c \|x - y\|$$

erfüllen.

Lemma 4.1.2. Sei $U \subset \mathbb{R}^n$ eine offenen Teilmenge, sei $f : U \rightarrow \mathbb{R}^n$ eine lokal Lipschitz-stetige Abbildung, und sei $K \subset U$ eine kompakte Teilmenge. Dann existiert eine Konstante $c > 0$ so dass alle $x, y \in K$ die Ungleichung $\|f(x) - f(y)\| \leq c \|x - y\|$ erfüllen.

Beweis. Wir nehmen an, die Behauptung sei falsch. Dann existieren (nach dem abzählbaren Auswahlaxiom) zwei Folgen $x_i, y_i \in K$, so dass für alle $i \in \mathbb{N}$ die Ungleichung

$$\|f(x_i) - f(y_i)\| > i \|x_i - y_i\| \tag{4.1.2}$$

erfüllt ist. Da K kompakt ist, können wir durch Übergang zu einer Teilfolge ohne Einschränkung der Allgemeinheit annehmen, dass die Folge $(x_i)_{i \in \mathbb{N}}$ konvergiert. Wir bezeichnen den Grenzwert mit

$$x_0 := \lim_{i \rightarrow \infty} x_i \in K.$$

Da die Funktion $K \rightarrow \mathbb{R} : x \mapsto \|f(x)\|$ stetig ist und K kompakt ist, existiert nach einem Satz aus Analysis I ein Element $\xi \in K$ so dass

$$C := \|f(\xi)\| \geq \|f(x)\| \quad \text{für alle } x \in K. \tag{4.1.3}$$

Aus (4.1.2) und (4.1.3) folgt die Ungleichung

$$\|y_i - x_i\| \leq \frac{\|f(y_i) - f(x_i)\|}{i} \leq \frac{\|f(y_i)\| + \|f(x_i)\|}{i} \leq \frac{2C}{i}$$

für alle $i \in \mathbb{N}$. Also gilt $\lim_{i \rightarrow \infty} (y_i - x_i) = 0$ und daher konvergiert die Folge $(y_i)_{i \in \mathbb{N}}$ ebenfalls gegen x_0 .

Da f lokal Lipschitz-stetig ist, existieren reelle Zahlen $\varepsilon > 0$ und $c > 0$, so dass für alle $x, y \in \mathbb{R}^n$ folgendes gilt:

$$\begin{aligned} \|x - x_0\| < \varepsilon, \\ \|y - x_0\| < \varepsilon \end{aligned} \quad \implies \quad \begin{aligned} x, y \in U \text{ und} \\ \|f(x) - f(y)\| \leq c \|x - y\|. \end{aligned} \quad (4.1.4)$$

Da die Folgen $(x_i)_{i \in \mathbb{N}}$ und $(y_i)_{i \in \mathbb{N}}$ beide gegen x_0 konvergieren, existiert eine natürliche Zahl $i_0 \in \mathbb{N}$, so dass jedes $i \in \mathbb{N}$ mit $i \geq i_0$ die Ungleichungen

$$i \geq c, \quad \|x_i - x_0\| < \varepsilon, \quad \|y_i - x_0\| < \varepsilon$$

erfüllt. Für $i \in \mathbb{N}$ mit $i \geq i_0$ folgt daraus nach (4.1.2) die Ungleichung

$$\|f(x_i) - f(y_i)\| > i \|x_i - y_i\| \geq c \|x_i - y_i\|.$$

Diese steht im Widerspruch zu (4.1.4) und beweist damit Lemma 4.1.2. \square

Lemma 4.1.3. *Sei $U \subset \mathbb{R}^n$ eine offene Menge und sei $K \subset U$ eine kompakte Menge. Dann gilt folgendes.*

(i) *Für jedes $\varepsilon > 0$ ist die Menge*

$$K_\varepsilon := \{y \in \mathbb{R}^n \mid \text{es gibt ein } x \in K \text{ so dass } \|x - y\| \leq \varepsilon\}$$

kompakt.

(ii) *Es existiert ein $\varepsilon > 0$ mit $K_\varepsilon \subset U$.*

Beweis. Es gilt $\sup_{y \in K_\varepsilon} \|y\| = \sup_{x \in K} \|x\| + \varepsilon < \infty$ und daher ist K_ε beschränkt. Wir zeigen nun, dass K_ε auch abgeschlossen ist. Sei $y_i \in K_\varepsilon$ eine Folge, die gegen $y \in \mathbb{R}^n$ konvergiert. Dann gibt es eine Folge $x_i \in K$ so dass

$$\|x_i - y_i\| \leq \varepsilon \quad \text{für alle } i \in \mathbb{N}.$$

Da K kompakt ist können wir durch Übergang zu einer Teilfolge ohne Einschränkung der Allgemeinheit voraussetzen, dass die Folge x_i gegen ein Element $x \in K$ konvergiert. Daraus folgt

$$\|x - y\| = \lim_{i \rightarrow \infty} \|x_i - y_i\| \leq \varepsilon.$$

Also ist $y \in K_\varepsilon$. Damit ist gezeigt, dass K_ε abgeschlossen und beschränkt ist. Nach dem Satz von Heine–Borel ist K_ε damit kompakt.

Wir beweisen Teil (ii) indirekt und nehmen an, dass $K_\varepsilon \not\subset U$ ist für jedes $\varepsilon > 0$. Dann existiert nach dem abzählbaren Auswahlaxiom eine Folge

$$y_i \in \mathbb{R}^n \setminus U$$

so dass $y_i \in K_{1/i}$ ist für jedes $i \in \mathbb{N}$. Also existiert, wiederum nach dem abzählbaren Auswahlaxiom, eine Folge $x_i \in K$ so dass

$$\|x_i - y_i\| \leq \frac{1}{i}$$

ist für jedes $i \in \mathbb{N}$. Da K kompakt ist, können wir durch Übergang zu einer Teilfolge ohne Einschränkung der Allgemeinheit annehmen, dass die Folge x_i gegen ein Element $x \in K$ konvergiert. Daraus folgt

$$x = \lim_{i \rightarrow \infty} x_i = \lim_{i \rightarrow \infty} y_i \notin U,$$

da $\mathbb{R}^n \setminus U$ abgeschlossen ist. Also ist $K \not\subset U$ im Gegensatz zu unserer Voraussetzung. Damit ist Lemma 4.1.3 bewiesen. \square

Der Schrankensatz (Satz 2.4.1) besagt, dass jede stetig differenzierbare Abbildung $f : U \rightarrow \mathbb{R}^n$ lokal Lipschitz-stetig ist.

Der folgende Satz aus der Analysis I Vorlesung garantiert die Existenz und Eindeutigkeit von Lösungen des Anfangswertproblems (4.1.1) unter der Voraussetzung, dass das Vektorfeld $f : U \rightarrow \mathbb{R}^n$ lokal Lipschitz-stetig ist.

Satz 4.1.4 (Existenz und Eindeutigkeit). *Sei $U \subset \mathbb{R}^n$ eine offene Teilmenge, sei $f : U \rightarrow \mathbb{R}^n$ eine lokal Lipschitz-stetige Abbildung, und sei $K \subset U$ eine kompakte Teilmenge. Dann gilt folgendes.*

(i) *Es existiert ein offenes Intervall $I \subset \mathbb{R}$ mit $0 \in I$, so dass das Anfangswertproblem (4.1.1) für jedes $x_0 \in K$ eine Lösung $x : I \rightarrow U$ besitzt.*

(ii) *Ist $I \subset \mathbb{R}$ ein offenes Intervall mit $0 \in I$ und sind $x, y : I \rightarrow U$ zwei Lösungen von (4.1.1) mit $x_0 \in U$, so gilt $x(t) = y(t)$ für alle $t \in I$.*

Beweis. Siehe Seite 137. \square

Seien f und x_0 wie in Satz 4.1.4, seien $I, J \subset \mathbb{R}$ zwei offenen Intervalle mit $0 \in I \cap J$, und seien $x : I \rightarrow U$ und $y : J \rightarrow U$ zwei Lösungen von (4.1.1). Dann gilt $x(t) = y(t)$ für alle $t \in I \cap J$ nach Teil (ii) von Satz 4.1.4. Daraus folgt, dass die Gleichung (4.1.1) auch eine Lösung auf dem Intervall $I \cup J$ besitzt, die durch $x(t)$ für $t \in I$ und durch $y(t)$ für $t \in J$ gegeben ist. Dies führt zu dem Konzept des *maximalen Existenzintervalls* in Definition 4.2.1.

Lemma 4.1.5. *Seien $f : U \rightarrow \mathbb{R}^n$ und $x_0 \in U$ wie in Satz 4.1.4, sei $I \subset \mathbb{R}$ ein offenes Intervall mit $0 \in I$, und sei $x : I \rightarrow U$ eine stetige Abbildung. Dann sind folgende Aussagen äquivalent.*

(i) *Die Abbildung $x : I \rightarrow U$ ist eine Lösung von (4.1.1)*

(ii) *Die Abbildung $x : I \rightarrow U$ erfüllt die Integralgleichung*

$$x(t) = x_0 + \int_{t_0}^t f(x(s)) ds \quad \text{für alle } t \in I. \quad (4.1.5)$$

Beweis. Dies folgt direkt aus dem Hauptsatz der Infinitesimalrechnung, beziehungsweise Teil (v) und (vi) von Lemma 2.4.4. \square

Beweis von Satz 4.1.4. Da $K \subset U$ eine kompakte Teilmenge ist, existiert nach Lemma 4.1.3 eine Zahl $\varepsilon > 0$, so dass

$$K_\varepsilon = \{x \in \mathbb{R}^n \mid \text{es gibt ein } x_0 \in K \text{ mit } \|x - x_0\| \leq \varepsilon\} \subset U. \quad (4.1.6)$$

Da K_ε nach Lemma 4.1.3 eine kompakte Teilmenge von U ist, existiert nach Lemma 4.1.2 eine Konstante $c > 0$, so dass

$$\|f(x) - f(y)\| \leq c \|x - y\| \quad \text{für alle } x, y \in K_\varepsilon. \quad (4.1.7)$$

Außerdem ist die Funktion $K \rightarrow \mathbb{R} : x \mapsto \|f(x)\|$ stetig und daher nach einem Satz aus der Analysis I Vorlesung beschränkt. Definiere

$$M := \sup_{x \in K} \|f(x)\| + c\varepsilon \quad (4.1.8)$$

und wähle $\delta > 0$ so klein, dass

$$\delta M < \varepsilon, \quad \delta c < 1. \quad (4.1.9)$$

Sei $x_0 \in K$ gegeben. Wir beweisen in fünf Schritten die Existenz und Eindeutigkeit einer Lösung $x : [-\delta, \delta] \rightarrow U$ des Anfangswertproblems (4.1.1).

Schritt 1. *Sei $x \in \mathbb{R}^n$ mit $\|x - x_0\| \leq \varepsilon$. Dann gilt*

$$x \in U, \quad \|f(x)\| \leq M.$$

Nach (4.1.6) gilt $x \in U$, und nach (4.1.7) und (4.1.8) gilt

$$\begin{aligned} \|f(x)\| &\leq \|f(x_0)\| + \|f(x) - f(x_0)\| \\ &\leq \|f(x_0)\| + c \|x - x_0\| \\ &\leq \|f(x_0)\| + c\varepsilon \\ &= M. \end{aligned}$$

Damit ist Schritt 1 bewiesen.

Schritt 2. Sei $x : [-\delta, \delta] \rightarrow U$ eine Lösung von (4.1.1). Dann gilt

$$\|x(t) - x_0\| < \varepsilon$$

für alle $t \in [-\delta, \delta]$.

Wir beweisen Schritt 2 durch ein Widerspruchsargument und nehmen an, dass eine reelle Zahl $t \in [0, \delta]$ existiert mit $\|x(t) - x_0\| \geq \varepsilon$. Sei

$$\tau := \inf \{t \in [0, \delta] \mid \|x(t) - x_0\| \geq \varepsilon\}.$$

Dann gilt $0 < \tau \leq \delta$ und

$$\|x(t) - x_0\| < \varepsilon \quad \text{für alle } t \in [0, \tau), \quad (4.1.10)$$

$$\|x(\tau) - x_0\| = \varepsilon. \quad (4.1.11)$$

Daraus folgt

$$\|x(\tau) - x_0\| = \left\| \int_0^\tau f(x(t)) dt \right\| \leq \int_0^\tau \|f(x(t))\| dt \leq \tau M < \varepsilon.$$

Hier folgt der erste Schritt aus Lemma 4.1.5, der zweite Schritt aus Teil (iii) von Lemma 2.4.4, der dritte Schritt aus (4.1.10) und Schritt 1, und der letzte Schritt folgt aus (4.1.9). Diese Ungleichung steht im Widerspruch zu (4.1.11) und damit ist gezeigt, dass $\|x(t) - x_0\| < \varepsilon$ ist für alle $t \in [0, \delta]$. Das gleiche Argument kann für $t \in [-\delta, 0]$ verwendet werden und damit ist Schritt 2 bewiesen.

Schritt 3. Sei

$$\mathcal{X} := \left\{ x : [-\delta, \delta] \rightarrow \mathbb{R}^n \left| \begin{array}{l} x \text{ ist stetig und} \\ \|x(t) - x_0\| \leq \varepsilon \\ \text{für alle } t \in [-\delta, \delta] \end{array} \right. \right\} \quad (4.1.12)$$

und definiere $d : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ durch

$$d(x, y) := \|x - y\|_\infty := \sup_{|t| \leq \delta} \|x(t) - y(t)\| \quad (4.1.13)$$

für $x, y \in \mathcal{X}$. Dann ist (\mathcal{X}, d) ein nichtleerer vollständiger metrischer Raum.

Der Raum \mathcal{X} ist nichtleer, da die konstante Funktion $x(t) = x_0$ ein Element von \mathcal{X} ist. Daß die Abbildung $d : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ eine Abstandsfunktion ist, folgt direkt aus den Definitionen. Daß der metrische Raum (\mathcal{X}, d) vollständig ist, folgt daraus, dass jede Cauchy-Folge reeller Zahlen konvergiert und dass der Grenzwert einer gleichmäßig konvergenten Folge stetiger Funktionen, nach einem Satz aus der Analysis I Vorlesung, selbst wieder eine stetige Funktion ist. Damit ist Schritt 3 bewiesen.

Schritt 4. Sei \mathcal{X} durch Gleichung (4.1.12) in Schritt 3 gegeben und sei $x \in \mathcal{X}$. Definiere die Funktion $\mathcal{F}(x) : [-\delta, \delta] \rightarrow \mathbb{R}^n$ durch

$$(\mathcal{F}(x))(t) := x_0 + \int_0^t f(x(s)) ds \quad (4.1.14)$$

für $-\delta \leq t \leq \delta$. Dann gilt $\mathcal{F}(x) \in \mathcal{X}$.

Für $-\delta \leq s \leq \delta$ gilt $\|x(s) - x_0\| \leq \varepsilon$ und daher $x(s) \in U$ und $\|f(x(s))\| \leq M$ nach Schritt 1. Daraus folgt, dass die Funktion $\mathcal{F}(x)$ in (4.1.14) wohldefiniert ist und für $0 \leq t \leq \delta$ der Ungleichung

$$\begin{aligned} \|(\mathcal{F}(x))(t) - x_0\| &= \left\| \int_0^t f(x(s)) ds \right\| \\ &\leq \int_0^t \|f(x(s))\| ds \\ &\leq tM \\ &< \varepsilon \end{aligned}$$

genügt. Hier folgt der zweite Schritt aus Teil (iii) von Lemma 2.4.4, und der letzte Schritt aus (4.1.9). Ebenso gilt $\|(\mathcal{F}(x))(t) - x_0\| < \varepsilon$ für $-\delta \leq t \leq 0$. Außerdem ist $\mathcal{F}(x)$ stetig und daher gilt $\mathcal{F}(x) \in \mathcal{X}$. Damit ist Schritt 4 bewiesen.

Schritt 5. Sei (\mathcal{X}, d) der nichtleere vollständige metrische Raum in Schritt 3 und sei $\mathcal{F} : \mathcal{X} \rightarrow \mathcal{X}$ die Abbildung in Schritt 4. Dann gilt

$$\|\mathcal{F}(x) - \mathcal{F}(y)\|_\infty \leq \delta c \|x - y\|_\infty$$

für alle $x, y \in \mathcal{X}$ und daher ist \mathcal{F} eine Kontraktion.

Seien $x, y \in \mathcal{X}$. Dann gilt

$$\begin{aligned} \|(\mathcal{F}(x))(t) - (\mathcal{F}(y))(t)\| &= \left\| \int_0^t (f(x(s)) - f(y(s))) ds \right\| \\ &\leq \int_0^t \|f(x(s)) - f(y(s))\| ds \\ &\leq \int_0^t c \|x(s) - y(s)\| ds \\ &\leq \delta c \|x - y\|_\infty \end{aligned}$$

für $0 \leq t \leq \delta$. Genauso argumentiert man für $-\delta \leq t \leq 0$ und daraus folgt die Ungleichung $\|\mathcal{F}(x) - \mathcal{F}(y)\|_\infty \leq \delta c \|x - y\|_\infty$. Da $\delta c < 1$ ist, folgt daraus, dass $\mathcal{F} : \mathcal{X} \rightarrow \mathcal{X}$ eine Kontraktion ist. Damit ist Schritt 5 bewiesen.

Nach Schritt 5 und dem Banachschen Fixpunktsatz E.0.1 besitzt \mathcal{F} einen eindeutigen Fixpunkt

$$x = \mathcal{F}(x) \in \mathcal{X}.$$

Dieser Fixpunkt ist nach (4.1.6) eine stetige Funktion

$$x : [-\delta, \delta] \rightarrow U,$$

welche nach Definition von \mathcal{F} die Gleichung

$$x(t) = (\mathcal{F}(x))(t) = x_0 + \int_0^t f(x(s)) ds$$

für alle $t \in [-\delta, \delta]$ erfüllt. Dies ist die Gleichung (4.1.5). Also folgt aus Lemma 4.1.5, dass x eine Lösung von (4.1.1) ist. Umgekehrt folgt aus Schritt 2 und Lemma 4.1.5, dass jede Lösung $x : [-\delta, \delta] \rightarrow U$ von (4.1.1) ein Fixpunkt von \mathcal{F} ist. Damit ist die Existenz und Eindeutigkeit einer Lösung von (4.1.1) auf dem Intervall $[-\delta, \delta]$ für jeden Anfangswert $x_0 \in K$ bewiesen.

Nun sei $x_0 \in U$ gegeben, sei $I \subset \mathbb{R}$ ein offenes Intervall mit $0 \in I$, und seien $x, y : I \rightarrow U$ zwei Lösungen von (4.1.1). Wir nehmen an, es gäbe ein Element $t \in I$ mit

$$x(t) \neq y(t).$$

Dann gilt $t \neq 0$. Im Fall $t > 0$ definieren wir

$$t_0 := \sup \{t \in I \mid t > 0 \text{ und } x(s) = y(s) \text{ für alle } s \in [0, t]\}.$$

Dann gilt $t_0 \in I$ und $t_0 > 0$ und $x(s) = y(s)$ für $0 \leq s \leq t_0$. Nun folgt aber aus der bereits bewiesenen Eindeutigkeit mit dem Anfangswert $x(t_0)$, angewendet auf die Lösungen $x(t - t_0)$ und $y(t - t_0)$, dass eine Zahl $\delta > 0$ existiert mit $t_0 - \delta, t_0 + \delta \in I$ und $x(t_0 + t) = y(t_0 + t)$ für $-\delta \leq t \leq \delta$. Dies steht im Widerspruch zur Definition von t_0 . Ebenso erhält man einen Widerspruch im Fall $t < 0$ und damit ist Satz 4.1.4 bewiesen. \square

In Satz 4.1.4 gilt die Existenzaussage in Teil (i) auch dann, wenn das Vektorfeld $f : U \rightarrow \mathbb{R}^n$ nur stetig, nicht aber Lipschitz-stetig ist. Allerdings erfordert diese allgemeinere Aussage einen anderen Beweis. Zudem kann bei der Eindeutigkeitsaussage in Teil (ii) auf die Lipschitz-Stetigkeit nicht verzichtet werden, wie das folgende Beispiel zeigt.

Beispiel 4.1.6. Sei $n = 1$ und sei $f : \mathbb{R} \rightarrow \mathbb{R}$ das stetige Vektorfeld

$$f(x) := \sqrt{|x|} \quad \text{für } x \in \mathbb{R}. \quad (4.1.15)$$

Dann sind die Funktionen

$$x(t) := 0, \quad y(t) := \begin{cases} t^2/4, & \text{falls } t \geq 0, \\ 0, & \text{falls } t \leq 0, \end{cases}$$

zwei verschiedene Lösungen des Anfangswertproblems (4.1.1) mit $x_0 = 0$.

Beispiel 4.1.7. Für $n = 1$ betrachten wir das Anfangswertproblem

$$\dot{x} := -x^2 \quad x(0) = x_0. \quad (4.1.16)$$

Ist $x_0 = 0$ so ist $x(t) = 0$ die eindeutige Lösung von (4.1.16) und diese existiert auf ganz \mathbb{R} . Ist $x_0 \neq 0$ so ist jede Lösung $x : I \rightarrow \mathbb{R}$ von (4.1.16) überall ungleich Null, hat daher eine negative Ableitung, und ist daher nach dem Mittelwertsatz strikt monoton fallend. Daraus ergibt sich, dass die Funktion $I \rightarrow \mathbb{R} : t \mapsto x(t)$ eine differenzierbare Umkehrfunktion $x \mapsto t(x)$ besitzt. Da $\dot{x}(t) = -x(t)^2$ ist für alle t , hat die Umkehrfunktion die Ableitung $dt/dx = -1/x^2$. Nach dem Hauptsatz der Infinitesimalrechnung gilt

$$t = t(x) - t(x_0) = - \int_{x_0}^x \frac{d\xi}{\xi^2} = \frac{1}{x} - \frac{1}{x_0}.$$

Daher ist die Lösung von (4.1.16) im Fall $x_0 > 0$ durch die Formel

$$x(t) = \frac{x_0}{1 + tx_0}, \quad -\frac{1}{x_0} < t < \infty \quad (4.1.17)$$

gegeben. Diese Lösung läßt sich auf kein grösseres Intervall fortsetzen.

Beispiel 4.1.8. Sei $A \in \mathbb{R}^{n \times n}$ und $x_0 \in \mathbb{R}^n$. Dann ist die eindeutige Lösung des Anfangswertproblems

$$\dot{x} = Ax, \quad x(0) = x_0 \quad (4.1.18)$$

durch die Exponentialmatrix gegeben, das heißt

$$x(t) = e^{tA} x_0 = \sum_{k=0}^{\infty} \frac{t^k A^k x_0}{k!} \quad \text{für } t \in \mathbb{R}. \quad (4.1.19)$$

4.2 Der Fluss eines Vektorfeldes

Von nun an nehmen wir an, dass eine offenen Teilmenge $U \subset \mathbb{R}^n$ und ein lokal Lipschitz-stetiges Vektorfeld $f : U \rightarrow \mathbb{R}^n$ gegeben sind.

Definition 4.2.1. Sei $x_0 \in U$. Das **maximale Existenzintervall** von x_0 (für Gleichung (4.1.1)) ist die Menge

$$I(x_0) := \bigcup \left\{ I \subset \mathbb{R} \mid \begin{array}{l} I \text{ ist ein offenes Intervall mit } 0 \in I \\ \text{und das Anfangswertproblem (4.1.1)} \\ \text{besitzt eine Lösung } x : I \rightarrow U \end{array} \right\}. \quad (4.2.1)$$

Nach Satz 4.1.4 ist $I(x_0) \subset \mathbb{R}$ ein offenes Intervall mit $0 \in I(x_0)$, das Anfangswertproblem (4.1.1) besitzt eine Lösung $x : I(x_0) \rightarrow U$, diese Lösung ist eindeutig, und es existiert keine Lösung von (4.1.1) auf einem echt grösseren Intervall.

Definition 4.2.2. Sei

$$\Omega := \{(t, x_0) \in \mathbb{R} \times U \mid t \in I(x_0)\}. \quad (4.2.2)$$

Der **Fluss** von f ist die Abbildung

$$\phi : \Omega \rightarrow U,$$

die jedem Paar $(t, x_0) \in \Omega$ die eindeutige Lösung von (4.1.1) zum Zeitpunkt t zuordnet, das heißt $\phi(t, x_0) := x(t)$, wobei $x : I(x_0) \rightarrow U$ die eindeutige Lösung von (4.1.1) ist. Mit anderen Worten, $\phi : \Omega \rightarrow U$ ist die eindeutige Abbildung die nach t partiell differenzierbar ist und die Gleichungen

$$\partial_t \phi(t, x) = f(\phi(t, x)), \quad \phi(0, x_0) = x_0 \quad (4.2.3)$$

für alle $(t, x) \in \Omega$ und alle $x_0 \in U$ erfüllt.

Beispiel 4.2.3. Sei $A \in \mathbb{R}^{n \times n}$ und sei $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ das lineare Vektorfeld

$$f(x) := Ax$$

für $x \in \mathbb{R}^n$. Dann ist $\Omega = \mathbb{R} \times \mathbb{R}^n$ und der Fluss $\phi : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ von f ist nach Beispiel 4.1.8 die Abbildung

$$\phi(t, x) = e^{tA}x$$

für $t \in \mathbb{R}$ und $x \in \mathbb{R}^n$.

Beispiel 4.2.4. Sei $n = 1$, $U = \mathbb{R}$, und sei $f : \mathbb{R} \rightarrow \mathbb{R}$ das Vektorfeld

$$f(x) := -x^2 \quad \text{für } x \in \mathbb{R}.$$

(Siehe Beispiel 4.1.7.) Für $x_0 \in \mathbb{R}$ ist

$$x(t) := \frac{x_0}{1 + tx_0}$$

die eindeutige Lösung von (4.1.1) auf dem Intervall

$$I(x_0) = \begin{cases} (-x_0^{-1}, \infty), & \text{für } x_0 > 0, \\ \mathbb{R}, & \text{für } x_0 = 0, \\ (-\infty, -x_0^{-1}), & \text{für } x_0 < 0. \end{cases}$$

Der Fluss von f ist daher durch die Gleichungen

$$\Omega = \{(t, x) \in \mathbb{R}^2 \mid tx > -1\}, \quad \phi(t, x) = \frac{x}{1 + tx}$$

gegeben.

Lemma 4.2.5. Sei $x_0 \in U$ und $t_0 \in I(x_0)$. Dann gilt

$$I(\phi(t_0, x_0)) = I(x_0) - t_0 \tag{4.2.4}$$

und

$$\phi(s, \phi(t_0, x_0)) = \phi(t_0 + s, x_0) \tag{4.2.5}$$

für alle $s \in I(\phi(t_0, x_0))$.

Beweis. Sei

$$y_0 := \phi(t_0, x_0)$$

und

$$J := I(x_0) - t_0 = \{t - t_0 \mid t \in I(x_0)\} = \{s \in \mathbb{R} \mid t_0 + s \in I(x_0)\}.$$

Dann ist $J \subset \mathbb{R}$ ein offenes Intervall und $0 \in J$. Definiere $y : J \rightarrow \mathbb{R}$ durch

$$y(s) := \phi(t_0 + s, x_0) \quad \text{für } s \in J.$$

Dann ist y stetig differenzierbar und es gilt $\dot{y}(s) = f(y(s))$ für alle $s \in J$ und $y(0) = y_0$. Also ist $J \subset I(y_0)$ und $y(s) = \phi(s, y_0)$ für alle $s \in J$. Damit haben wir die Inklusion

$$I(x_0) - t_0 \subset I(\phi(t_0, x_0))$$

und die Gleichung (4.2.5) für alle $s \in I(x_0) - t_0$ bewiesen.

Es bleibt zu zeigen, dass die umgekehrte Inklusion gilt, das heißt

$$I(\phi(t_0, x_0)) \subset I(x_0) - t_0.$$

Dazu bemerken wir, dass nach dem bisher gezeigten folgendes gilt:

$$-t_0 \in I(x_0) - t_0 \subset I(y_0), \quad \phi(-t_0, y_0) = \phi(0, x_0) = x_0.$$

Daher ergibt sich aus dem ersten Teil des Beweises, mit (t_0, x_0) durch $(-t_0, y_0)$ ersetzt, die Inklusion $I(y_0) - (-t_0) \subset I(\phi(-t_0, y_0))$. Das heißt

$$I(\phi(t_0, x_0)) + t_0 \subset I(x_0)$$

und daher $I(\phi(t_0, x_0)) \subset I(x_0) - t_0$. Damit ist Lemma 4.2.5 bewiesen. \square

Satz 4.2.6. Sei $x_0 \in U$ ein Anfangswert mit

$$I(x_0) \cap [0, \infty) = [0, b), \quad 0 < b < \infty.$$

Dann existiert für jede kompakte Menge $K \subset U$ eine Zahl $0 \leq t_K < b$ mit

$$t_K < t < b \quad \implies \quad \phi(t, x_0) \notin K. \quad (4.2.6)$$

Mit anderen Worten, wenn die Lösung des Anfangswertproblems (4.1.1) nur auf einem endlichen Zeitintervall existiert, muss sie jede kompakte Teilmenge von U verlassen.

Beweis. Nach Satz 4.1.4 existiert eine Zahl $\delta > 0$ mit

$$(-\delta, \delta) \subset I(x) \quad \text{für alle } x \in K.$$

Hier können wir $\delta > 0$ so klein wählen, dass $\delta \leq b$ ist. Dann erfüllt die Zahl

$$t_K := b - \delta$$

die Bedingung (4.2.6). Ist nämlich t eine reelle Zahl mit

$$b - \delta < t < b$$

so gilt nach Lemma 4.2.5

$$I(\phi(t, x_0)) = I(x_0) - t.$$

Daraus folgt

$$I(\phi(t, x_0)) \cap [0, \infty) = [0, b - t).$$

Da $b - t < \delta$ ist, gilt $(-\delta, \delta) \not\subset I(\phi(t, x_0))$ und daher kann $\phi(t, x_0)$ kein Element von K sein. Damit ist Satz 4.2.6 bewiesen. \square

4.3 Stetigkeit des Flusses

Unser Ziel ist es, zu zeigen, dass der Fluss eines lokal Lipschitz-stetigen Vektorfeldes selbst ebenfalls eine lokal Lipschitz-stetige Abbildung ist.

Satz 4.3.1. *Sei $U \subset \mathbb{R}^n$ eine offene Menge, sei $f : U \rightarrow \mathbb{R}^n$ ein lokal Lipschitz-stetiges Vektorfeld, und sei $\phi : \Omega \rightarrow U$ der Fluss von f . Dann ist ϕ lokal Lipschitz-stetig.*

Beweis. Siehe Seite 146. □

Der Beweis beruht auf dem **Lemma von Gronwall**.

Lemma 4.3.2 (Gronwall). *Sei $I \subset \mathbb{R}$ ein Intervall mit $0 \in I$, seien $A, B \geq 0$, und sei $g : I \rightarrow [0, \infty)$ eine stetige Funktion, die die Ungleichung*

$$g(t) \leq A + B \left| \int_0^t g(s) ds \right| \quad (4.3.1)$$

für alle $t \in I$ erfüllt. Dann gilt

$$g(t) \leq Ae^{B|t|} \quad (4.3.2)$$

für alle $t \in I$.

Beweis. Definiere die Funktion $G : I \cap [0, \infty) \rightarrow [0, \infty)$ durch

$$G(t) := A + B \int_0^t g(s) ds \quad \text{für } t \in I \text{ mit } t \geq 0.$$

Sie erfüllt die Ungleichung $g(t) \leq G(t)$ und daher

$$\dot{G}(t) = Bg(t) \leq BG(t)$$

für alle $t \in I$ mit $t \geq 0$. Daraus folgt

$$\frac{d}{dt}(e^{-Bt}G(t)) = e^{-Bt}(\dot{G}(t) - BG(t)) \leq 0$$

und daher

$$e^{-Bt}G(t) \leq G(0) = A$$

für alle $t \in I$ mit $t \geq 0$. Also gilt

$$g(t) \leq G(t) \leq Ae^{Bt}$$

für alle $t \in I$ mit $t \geq 0$. Im Fall $t \leq 0$ ergibt sich die gewünschte Ungleichung indem man die Funktion g durch die Funktion $t \mapsto g(-t)$ ersetzt. Damit ist Lemma 4.3.2 bewiesen. □

Beweis von Satz 4.3.1. Der Beweis hat drei Schritte. Der erste Schritt verwendet die Notation von Lemma 4.1.3.

Schritt 1. Sei $K \subset U$ kompakt und sei $\varepsilon > 0$ so dass $K_\varepsilon \subset U$ ist. Dann existiert ein $\delta > 0$ mit $[-\delta, \delta] \times K \subset \Omega$ und $\phi([-\delta, \delta] \times K) \subset K_\varepsilon$.

Nach Voraussetzung ist $U_\varepsilon := \bigcup_{x \in K} B_\varepsilon(x)$ eine offene Teilmenge von U . Nach Teil (i) von Satz 4.1.4, mit U_ε statt U , existiert dann eine Zahl $\delta > 0$ mit $[-\delta, \delta] \subset I(x_0)$ und $\phi(t, x_0) \in U_\varepsilon$ für alle $x_0 \in K$ und alle $t \in [-\delta, \delta]$. Damit ist Schritt 1 bewiesen.

Schritt 2. Seien K, ε, δ wie in Schritt 1. Dann ist ϕ Lipschitz-stetig auf der Menge $[-\delta, \delta] \times K \subset \Omega$.

Da f lokal Lipschitz-stetig und $K_\varepsilon \subset U$ nach Lemma 4.1.3 kompakt ist, ist $f|_{K_\varepsilon}$ nach Lemma 4.1.2 Lipschitz-stetig, und das heißt

$$c := \sup_{\substack{x, y \in K_\varepsilon \\ x \neq y}} \frac{\|f(x) - f(y)\|}{\|x - y\|} < \infty, \quad M := \sup_{x \in K_\varepsilon} \|f(x)\| < \infty. \quad (4.3.3)$$

Seien $x_0, y_0 \in K$ gegeben, und seien $x, y : [-\delta, \delta] \rightarrow U$ die zugehörigen Lösungen von (4.1.1), das heißt $x(t) := \phi(t, x_0)$ und $y(t) := \phi(t, y_0)$ für $t \in [-\delta, \delta]$. Dann gilt $x(t), y(t) \in K_\varepsilon$ für $t \in [-\delta, \delta]$ nach Schritt 1 und daher

$$\|x(s) - x(t)\| = \left\| \int_s^t f(x(r)) dr \right\| \leq \int_s^t \|f(x(r))\| dr \leq M|s - t|$$

für alle $s, t \in [-\delta, \delta]$ mit $s < t$. Außerdem gilt für alle $t \in [-\delta, \delta]$

$$\begin{aligned} \|x(t) - y(t)\| &= \left\| x_0 - y_0 + \int_0^t (f(x(s)) - f(y(s))) ds \right\| \\ &\leq \|x_0 - y_0\| + \left\| \int_0^t (f(x(s)) - f(y(s))) ds \right\| \\ &\leq \|x_0 - y_0\| + \left| \int_0^t \|f(x(s)) - f(y(s))\| ds \right| \\ &\leq \|x_0 - y_0\| + c \left| \int_0^t \|x(s) - y(s)\| ds \right|. \end{aligned}$$

Daraus folgt $\|x(t) - y(t)\| \leq e^{c|t|} \|x_0 - y_0\|$ nach Lemma 4.3.2 und daher

$$\begin{aligned} \|\phi(s, x_0) - \phi(t, y_0)\| &\leq \|x(s) - x(t)\| + \|x(t) - y(t)\| \\ &\leq M|s - t| + e^{c\delta} \|x_0 - y_0\| \end{aligned}$$

für alle $x_0, y_0 \in K$ und alle $s, t \in [-\delta, \delta]$. Damit ist Schritt 2 bewiesen.

Schritt 3. Die Menge $\Omega \subset \mathbb{R} \times U$ ist offen und die Abbildung $\phi : \Omega \rightarrow U$ ist lokal Lipschitz-stetig.

Sei $(t_0, x_0) \in \Omega$ mit $t_0 \geq 0$ und definiere

$$\Gamma := \{\phi(t, x_0) \mid 0 \leq t \leq t_0\}.$$

Dies ist eine kompakte Teilmenge von U . Daher existiert nach Lemma 4.1.3 eine Konstante $\varepsilon > 0$ so dass

$$K := \{y \in \mathbb{R}^n \mid \text{es gibt ein } x \in \Gamma \text{ so dass } \|x - y\| \leq \varepsilon\} \subset U.$$

Außerdem ist die Menge K kompakt, nach Lemma 4.1.3, und die Menge

$$V := \{y \in \mathbb{R}^n \mid \text{es gibt ein } x \in \Gamma \text{ so dass } \|x - y\| < \varepsilon\} = \bigcup_{x \in \Gamma} B_\varepsilon(x)$$

ist offen. Nach Schritt 1 und Schritt 2 existiert eine Konstante $\delta > 0$ so dass $[-\delta, \delta] \times K \subset \Omega$ und ϕ auf $[-\delta, \delta] \times K$ Lipschitz-stetig ist. Wähle $N \in \mathbb{N}$ so gross, dass

$$\tau := \frac{t_0}{N} < \delta.$$

Dann ist $[0, \tau] \times V \subset \Omega$ und wir definieren die Abbildung $\phi_\tau : V \rightarrow U$ durch

$$\phi_\tau(y) := \phi(\tau, y) \quad \text{für } y \in V.$$

Diese Abbildung ist stetig nach Schritt 2. Wir definieren nun eine endliche Folge offener Mengen $V_0 \supset V_1 \supset \dots \supset V_N$ durch

$$\begin{aligned} V_0 &:= V, \\ V_1 &:= \phi_\tau^{-1}(V_0) = \{y \in V \mid \phi_\tau(y) \in V\} \\ V_2 &:= \phi_\tau^{-1}(V_1) = \{y \in V \mid \phi_\tau(y) \in V, \phi_\tau(\phi_\tau(y))\} \\ &\vdots \\ V_N &:= \phi_\tau^{-1}(V_{N-1}) = \{y \in V \mid \phi_\tau^k(y) \in V \text{ für } k = 1, \dots, N\}. \end{aligned}$$

Hier bezeichnen wir mit $\phi_\tau^k := \phi_\tau \circ \dots \circ \phi_\tau$ die k -fache Komposition der Abbildung $\phi_\tau : V \rightarrow U$ mit sich selbst. Ihr Definitionsgebiet ist genau die Menge V_{k-1} . Die Mengen $V_k \subset U$ sind offen, da $\phi_\tau : V \rightarrow U$ stetig ist. Außerdem folgt aus Lemma 4.2.5, dass $\phi_\tau^k(x_0) = \phi(k\tau, x_0) \in \Gamma \subset V$ ist für $k = 1, \dots, N$ und daher gilt $x_0 \in V_k$ für $k = 0, 1, \dots, N$.

Es folgt ebenfalls aus Lemma 4.2.5, dass für alle $x \in V_N$ folgendes gilt:

$$t_0 = N\tau \in I(x), \quad \phi(t_0, x) \in V.$$

Daraus folgt $[-\delta, \delta] \subset I(\phi(t_0, x))$ für alle $x \in V_N$ und daher, nach Lemma 4.2.5,

$$[t_0 - \delta, t_0 + \delta] \times V_N \subset \Omega.$$

Da ϕ auf $[-\delta, \delta] \times K$ Lipschitz-stetig ist, existiert eine Konstante $c > 0$, so dass alle $s, s' \in [-\delta, \delta]$ und alle $y, y' \in K$ die Ungleichung

$$\|\phi(s, y) - \phi(s', y')\| \leq c|s - s'| + c\|y - y'\| \quad (4.3.4)$$

erfüllen. Für $x, x' \in V_N$ und $t, t' \in (t_0 - \delta, t_0 + \delta)$ folgt daraus die Ungleichung

$$\begin{aligned} \|\phi(t, x) - \phi(t', x')\| &= \|\phi(t - t_0, \phi_\tau^N(x)) - \phi(t' - t_0, \phi_\tau^N(x'))\| \\ &\leq c|t - t'| + c\|\phi_\tau^N(x) - \phi_\tau^N(x')\| \\ &\leq c|t - t'| + c^{N+1}\|x - x'\|. \end{aligned}$$

Hier folgt die letzte Ungleichung durch N -malige Anwendung der Ungleichung (4.3.4). Damit haben wir gezeigt, dass es für jedes Element $(t_0, x_0) \in \Omega$ mit $t_0 \geq 0$ eine offene Menge $W = (t_0 - \delta, t_0 + \delta) \times V_N \subset \mathbb{R} \times \mathbb{R}^n$ gibt, so dass $(t_0, x_0) \in W \subset \Omega$ ist und die Restriktion $\phi|_W : W \rightarrow U$ Lipschitz-stetig ist. Für $t_0 < 0$ folgt dieselbe Aussage, indem man f durch $-f$ und $\phi(t, x)$ durch $\phi(-t, x)$ ersetzt. Damit sind Schritt 3 und Satz 4.3.1 bewiesen. \square

Bemerkung 4.3.3. Für jedes $t \in \mathbb{R}$ ist die Menge

$$U_t := \{x_0 \in U \mid t \in I(x_0)\}$$

offen nach Satz 4.3.1. Außerdem folgt aus Lemma 4.2.5, dass für jedes $x \in U_t$ gilt, dass $-t \in I(x) - t = I(\phi(t, x))$ und daher $\phi(t, x) \in U_{-t}$ ist. Es ist oft nützlich, für den Fluss von f die Notation

$$\phi_t : U_t \rightarrow U_{-t}, \quad \phi_t(x) := \phi(t, x),$$

zu verwenden. Nach Satz 4.3.1 ist die Abbildung $\phi_t : U_t \rightarrow U_{-t}$ ein Homöomorphismus mit der Umkehrabbildung

$$\phi_t^{-1} = \phi_{-t}.$$

Außerdem gilt $U_0 = U$, $\phi_0 = \text{id} : U \rightarrow U$, und

$$\phi_{s+t} = \phi_s \circ \phi_t \quad \text{auf } U_{s+t} \cap \phi_t^{-1}(U_s)$$

für alle $s, t \in \mathbb{R}$ nach Lemma 4.2.5. Diese Notation vereinfacht sich erheblich, wenn $I(x) = \mathbb{R}$ ist für alle $x \in U$. Dann ist $U_t = U$ für alle $t \in \mathbb{R}$.

4.4 Differenzierbarkeit des Flusses

Satz 4.4.1. *Sei $U \subset \mathbb{R}^n$ offen und sei $f : U \rightarrow \mathbb{R}^n$ ein C^ℓ Vektorfeld mit $\ell \in \mathbb{N} \cup \{\infty\}$. Dann ist der Fluss $\phi : \Omega \rightarrow U$ von f eine C^ℓ -Abbildung.*

Beweis. Siehe Seite 151. □

Bevor wir diesen Satz beweisen, formulieren wir zunächst ein Lemma, welches die stetige Differenzierbarkeit des Flusses voraussetzt. Dieses Lemma liefert eine Formel für die Ableitung von ϕ und diese Formel wird dann verwendet, um Satz 4.4.1 zu beweisen.

Lemma 4.4.2. *Sei $U \subset \mathbb{R}^n$ offen, sei $f : U \rightarrow \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld, und sei der Fluss $\phi : \Omega \rightarrow U$ von f eine stetig differenzierbare Abbildung. Seien $x_0 \in U$ und $\xi_0 \in \mathbb{R}^n$ gegeben. Sei $I := I(x_0)$ und definiere die Abbildungen $x : I \rightarrow U$ und $\xi : I \rightarrow \mathbb{R}^n$ durch*

$$x(t) := \phi_t(x_0), \quad \xi(t) := d\phi_t(x_0)\xi_0 \quad (4.4.1)$$

für $t \in I$. Dann sind x und ξ stetig differenzierbar und es gilt

$$\begin{aligned} \dot{x}(t) &= f(x(t)), & x(0) &= x_0, \\ \dot{\xi}(t) &= df(x(t))\xi(t), & \xi(0) &= \xi_0, \end{aligned} \quad (4.4.2)$$

für alle $t \in I$.

Beweis. Die partielle Ableitung $\partial\phi/\partial t = f \circ \phi : \Omega \rightarrow \mathbb{R}^n$ existiert nach Definition des Flusses und ist stetig differenzierbar nach Voraussetzung. Außerdem gilt nach der Kettenregel

$$\frac{\partial}{\partial x_i} \frac{\partial \phi}{\partial t}(t, x) = df(\phi(t, x)) \frac{\partial \phi}{\partial x_i}(t, x) \quad \text{für alle } (t, x) \in \Omega.$$

Nach dem Satz von Schwarz (über die Vertauschbarkeit der zweiten Ableitungen) folgt daraus, dass die Abbildung $\partial\phi/\partial x_i : \Omega \rightarrow \mathbb{R}^n$ nach t partiell differenzierbar ist und die Gleichung

$$\frac{\partial}{\partial t} \frac{\partial \phi}{\partial x_i}(t, x) = \frac{\partial}{\partial x_i} \frac{\partial \phi}{\partial t}(t, x) = df(\phi(t, x)) \frac{\partial \phi}{\partial x_i}(t, x)$$

für alle $(t, x) \in \Omega$ erfüllt. Ist also $\xi_0 = (\xi_{01}, \dots, \xi_{0n}) \in \mathbb{R}^n$, dann ist die Funktion $\xi(t) = d\phi_t(x_0)\xi_0 = \sum_{i=1}^n \frac{\partial \phi}{\partial x_i}(t, x_0)\xi_{0i}$ stetig differenzierbar mit Ableitung

$$\dot{\xi}(t) = \sum_{i=1}^n \frac{\partial}{\partial t} \frac{\partial \phi}{\partial x_i}(t, x_0)\xi_{0i} = \sum_{i=1}^n df(\phi(t, x_0)) \frac{\partial \phi}{\partial x_i}(t, x_0)\xi_{0i} = df(x(t))\xi(t).$$

Damit ist Lemma 4.4.2 bewiesen. □

Lemma 4.4.3. *Sei $I \subset \mathbb{R}$ ein offenes Intervall mit $0 \in I$ und $A : I \rightarrow \mathbb{R}^{n \times n}$ eine stetige Abbildung. Dann besitzt das Anfangswertproblem*

$$\dot{\xi}(t) = A(t)\xi(t), \quad \xi(0) = \xi_0, \quad (4.4.3)$$

für jedes $\xi_0 \in \mathbb{R}^n$ eine eindeutige, stetig differenzierbare Lösung $\xi : I \rightarrow \mathbb{R}^n$.

Beweis. Sei $T \in I$ mit $T > 0$ und definiere

$$c := 2 \sup_{0 \leq t \leq T} \|A(t)\|, \quad \|A(t)\| := \sup_{\xi \in \mathbb{R}^n \setminus \{0\}} \frac{\|A(t)\xi\|}{\|\xi\|}.$$

Sei $\mathcal{X} := C([0, T], \mathbb{R}^n)$ der Raum aller stetigen Abbildungen $\xi : [0, T] \rightarrow \mathbb{R}^n$. Dies ist ein Banachraum mit der Normfunktion

$$\|\xi\|_c := \sup_{0 \leq t \leq T} e^{-ct} \|\xi(t)\| \quad \text{für } \xi \in \mathcal{X}.$$

Definiere die Abbildung $\mathcal{F} : \mathcal{X} \rightarrow \mathcal{X}$ durch

$$\mathcal{F}(\xi)(t) := \xi_0 + \int_0^t A(s)\xi(s) ds \quad \text{für } 0 \leq t \leq T.$$

Wir beweisen die Ungleichung

$$\|\mathcal{F}(\xi) - \mathcal{F}(\eta)\|_c \leq \frac{1}{2} \|\xi - \eta\|_c \quad \text{für alle } \xi, \eta \in \mathcal{X}. \quad (4.4.4)$$

Seien also $\xi, \eta \in \mathcal{X}$ gegeben und sei $t \in [0, T]$. Dann gilt

$$\begin{aligned} \|\mathcal{F}(\xi)(t) - \mathcal{F}(\eta)(t)\| &= \left\| \int_0^t A(s)(\xi(s) - \eta(s)) ds \right\| \\ &\leq \int_0^t \|A(s)\| \|\xi(s) - \eta(s)\| ds \\ &\leq \frac{1}{2} \int_0^t c \|\xi(s) - \eta(s)\| ds \\ &= \frac{1}{2} \int_0^t c e^{cs} e^{-cs} \|\xi(s) - \eta(s)\| ds \\ &\leq \frac{1}{2} \int_0^t c e^{cs} ds \|\xi - \eta\|_c \\ &\leq \frac{1}{2} e^{ct} \|\xi - \eta\|_c \end{aligned}$$

für alle $t \in [0, T]$. Daraus folgt die Ungleichung (4.4.4) indem man beide Seiten mit e^{-ct} multipliziert und das Supremum über alle $t \in [0, T]$ bildet.

Aus der Ungleichung (4.4.4) folgt nach dem Banachschen Fixpunktsatz (Satz E.0.1), dass die Abbildung

$$\mathcal{F} : \mathcal{X} \rightarrow \mathcal{X}$$

einen eindeutigen Fixpunkt

$$\xi = \mathcal{F}(\xi)$$

besitzt. Dieser Fixpunkt ist die eindeutige Lösung $\xi : [0, T] \rightarrow \mathbb{R}^n$ des Anfangswertproblems (4.4.3) auf dem Intervall $[0, T]$. Da eine solche eindeutige Lösung auf jedem kompakten Teilintervall $[0, T] \subset I$ existiert, existiert sie auch auf $I \cap [0, \infty)$. Die Existenz einer eindeutigen Lösung auf ganz I folgt nun dadurch, dass man die Funktion $A : I \rightarrow \mathbb{R}^{n \times n}$ durch die Funktion $-I \rightarrow \mathbb{R}^{n \times n} : t \mapsto -A(-t)$ ersetzt. Damit ist Lemma 4.4.3 bewiesen. \square

Lemma 4.4.2 zeigt, dass die Linearisierung $\xi(t) := d\phi_t(x_0)\xi_0$ des Flusses, falls sie existiert, die linearisierte Differentialgleichung (4.4.2) löst. Lemma 4.4.3 zeigt, dass die zweite Gleichung in (4.4.2) immer auf dem gesamten Existenzintervall $I = I(x_0)$ eine Lösung $\xi : I \rightarrow \mathbb{R}^n$ besitzt. Die Aufgabe im Beweis von Satz 4.4.1 besteht nun darin, zu zeigen, dass diese Lösung auch tatsächlich durch die Ableitung von ϕ_t mittels (4.4.1) gegeben ist.

Beweis von Satz 4.4.1. Wir beweisen die Behauptung zunächst für $\ell = 1$. Sei also $f : U \rightarrow \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld. Zu zeigen ist, dass der Fluss $\phi : \Omega \rightarrow U$ von f ebenfalls stetig differenzierbar ist. Sei $x_0 \in U$ und $I := I(x_0) \subset \mathbb{R}$. Definiere die Abbildungen

$$x : I \rightarrow U, \quad A : I \rightarrow \mathbb{R}^{n \times n}$$

durch

$$x(t) := \phi_t(x_0), \quad A(t) := df(x(t)) \quad (4.4.5)$$

für $t \in I$. Nach Lemma 4.4.3 existiert eine eindeutige, stetig differenzierbare Abbildung $\Phi : I \rightarrow \mathbb{R}^{n \times n}$ so dass

$$\dot{\Phi}(t) := A(t)\Phi(t), \quad \Phi(0) = \mathbb{1}, \quad (4.4.6)$$

für alle $t \in I$. Wir beweisen zunächst, dass, für jedes $t \in I$, die Abbildung $\phi_t : U_t \rightarrow U_{-t}$ an der Stelle $x_0 \in U_t$ differenzierbar ist mit der Jacobi-Matrix

$$d\phi_t(x_0) = \Phi(t) \quad \text{für } t \in I. \quad (4.4.7)$$

Sei $T \in I$ mit $T > 0$. Dann ist folgendes zu zeigen.

Behauptung 1: Für jedes $\varepsilon > 0$ existiert ein $\delta > 0$ so dass für alle $t \in [0, T]$ und alle $\xi_0 \in \mathbb{R}^n$ mit $\|\xi_0\| < \delta$ gilt, dass $x_0 + \xi_0 \in U_t$ ist und

$$\|\phi_t(x_0 + \xi_0) - \phi_t(x_0) - \Phi(t)\xi_0\| \leq \varepsilon \|\xi_0\|. \quad (4.4.8)$$

Beweis von Behauptung 1. Sei eine reelle Zahl $\varepsilon > 0$ gegeben. Wir wählen die Konstante $\delta > 0$ in den folgenden vier Schritten.

(a) Nach Satz 4.3.1 existiert eine Konstante $r > 0$ mit

$$[0, T] \times \{x \in \mathbb{R}^n \mid \|x - x_0\| \leq r\} \subset \Omega.$$

(b) Nach Satz 4.3.1 existiert eine Konstante $c > 0$, so dass

$$\|\phi_t(x_0 + \xi_0) - \phi_t(x_0)\| \leq c \|\xi_0\|, \quad \|df(x(t))\| \leq c$$

für alle $t \in [0, T]$ und alle $\xi_0 \in \mathbb{R}^n$ mit $\|\xi_0\| \leq r$.

(c) Es existiert eine Konstante $\rho > 0$, so dass für alle $t \in [0, T]$ und alle $h \in \mathbb{R}^n$ mit $\|h\| < \rho$ gilt, dass $x(t) + h \in U$ ist und

$$\|df(x(t) + h) - df(x(t))\| \leq e^{-cT} \varepsilon.$$

Hier verwenden wir die Tatsache dass eine stetige Abbildung auf jeder kompakten Teilmenge ihres Definitionsgebietes gleichmäßig stetig ist.

(d) Wähle $\delta > 0$ so dass $\delta \leq r$ und $c\delta \leq \rho$ ist.

Wir zeigen, dass Behauptung 1 mit diesem δ gilt. Für $t \in [0, T]$ und $h \in \mathbb{R}^n$ mit $\|h\| < \rho$ sei

$$R(t, h) := f(x(t) + h) - f(x(t)) - df(x(t))h. \quad (4.4.9)$$

Dann gilt

$$R(t, h) = \int_0^1 (df(x(t) + sh) - df(x(t)))h \, ds$$

und daher, nach (c),

$$\|R(t, h)\| \leq \int_0^1 \|df(x(t) + sh) - df(x(t))\| \|h\| \, ds \leq e^{-cT} \varepsilon \|h\| \quad (4.4.10)$$

für alle $t \in [0, T]$ und alle $h \in \mathbb{R}^n$ mit $\|h\| < \rho$.

Sei nun $\xi_0 \in \mathbb{R}^n$ mit $\|\xi_0\| < \delta$ und definiere $\eta : [0, T] \rightarrow \mathbb{R}^n$ durch

$$\eta(t) := \phi_t(x_0 + \xi_0) - \phi_t(x_0) - \Phi(t)\xi_0 \quad \text{für } 0 \leq t \leq T. \quad (4.4.11)$$

Dann ist $\eta(0) = 0$ und

$$\begin{aligned} \dot{\eta}(t) &= f(\phi_t(x_0 + \xi_0)) - f(\phi_t(x_0)) - A(t)\Phi(t)\xi_0 \\ &= f(\phi_t(x_0 + \xi_0)) - f(\phi_t(x_0)) - A(t)(\phi_t(x_0 + \xi_0) - \phi_t(x_0)) + A(t)\eta(t) \\ &= R(t, \phi_t(x_0 + \xi_0) - \phi_t(x_0)) + A(t)\eta(t). \end{aligned}$$

Nach (a), (b), und (d) gilt $x_0 + \xi_0 \in U_t$ und

$$\|\phi_t(x_0 + \xi_0) - \phi_t(x_0)\| \leq c \|\xi_0\| < c\delta \leq \rho \quad \text{für } 0 \leq t \leq T.$$

Nach (4.4.10) folgt daraus die Ungleichung

$$\begin{aligned} \|R(t, \phi_t(x_0 + \xi_0) - \phi_t(x_0))\| &\leq e^{-cT} \varepsilon \|\phi_t(x_0 + \xi_0) - \phi_t(x_0)\| \\ &\leq ce^{-cT} \varepsilon \|\xi_0\| \end{aligned}$$

für $0 \leq t \leq T$. Daraus wiederum folgt

$$\begin{aligned} \|\dot{\eta}(t)\| &\leq \|R(t, \phi_t(x_0 + \xi_0) - \phi_t(x_0))\| + \|A(t)\eta(t)\| \\ &\leq ce^{-cT} \varepsilon \|\xi_0\| + c \|\eta(t)\| \\ &\leq ce^{-cT} \varepsilon \|\xi_0\| + c \int_0^t \|\dot{\eta}(s)\| ds \end{aligned}$$

für $0 \leq t \leq T$. Nach Gronwall's Lemma 4.3.2 gilt daher die Ungleichung

$$\|\dot{\eta}(t)\| \leq ce^{-cT} \varepsilon \|\xi_0\| e^{ct}$$

und daraus folgt

$$\begin{aligned} \|\eta(t)\| &\leq \int_0^t \|\dot{\eta}(s)\| ds \\ &\leq e^{-cT} \varepsilon \|\xi_0\| \int_0^t ce^{cs} ds \\ &\leq \varepsilon \|\xi_0\| \end{aligned}$$

für $0 \leq t \leq T$. Da $\eta(t)$ durch (4.4.11) definiert war, ist Behauptung 1 damit bewiesen. Damit haben wir gezeigt, dass ϕ_t für $t > 0$ differenzierbar ist und der Beweis für $t < 0$ ist analog.

Als nächstes zeigen wir, dass die Abbildung $\Omega \rightarrow \mathbb{R}^{n \times n} : (t, x) \mapsto d\phi_t(x)$ stetig ist. Sei also $x_0 \in U$ und sei $T \in I(x_0)$ mit $T > 0$. Definiere die Abbildungen $A, \Phi : [0, T] \rightarrow \mathbb{R}^{n \times n}$ durch

$$A(t) := df(\phi_t(x_0)), \quad \Phi(t) := d\phi_t(x_0)$$

für $0 \leq t \leq T$.

Behauptung 2: Für jedes $\varepsilon > 0$ existiert ein $\delta > 0$, so dass für alle $t \in [0, T]$ und alle $y_0 \in \mathbb{R}^n$ mit $\|x_0 - y_0\| < \delta$ gilt, dass $y_0 \in U_t$ ist und

$$\|d\phi_t(x_0) - d\phi_t(y_0)\| < \varepsilon. \quad (4.4.12)$$

Beweis von Behauptung 2. Sei eine Konstante $0 < \varepsilon \leq 1$ gegeben. Wir wählen die Konstante $\delta > 0$ in den folgenden drei Schritten.

(a) Wähle $c > 0$ so dass mit

$$\|A(t)\| + 1 \leq c, \quad \|\Phi(t)\| \leq c \quad \text{für } 0 \leq t \leq T.$$

(b) Es existiert eine Konstante $\rho > 0$, so dass für alle $t \in [0, T]$ und alle $y \in \mathbb{R}^n$ mit $\|y - \phi_t(x_0)\| < \rho$ gilt, dass $y \in U$ ist und

$$\|df(y) - df(\phi_t(x_0))\| < e^{-cT}\varepsilon.$$

(c) Nach Satz 4.3.1 existiert eine Konstante $\delta > 0$, so dass für alle $t \in [0, T]$ und alle $y_0 \in \mathbb{R}^n$ mit $\|x_0 - y_0\| < \delta$ gilt, dass $y_0 \in U_t$ ist und

$$\|\phi_t(x_0) - \phi_t(y_0)\| < \rho.$$

Sei nun $y_0 \in \mathbb{R}^n$ mit $\|x_0 - y_0\| < \delta$ und definiere $B, \Psi : [0, T] \rightarrow \mathbb{R}^{n \times n}$ durch

$$B(t) := df(\phi_t(y_0)), \quad \Psi(t) := d\phi_t(y_0)$$

für $0 \leq t \leq T$. Dann ist Ψ nach dem bisher bewiesenen stetig differenzierbar und es gilt $\dot{\Psi}(t) = B(t)\Psi(t)$ für $0 \leq t \leq T$ und $\Psi(0) = \mathbb{1}$. Außerdem gilt nach (c), dass $\|\phi_t(x_0) - \phi_t(y_0)\| < \rho$ ist für $0 \leq t \leq T$. Daraus folgt nach (b)

$$\|A(t) - B(t)\| \leq e^{-cT}\varepsilon \leq 1 \quad \text{für } 0 \leq t \leq T$$

und daraus wiederum folgt nach (a)

$$\|B(t)\| \leq \|A(t)\| + 1 \leq c \quad \text{für } 0 \leq t \leq T.$$

Aus diesen beiden Ungleichungen folgt

$$\begin{aligned}
\|\dot{\Phi}(t) - \dot{\Psi}(t)\| &= \|A(t)\Phi(t) - B(t)\Psi(t)\| \\
&\leq \|A(t)\Phi(t) - B(t)\Phi(t)\| + \|B(t)\Phi(t) - B(t)\Psi(t)\| \\
&< ce^{-cT}\varepsilon + c\|\Phi(t) - \Psi(t)\| \\
&\leq ce^{-cT}\varepsilon + c\int_0^T \|\dot{\Phi}(s) - \dot{\Psi}(s)\| ds
\end{aligned}$$

für $0 \leq t \leq T$. Nach Gronwall's Lemma 4.3.2 folgt daraus die Ungleichung $\|\dot{\Phi}(t) - \dot{\Psi}(t)\| < e^{-cT}ce^{cT}\varepsilon$ und daraus ergibt sich durch Integration die Ungleichung $\|\Phi(t) - \Psi(t)\| < \varepsilon$ für $0 \leq t \leq T$. Damit ist Behauptung 2 bewiesen. Wir haben also gezeigt, dass die Abbildung $(t, x) \mapsto d\phi_t(x)$ für $t \geq 0$ stetig ist und der Beweis für $t \leq 0$ ist analog. Damit sind alle partiellen Ableitungen $\partial_t \phi = f \circ \phi$ und $\partial \phi / \partial x_i$ für $i = 1, \dots, n$ stetig, und somit ist $\phi : \Omega \rightarrow U$ stetig differenzierbar.

Wir haben damit Satz 4.4.1 für $\ell = 1$ bewiesen. Sei nun $\ell \in \mathbb{N}$ und nehmen wir an, der Satz gelte für dieses ℓ (und jedes C^ℓ Vektorfeld auf jeder offenen Teilmenge eines Euklidischen Raumes beliebiger Dimension). Sei $f : U \rightarrow \mathbb{R}^n$ ein $C^{\ell+1}$ Vektorfeld und sei $\phi : \Omega \rightarrow \mathbb{R}^n$ der Fluss von f . Dann ist ϕ eine C^ℓ -Abbildung nach Induktionsannahme, und es ist zu zeigen, dass ϕ eine $C^{\ell+1}$ -Abbildung ist. Dazu definieren wir die offene Menge

$$\tilde{U} := U \times \mathbb{R}^n \subset \mathbb{R}^n \times \mathbb{R}^n$$

und die Abbildung $\tilde{f} : \tilde{U} \rightarrow \mathbb{R}^n \times \mathbb{R}^n$ durch

$$\tilde{f}(x, \xi) := (f(x), df(x)\xi) \quad \text{für } x \in U \text{ und } \xi \in \mathbb{R}^n.$$

Außerdem sei $\tilde{\Omega} := \Omega \times \mathbb{R}^n$ und sei die Abbildung $\tilde{\phi} : \tilde{\Omega} \rightarrow \tilde{U}$ durch

$$\tilde{\phi}(t, x, \xi) := (\phi_t(x), d\phi_t(x)\xi) \quad \text{für } (t, x) \in \Omega \text{ und } \xi \in \mathbb{R}^n$$

definiert. Dann ist \tilde{f} ein C^ℓ Vektorfeld und $\tilde{\phi}$ ist der Fluss von \tilde{f} wie wir oben bewiesen haben. Nach der Induktionsannahme ist $\tilde{\phi}$ daher eine C^ℓ -Abbildung. Daraus folgt, dass alle partiellen Ableitungen $\partial_t \phi = f \circ \phi$ und $\partial \phi / \partial x_i$ für $i = 1, \dots, n$ C^ℓ -Abbildungen sind, und damit ist gezeigt dass ϕ eine $C^{\ell+1}$ Abbildung ist.

Dieses Induktionsargument zeigt, dass die Behauptung des Satzes für jedes $\ell \in \mathbb{N}$ gilt. Daß sie auch für $\ell = \infty$ gilt, folgt daraus, dass eine Funktion genau dann glatt ist, wenn sie, für jedes $\ell \in \mathbb{N}$, ℓ mal stetig differenzierbar ist. Damit ist Satz 4.4.1 bewiesen. \square

Im Fall $\ell = \infty$ sagt Satz 4.4.1, dass der Fluss $\phi : \Omega \rightarrow U$ eines glatten Vektorfeldes $f : U \rightarrow \mathbb{R}^n$ ebenfalls eine glatte Abbildung ist, und insbesondere, dass die Abbildung $\phi_t : U_t \rightarrow U_{-t}$ in Bemerkung 4.3.3 für jedes $t \in \mathbb{R}$ ein glatter Diffeomorphismus ist. Dies liefert eine in vielen Situationen nützliche Methode zur Konstruktion glatter Diffeomorphismen, indem man zunächst ein glattes Vektorfeld mit den gewünschten Eigenschaften konstruiert, und anschließend die Diffeomorphismen über die Lösungen der Differentialgleichung definiert. Außerdem können wir Satz 4.4.1 verwenden, um zu zeigen, dass die Exponentialabbildung glatt ist.

Korollar 4.4.4. *Die Exponentialabbildung $\exp : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$ ist glatt.*

Beweis. Sei $f : \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n}$ das durch

$$f(A, B) := (0, AB) \quad \text{für } A, B \in \mathbb{R}^{n \times n}$$

definierte Vektorfeld. Der Fluss von f ist die durch

$$\phi(t, A, B) := (A, e^{At}B) \quad \text{für } t \in \mathbb{R} \text{ und } A, B \in \mathbb{R}^{n \times n}$$

gegebene Abbildung $\phi : \mathbb{R} \times \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n}$. Da f glatt ist, folgt aus Satz 4.4.1, dass ϕ glatt ist. Daher ist auch die Abbildung

$$\mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n} : A \mapsto \phi(1, A, \mathbb{1}) = (A, e^A)$$

glatt. Die Exponentialabbildung $\exp : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$ ist die Komposition dieser Abbildung mit der Projektion auf den zweiten Faktor und ist daher ebenfalls glatt. Damit ist Korollar 4.4.4 bewiesen. \square

Kapitel 5

Mehrfache Integrale

Das Ziel dieses Kapitels ist es, das Riemannsches Integral einer Funktion von mehreren Variablen über einer kompakten Jordan-messbaren Menge zu definieren, sowie die grundlegenden Eigenschaften des Riemanschen Integrals herzuleiten.

5.1 Das Riemannsches Integral

Zur Definition des Riemanschen Integrals bedarf es der Einführung einiger Grundbegriffe, wie des einer Partition in höheren Dimensionen.

Partitionen

Seien $a = (a_1, \dots, a_n)$ und $b = (b_1, \dots, b_n)$ zwei reelle Vektoren mit $a_i < b_i$ für $i = 1, \dots, n$. Die Menge

$$\begin{aligned} Q &:= Q(a, b) \\ &:= \{x \in \mathbb{R}^n \mid a_i < x_i < b_i \text{ für } i = 1, \dots, n\} \\ &= (a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_n, b_n) \end{aligned} \quad (5.1.1)$$

heißt (**achsenparalleler**) **offener Quader**. Das **Volumen von Q** ist die Zahl

$$\text{Vol}_n(Q) := \prod_{i=1}^n (b_i - a_i). \quad (5.1.2)$$

Es ist hervorzuheben, dass das Volumen damit bislang **nur** für achsenparallele offene Quader definiert ist. Ist $Q \subset \mathbb{R}^n$ durch (5.1.1) definiert, so ist der

Abschluss von Q (bezüglich der Euklidischen Norm auf dem \mathbb{R}^n) die Menge

$$\begin{aligned}\overline{Q} &= \{x \in \mathbb{R}^n \mid a_i \leq x_i \leq b_i \text{ für } i = 1, \dots, n\} \\ &= [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]\end{aligned}\tag{5.1.3}$$

(Übung: beweisen Sie diese Formel.) Jede solche Menge heißt (**achsenparalleler**) **abgeschlossener Quader**. Es ist oft nützlich, das **Volumen von \overline{Q}** ebenfalls durch die rechte Seite der Gleichung (5.1.2) zu definieren. Allgemeiner, definieren wir

$$\text{Vol}_n(B) := \prod_{i=1}^n (b_i - a_i) \quad \text{für } Q \subset B \subset \overline{Q}.$$

Definition 5.1.1. Sei $B \subset \mathbb{R}^n$ eine kompakte Teilmenge. Eine **Partition von B** ist eine endliche Menge $P = \{P_1, \dots, P_k\}$ von offenen achsenparallelen Quadern so dass

$$B = \bigcup_{i=1}^k \overline{P}_i, \quad P_i \cap P_j = \emptyset \quad \text{für } i \neq j.$$

Die Menge der Partitionen von B wird mit $\mathcal{P}(B)$ bezeichnet.

Ein Beispiel einer Partition im Fall $n = 2$ ist in Abbildung 5.1 dargestellt. Es ist zu beachten, dass nicht jede Teilmenge des \mathbb{R}^n eine Partition besitzt. Wenn $B \subset \mathbb{R}^n$ eine Partition besitzt so ist B notwendigerweise kompakt und nichtleer, und auch für nichtleere kompakte Teilmengen des \mathbb{R}^n ist die Existenz einer Partition eine *seltene* Eigenschaft.

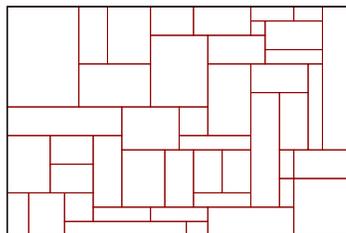


Abbildung 5.1: Eine Partition eines Rechtecks.

Beispiel 5.1.2. Im Fall $n = 1$ kann eine Partition eines kompakten Intervalls $B = [a, b]$ durch eine endliche Folge $a = t_0 < t_1 < t_2 < \cdots < t_k = b$ repräsentiert werden mit $P_i = (t_{i-1}, t_i)$ (siehe [5]).

Übung 5.1.3. Die abgeschlossene Einheitskreisscheibe im \mathbb{R}^2 besitzt keine Partition. Ein abgeschlossenes Dreieck im \mathbb{R}^2 besitzt keine Partition.

Übung 5.1.4. Eine Teilmenge $B \subset \mathbb{R}^n$ besitzt genau dann eine Partition, wenn sie eine endliche Vereinigung abgeschlossener achsenparalleler Quader ist. Eine solche Menge heißt **Quadergebäude** (Definition 5.6.1).

Übung 5.1.5. Seien $K \subset U \subset \mathbb{R}^n$, so dass U offen und K kompakt ist. Dann existiert eine Teilmenge $B \subset \mathbb{R}^n$ mit $\mathcal{P}(B) \neq \emptyset$ und $K \subset B \subset U$.

Lemma 5.1.6. Sei $B \subset \mathbb{R}^n$ eine kompakte Teilmenge und $P = \{P_1, \dots, P_k\}$ und $Q = \{Q_1, \dots, Q_\ell\}$ zwei Partitionen von B . Dann ist auch die Menge

$$P \wedge Q := \{P_i \cap Q_j \mid i = 1, \dots, k, j = 1, \dots, \ell, P_i \cap Q_j \neq \emptyset\} \quad (5.1.4)$$

eine Partition von B . Die Menge $\{P_i \cap Q_j \mid i = 1, \dots, k, P_i \cap Q_j \neq \emptyset\}$ ist für jedes j eine Partition von $\overline{Q_j}$, und $\{P_i \cap Q_j \mid j = 1, \dots, \ell, P_i \cap Q_j \neq \emptyset\}$ ist für jedes i eine Partition von $\overline{P_i}$.

Beweis. Es folgt aus den Definitionen, dass die Menge $P_i \cap Q_j$, wenn sie nichtleer ist, ein offener achsenparalleler Quader ist. Wir beweisen folgendes.

Behauptung: Für jedes $x \in \overline{P_i}$ existiert ein j so dass $x \in \overline{P_i \cap Q_j}$.

Nach Voraussetzung existiert eine Folge $x_\nu \in P_i$, die gegen x konvergiert. Da $P_i \subset B$ ist, gilt $x_\nu \in \overline{Q_1} \cup \dots \cup \overline{Q_\ell}$ für alle ν . Es gibt also (mindestens) ein $j \in \{1, \dots, \ell\}$ so dass $\overline{Q_j}$ unendlich viele Folgenglieder x_ν enthält. Durch Übergang zu einer Teilfolge können wir annehmen, dass $x_\nu \in \overline{Q_j}$ ist für alle ν . Wähle eine Folge $0 < \varepsilon_\nu < 1/\nu$ so dass

$$B_{\varepsilon_\nu}(x_\nu) \subset P_i$$

ist für alle ν . Dann gilt auch $P_{\varepsilon_\nu}(x_\nu) \cap Q_j \neq \emptyset$ für alle ν . Nach dem abzählbaren Auswahlaxiom gibt es also eine Folge $y_\nu \in Q_j$ so dass $\|x_\nu - y_\nu\| < \varepsilon_\nu$ ist für alle ν . Dann gilt $y_\nu \in P_i \cap Q_j$ und

$$x = \lim_{\nu \rightarrow \infty} x_\nu = \lim_{\nu \rightarrow \infty} y_\nu.$$

Also ist $x \in \overline{P_i \cap Q_j}$ und damit ist die Behauptung bewiesen.

Die Behauptung besagt, dass

$$\overline{P_i} = \bigcup_{j=1}^{\ell} \overline{P_i \cap Q_j}$$

ist für alle i . Ebenso gilt $\overline{Q_j} = \bigcup_{i=1}^k \overline{P_i \cap Q_j}$ und

$$B = \bigcup_{i=1}^k \overline{P_i} = \bigcup_{i=1}^k \bigcup_{j=1}^{\ell} \overline{P_i \cap Q_j}.$$

Damit ist Lemma 5.1.6 bewiesen. \square

Lemma 5.1.7. *Sei $Q = Q(a, b) \subset \mathbb{R}^n$ ein offener achsenparalleler Quader und $P = (P_1, \dots, P_k)$ eine Partition von $B := \overline{Q}$. Dann gilt*

$$\text{Vol}_n(Q) = \sum_{i=1}^k \text{Vol}_n(P_i).$$

Beweis. Der Beweis beruht auf einer Idee von John von Neumann (der von 1921 bis 1923 an der ETH studierte). Für jedes $\varepsilon > 0$ betrachten wir das Gitter

$$\Lambda_\varepsilon := \varepsilon\mathbb{Z}^n$$

und untersuchen die Anzahl der Gitterpunkte die dem offenen Quader Q , beziehungsweise dem abgeschlossenen Quader \overline{Q} angehören. Es gilt die folgende Ungleichung

$$\prod_{i=1}^n (b_i - a_i - \varepsilon) \leq \varepsilon^n \#(\Lambda_\varepsilon \cap Q) \leq \varepsilon^n \#(\Lambda_\varepsilon \cap \overline{Q}) \leq \prod_{i=1}^n (b_i - a_i + \varepsilon). \quad (5.1.5)$$

Zum Beweis von (5.1.5) betrachten wir zunächst den Fall $n = 1$. Seien also $a < b$ zwei reelle Zahlen und sei $\varepsilon > 0$. Wähle $\ell, m \in \mathbb{Z}$ so dass

$$(\ell - 1)\varepsilon < a \leq \ell\varepsilon, \quad m\varepsilon \leq b < (m + 1)\varepsilon.$$

Dann gilt $\varepsilon\mathbb{Z} \cap [a, b] = \{j\varepsilon \mid \ell \leq j \leq m\}$ und daher

$$\#(\varepsilon\mathbb{Z} \cap [a, b]) = m - \ell + 1 \leq \frac{b - a + \varepsilon}{\varepsilon}.$$

Nun wähle $\ell, m \in \mathbb{Z}$ so dass

$$(\ell - 1)\varepsilon \leq a < \ell\varepsilon, \quad m\varepsilon < b \leq (m + 1)\varepsilon.$$

Dann gilt $\varepsilon\mathbb{Z} \cap (a, b) = \{j\varepsilon \mid \ell \leq j \leq m\}$ und daher

$$\#(\varepsilon\mathbb{Z} \cap (a, b)) = m - \ell + 1 = (m + 1) - (\ell - 1) - 1 \geq \frac{b - a - \varepsilon}{\varepsilon}.$$

Daraus folgt die Ungleichung

$$b - a - \varepsilon \leq \varepsilon\#(\varepsilon\mathbb{Z} \cap (a, b)) \leq \varepsilon\#(\varepsilon\mathbb{Z} \cap [a, b]) \leq b - a + \varepsilon.$$

Ersetzen wir nun a, b durch a_i, b_i und bilden das Produkt über $i = 1, \dots, n$, so ergibt sich die Ungleichung (5.1.5).

Aus (5.1.5) folgt durch den Grenzübergang $\varepsilon \rightarrow 0$, dass

$$\text{Vol}_n(Q) = \lim_{\varepsilon \rightarrow 0} \varepsilon^n \#(\Lambda_\varepsilon \cap Q) = \lim_{\varepsilon \rightarrow 0} \varepsilon^n \#(\Lambda_\varepsilon \cap \overline{Q}). \quad (5.1.6)$$

Nun gilt offensichtlich

$$\sum_{i=1}^k \#(\Lambda_\varepsilon \cap P_i) \leq \#(\Lambda_\varepsilon \cap Q) \leq \sum_{i=1}^k \#(\Lambda_\varepsilon \cap \overline{P}_i)$$

Multiplizieren wir diese Ungleichung mit ε^n und bilden den Limes $\varepsilon \rightarrow 0$, so ergibt sich unter dreimaliger Verwendung von (5.1.6), dass

$$\begin{aligned} \sum_i \text{Vol}_n(P_i) &= \lim_{\varepsilon \rightarrow 0} \sum_i \varepsilon^n \#(\Lambda_\varepsilon \cap P_i) \\ &\leq \lim_{\varepsilon \rightarrow 0} \varepsilon^n \#(\Lambda_\varepsilon \cap Q) = \text{Vol}_n(Q) \\ &\leq \lim_{\varepsilon \rightarrow 0} \sum_i \varepsilon^n \#(\Lambda_\varepsilon \cap \overline{P}_i) \\ &= \sum_i \text{Vol}_n(P_i). \end{aligned}$$

Hieraus folgt sofort die Behauptung von Lemma 5.1.7. □

Obersummen und Untersummen

Definition 5.1.8. Sei $B \subset \mathbb{R}^n$ eine kompakte Menge und $P = \{P_1, \dots, P_k\}$ eine Partition von B . Sei $f : B \rightarrow \mathbb{R}$ eine beschränkte Funktion. Die **Obersumme von f und P** ist die Zahl

$$\overline{S}(f, P) := \sum_{i=1}^k \left(\sup_{\overline{P}_i} f \right) \text{Vol}_n(P_i). \quad (5.1.7)$$

Die **Untersumme von f und P** ist die Zahl

$$\underline{S}(f, P) := \sum_{i=1}^k \left(\inf_{\overline{P}_i} f \right) \text{Vol}_n(P_i). \quad (5.1.8)$$

Lemma 5.1.9. Sei $B \subset \mathbb{R}^n$ kompakt, sei $f : B \rightarrow \mathbb{R}$ eine beschränkte Funktion, und seien

$$P = \{P_1, \dots, P_k\}, \quad Q = \{Q_1, \dots, Q_\ell\}$$

Partitionen von B . Dann gilt

$$\underline{S}(f, P) \leq \underline{S}(f, P \wedge Q) \leq \overline{S}(f, P \wedge Q) \leq \overline{S}(f, Q).$$

Beweis. Es folgt aus Definition 5.1.8 sowie Lemma 5.1.6 and Lemma 5.1.7, dass

$$\begin{aligned} \underline{S}(f, P) &= \sum_i \left(\inf_{\overline{P}_i} f \right) \text{Vol}_n(P_i) \\ &= \sum_i \sum_j \left(\inf_{\overline{P}_i} f \right) \text{Vol}_n(P_i \cap Q_j) \\ &\leq \sum_i \sum_j \left(\inf_{\overline{P}_i \cap \overline{Q}_j} f \right) \text{Vol}_n(P_i \cap Q_j) = \underline{S}(f, P \wedge Q) \\ &\leq \sum_i \sum_j \left(\sup_{\overline{P}_i \cap \overline{Q}_j} f \right) \text{Vol}_n(P_i \cap Q_j) = \overline{S}(f, P \wedge Q) \\ &\leq \sum_j \sum_i \left(\sup_{\overline{Q}_j} f \right) \text{Vol}_n(P_i \cap Q_j) \\ &= \sum_j \left(\sup_{\overline{Q}_j} f \right) \text{Vol}_n(Q_j) \\ &= \overline{S}(f, Q). \end{aligned}$$

Damit ist Lemma 5.1.9 bewiesen. \square

Definition des Riemannsches Integrals

Nach Lemma 5.1.9 erfüllt jede kompakte Teilmenge $B \subset \mathbb{R}^n$ mit $\mathcal{P}(B) \neq \emptyset$ und jede beschränkte Funktion $f : B \rightarrow \mathbb{R}$ die Ungleichung

$$\sup_{P \in \mathcal{P}(B)} \underline{S}(f, P) \leq \inf_{Q \in \mathcal{P}(B)} \overline{S}(f, Q). \quad (5.1.9)$$

Definition 5.1.10. Sei $B \subset \mathbb{R}^n$ eine kompakte Menge mit $\mathcal{P}(B) \neq \emptyset$. Eine beschränkte Funktion $f : B \rightarrow \mathbb{R}$ heißt **Riemann-integrierbar**, wenn

$$\sup_{P \in \mathcal{P}(B)} \underline{S}(f, P) = \inf_{Q \in \mathcal{P}(B)} \overline{S}(f, Q)$$

ist. Falls $f : B \rightarrow \mathbb{R}$ Riemann-integrierbar ist, so wird die reelle Zahl

$$\int_B f(x) dx_1 \cdots dx_n := \sup_{P \in \mathcal{P}(B)} \underline{S}(f, P) = \inf_{Q \in \mathcal{P}(B)} \overline{S}(f, Q) \quad (5.1.10)$$

das (**Riemannsches**) **Integral von f über B** genannt.

Beispiel 5.1.11. Sei $B \subset \mathbb{R}^n$ eine kompakte Menge, die eine Partition besitzt. Als erstes Beispiel betrachten wir eine konstante Funktion $f : B \rightarrow \mathbb{R}$. Sei also $c \in \mathbb{R}$ so dass $f(x) = c$ ist für alle $x \in B$. Ist $P = \{P_1, \dots, P_k\}$ eine Partition von B so gilt $\inf_{\overline{P}_i} f = \sup_{\overline{P}_i} f = c$ für alle i und daher

$$\underline{S}(f, P) = \overline{S}(f, P) = c \sum_{i=1}^k \text{Vol}_n(P_i).$$

Daraus folgt nach Lemma 5.1.9, dass alle $P, Q \in \mathcal{P}(B)$ die Ungleichungen

$$\overline{S}(f, P) = \underline{S}(f, P) \leq \overline{S}(f, Q) = \underline{S}(f, Q) \leq \overline{S}(f, P)$$

erfüllen. Also sind Obersumme und Untersumme von f unabhängig von der Wahl der Partition. Daraus folgt, dass die Konstante Funktion $f \equiv c$ Riemann-integrierbar ist und die Gleichung

$$\int_B c dx_1 \cdots dx_n = c \sum_{i=1}^k \text{Vol}_n(P_i)$$

für jede Partition $P = \{P_1, \dots, P_k\} \in \mathcal{P}(B)$ gilt.

Definition 5.1.12 (Volumen eines Quadergebäudes). Sei $B \subset \mathbb{R}^n$ eine kompakte Menge mit $\mathcal{P}(B) \neq \emptyset$. Das Integral der konstanten Funktion $f \equiv 1$ über B wird das (n -dimensionale) **Volumen von B** genannt und mit

$$\text{Vol}_n(B) := \int_B 1 \, dx_1 \cdots dx_n \quad (5.1.11)$$

bezeichnet. Nach Beispiel 5.1.11 gilt $\text{Vol}_n(B) = \sum_{i=1}^k \text{Vol}_n(P_i)$ für jede Partition $P = \{P_1, \dots, P_k\} \in \mathcal{P}(B)$

Lemma 5.1.13. Sei $B \subset \mathbb{R}^n$ eine kompakte Teilmenge mit $\mathcal{P}(B) \neq \emptyset$ und sei $f : B \rightarrow \mathbb{R}$ eine beschränkte Funktion. Folgende Aussagen sind äquivalent.

(i) f ist Riemann-integrierbar.

(ii) Für jedes $\varepsilon > 0$ existiert eine Partition $P \in \mathcal{P}(B)$ mit

$$\overline{S}(f, P) - \underline{S}(f, P) < \varepsilon.$$

Proof. Wir zeigen zunächst dass (i) aus (ii) folgt. Sei also (ii) erfüllt. Dann gilt die Ungleichung

$$\inf_{Q \in \mathcal{P}(P)} \overline{S}(f, Q) < \sup_{P \in \mathcal{P}(P)} \underline{S}(f, P) + \varepsilon$$

für jedes $\varepsilon > 0$. Daraus folgt

$$\inf_{Q \in \mathcal{P}(P)} \overline{S}(f, Q) \leq \sup_{P \in \mathcal{P}(P)} \underline{S}(f, P).$$

Daher folgt aus der Ungleichung (5.1.9) dass f Riemann-integrierbar ist.

Wir zeigen dass (ii) aus (i) folgt. Sei also $f : B \rightarrow \mathbb{R}$ eine Riemann-integrierbare Funktion und sei

$$c := \int_B f(x) dx_1 \cdots dx_n.$$

Sei $\varepsilon > 0$. Dann existieren nach Definition des Riemannschen Integrals zwei Partitionen $P, Q \in \mathcal{P}(B)$ so dass

$$\underline{S}(f, P) > c - \frac{\varepsilon}{2}, \quad \overline{S}(f, Q) < c + \frac{\varepsilon}{2}.$$

Daraus folgt nach Lemma 5.1.9, dass

$$\overline{S}(f, P \wedge Q) - \underline{S}(f, P \wedge Q) \leq \overline{S}(f, Q) - \underline{S}(f, P) < \varepsilon$$

ist. Damit ist Lemma 5.1.13 bewiesen. \square

5.2 Jordansche Nullmengen und Stetigkeit

Mit Hilfe von Lemma 5.1.13 kann man leicht zeigen, dass jede stetige Funktion $f : B \rightarrow \mathbb{R}$ auf einer kompakten Teilmenge $B \subset \mathbb{R}^n$ mit $\mathcal{P}(B) \neq \emptyset$ Riemann-integrierbar ist. In diesem Abschnitt werden wir eine viel allgemeinere Aussage beweisen, welche die Riemannsche Integrierbarkeit einer beschränkten Funktion $f : B \rightarrow \mathbb{R}$ bereits dann garantiert, wenn sie ausserhalb einer sogenannten *Jordanschen Nullmenge* stetig ist.

Definition 5.2.1. Eine Teilmenge $A \subset \mathbb{R}^n$ heisst **Jordansche Nullmenge**, wenn für jedes $\varepsilon > 0$ endlich viele offene Quader $W_1, \dots, W_N \subset \mathbb{R}^n$ existieren mit $A \subset \bigcup_{\nu=1}^N W_\nu$ und $\sum_{\nu=1}^N \text{Vol}_n(W_\nu) < \varepsilon$.

Satz 5.2.2. Sei $B \subset \mathbb{R}^n$ eine kompakte Menge mit $\mathcal{P}(B) \neq \emptyset$, sei $A \subset B$ eine Jordansche Nullmenge, und sei $f : B \rightarrow \mathbb{R}$ eine beschränkte Funktion.

(i) Ist f auf $B \setminus A$ stetig, dann ist f Riemann-integrierbar.

(ii) Ist $f(x) = 0$ für alle $x \in B \setminus A$ so ist $\int_B f = 0$.

Beweis. Siehe Seite 169. □

Es folgt direkt aus Definition 5.2.1, dass jede Jordansche Nullmenge beschränkt ist, dass jede Teilmenge einer Jordanschen Nullmenge ebenfalls eine Jordansche Nullmenge ist, und dass jede endliche Vereinigung Jordanscher Nullmengen selbst wieder eine Jordansche Nullmenge ist.

Lemma 5.2.3. Ist $A \subset \mathbb{R}^n$ eine Jordansche Nullmenge so ist $\overset{\circ}{A} = \emptyset$ und ihr Abschluss \overline{A} ist ebenfalls eine Jordansche Nullmenge.

Beweis. Ist $\overset{\circ}{A} \neq \emptyset$, so existiert ein achsenparalleler offener Quader $Q \subset A$ mit $\delta := \text{Vol}_n(Q) > 0$. Sind nun $W_1, \dots, W_N \subset \mathbb{R}^n$ achsenparallele offene Quader mit $A \subset \bigcup_{\nu=1}^N W_\nu$, so folgt aus dem von Neumannschen Argument im Beweis von Lemma 5.1.7 die Ungleichung

$$\delta = \lim_{\varepsilon \rightarrow 0} \varepsilon^n \#(\Lambda_\varepsilon \cap Q) \leq \lim_{\varepsilon \rightarrow 0} \sum_{\nu=1}^N \varepsilon^n \#(\Lambda_\varepsilon \cap W_\nu) = \sum_{\nu=1}^N \text{Vol}_n(W_\nu),$$

im Widerspruch zu der Annahme, dass A eine Jordansche Nullmenge ist.

Sei nun $\varepsilon > 0$ gegeben und seien $W_1, \dots, W_N \subset \mathbb{R}^n$ achsenparallele offene Quader mit $A \subset \bigcup_{\nu=1}^N W_\nu$ und $\sum_{\nu=1}^N \text{Vol}_n(W_\nu) < \frac{\varepsilon}{2}$. Für $\nu = 1, \dots, N$ sei W'_ν ein achsenparalleler offener Quader mit $\overline{W}_\nu \subset W'_\nu$ und $\text{Vol}_n(W'_\nu) \leq 2\text{Vol}_n(W_\nu)$. Dann gilt $\overline{A} \subset \bigcup_{\nu=1}^N \overline{W}_\nu \subset \bigcup_{\nu=1}^N W'_\nu$ und $\sum_{\nu=1}^N \text{Vol}_n(W'_\nu) < \varepsilon$. Damit ist Lemma 5.2.3 bewiesen. □

Beispiel 5.2.4 (Jordansche Nullmengen).

(i) Ist $B \subset \mathbb{R}^{n-1}$ eine beschränkte Menge und ist $c \in \mathbb{R}$, so ist die Menge

$$A := B \times \{c\} \subset \mathbb{R}^{n-1} \times \mathbb{R} = \mathbb{R}^n$$

eine Jordansche Nullmenge. **Beweis:** Wähle $R > 0$ und $\delta > 0$ mit

$$B \subset (-R, R)^{n-1}, \quad \delta < \frac{\varepsilon}{2^n R^{n-1}}.$$

Dann ist die Menge $W := (-R, R)^{n-1} \times (c - \delta, c + \delta)$ ein offener Quader der A enthält und es gilt $\text{Vol}_n(W) = (2R)^{n-1} 2\delta < \varepsilon$.

(ii) Sei $Q \subset \mathbb{R}^n$ ein achsenparalleler offener Quader. Dann ist ∂Q eine Jordansche Nullmenge nach (i).

(iii) Eine nichtleere offene Teilmenge des \mathbb{R}^n ist nach Lemma 5.2.3 keine Jordansche Nullmenge.

(iv) Die abzählbare beschränkte Menge $A := \mathbb{Q} \cap [0, 1] \subset \mathbb{R}$ ist keine Jordansche Nullmenge, denn $\overline{A} = [0, 1]$ ist keine Jordansche Nullmenge nach (iii).

(v) Die standard Cantor-Menge $K \subset [0, 1]$ ist eine Jordansche Nullmenge.

(vi) Die Konstruktion der Cantor-Menge läßt sich so modifizieren, dass man eine kompakte Menge $K' \subset [0, 1]$ erhält, die keine Jordansche Nullmenge ist. In diesem Fall ist $U := [0, 1] \setminus K'$ eine offene Teilmenge von \mathbb{R} deren Rand $\partial U = K'$ keine Jordansche Nullmenge ist.

Beispiel 5.2.5 (Riemann-Integrierbarkeit und Monotonie).

(i) Sei $f : [0, 1]^2 \rightarrow \mathbb{R}$ eine Funktion, die in beiden Variablen monoton ist, das heißt, für alle $x, y, x', y' \in \mathbb{R}$ gilt

$$x \leq x' \text{ und } y \leq y' \quad \implies \quad f(x, y) \leq f(x', y'). \quad (5.2.1)$$

Dann ist f Riemann-integrierbar: Ist $P = \{P_{ij} \mid i, j = 1, \dots, N\} \in \mathcal{P}([0, 1]^2)$ mit $P_{ij} := (\frac{i-1}{N}, \frac{i}{N}) \times (\frac{j-1}{N}, \frac{j}{N})$, so gilt $\overline{S}(f, P) - \underline{S}(f, P) = \frac{f(1,1) - f(0,0)}{N}$.

(ii) Sei $\mathbb{N} \rightarrow \mathbb{Q} \cap [0, 1] : i \mapsto q_i$ eine bijektive Abbildung und sei die Funktion $f : [0, 1]^2 \rightarrow \mathbb{R}$ durch $f(x, 0) := f(0, y) := 0$ für $x, y \in [0, 1]$ und

$$f(x, y) := \sum_{i \in \mathbb{N}, q_i \leq x} 2^{-i} + \sum_{j \in \mathbb{N}, q_j \leq y} 2^{-j} \quad (5.2.2)$$

für $x, y > 0$ definiert. Diese Funktion erfüllt (5.2.1) und ist daher Riemann-integrierbar. Außerdem ist f an der Stelle $(x, y) \in [0, 1]^2$ genau dann stetig, wenn x und y beide irrational sind. Daher ist die Menge der Unstetigkeitsstellen von f keine Jordansche Nullmenge.

Definition 5.2.6. Sei $B \subset \mathbb{R}^n$ eine kompakte Menge und $P = \{P_1, \dots, P_k\}$ eine Partition von B . Dann ist jeder der offenen Quader P_i ein Produkt offener Intervalle

$$P_i =: (a_{i1}, b_{i1}) \times (a_{i2}, b_{i2}) \times \cdots \times (a_{in}, b_{in})$$

mit $a_{ij} < b_{ij}$. Für $i = 1, \dots, k$ heißt die Zahl

$$\delta(P_i) := \max \{b_{i1} - a_{i1}, b_{i2} - a_{i2}, \dots, b_{in} - a_{in}\}$$

die **maximale Seitenlänge von P_i** . Die Zahl

$$\delta(P) := \max_{i=1, \dots, k} \delta(P_i) = \max_{\substack{i=1, \dots, k \\ j=1, \dots, n}} (b_{ij} - a_{ij})$$

heißt **Feinheit der Partition P** .

Übung 5.2.7. Sei $B \subset \mathbb{R}^n$ eine kompakte Menge mit $\mathcal{P}(B) \neq \emptyset$. Dann existiert für jedes $\delta_0 > 0$ eine Partition $P \in \mathcal{P}(B)$ mit $\delta(P) < \delta_0$.

Das folgende Lemma ist von grundlegender Bedeutung für das Verständnis Jordanscher Nullmengen.

Lemma 5.2.8 (Das Nullmengen-Lemma). Sei $B \subset \mathbb{R}^n$ eine kompakte Menge mit $\mathcal{P}(B) \neq \emptyset$, sei $A \subset \mathbb{R}^n$ eine Jordansche Nullmenge, und sei $\varepsilon > 0$. Dann existiert eine Zahl $\delta_0 > 0$ so dass für alle $P \in \mathcal{P}(B)$ gilt

$$\delta(P) < \delta_0 \quad \implies \quad \sum_{\overline{P_i} \cap A \neq \emptyset} \text{Vol}_n(P_i) < \varepsilon.$$

Beweis. Sei $\varepsilon > 0$ gegeben. Nach Lemma 5.2.3 ist \overline{A} eine Jordansche Nullmenge. Daher existieren offene Quader $W_1, \dots, W_N \subset \mathbb{R}^n$ so dass

$$\overline{A} \subset \bigcup_{\nu=1}^N W_\nu, \quad \sum_{\nu=1}^N \text{Vol}_n(W_\nu) < \varepsilon. \quad (5.2.3)$$

Schritt 1. Es existiert ein $\rho > 0$, so dass für jedes $a \in \overline{A}$ ein $\nu \in \{1, \dots, N\}$ existiert mit $B_\rho(a) \subset W_\nu$.

Für jedes $a \in \overline{A}$ existiert eine Zahl $\nu = \nu(a) \in \{1, \dots, N\}$ so dass $a \in W_\nu$ ist. Da W_ν eine offene Menge ist, gibt es weiterhin für jedes $a \in A$ eine

Zahl $\rho(a) > 0$ mit $B_{2\rho(a)}(a) \subset W_{\nu(a)}$. Dann bilden die Mengen $B_{\rho(a)}(a)$, $a \in \bar{A}$, eine offene Überdeckung der Menge \bar{A} (Abschnitt C.1). Diese ist abgeschlossen und beschränkt, und daher nach dem Satz von Heine–Borel folgenkompakt. Nach Satz C.1.2 existieren daher endlich viele Elemente $a_1, \dots, a_m \in \bar{A}$ mit $\bar{A} \subset \bigcup_{i=1}^m B_{\rho(a_i)}(a_i)$. Für $i = 1, \dots, m$ führen wir nun die Abkürzungen $\rho_i := \rho(a_i)$ und $\nu_i := \nu(a_i)$ ein. Sei $\rho := \min\{\rho_1, \dots, \rho_m\} > 0$. Dann gilt $B_{\rho_i+\rho}(a_i) \subset B_{2\rho_i}(a_i) \subset W_{\nu_i}$ für $i = 1, \dots, m$. Sei nun $a \in \bar{A}$. Dann existiert ein $i \in \{1, \dots, m\}$ mit $a \in B_{\rho_i}(a_i)$ und daher $B_\rho(a) \subset B_{\rho+\rho_i}(a_i) \subset W_{\nu_i}$. Damit ist Schritt 1 bewiesen.

Schritt 2. Sei $\rho > 0$ wie in Schritt 1 und $\delta_0 := \rho/\sqrt{n}$. Sei $P = \{P_1, \dots, P_k\}$ eine Partition von B mit $\delta(P) < \delta_0$ und sei $j \in \{1, \dots, k\}$. Dann gilt

$$\bar{P}_j \cap A \neq \emptyset \quad \implies \quad \exists \nu \in \{1, \dots, N\} \text{ mit } \bar{P}_j \subset W_\nu.$$

Sei $a \in \bar{P}_j \cap A$. Nach Schritt 1 existiert ein $\nu \in \{1, \dots, N\}$ mit $B_\rho(a) \subset W_\nu$. Da $\delta(P_j) < \delta_0$ ist, erfüllt jedes $x \in \bar{P}_j$ die Ungleichung

$$\|x - a\| = \sqrt{\sum_{i=1}^n (x_i - a_i)^2} \leq \sqrt{n} \delta(P_j) < \sqrt{n} \delta_0 = \rho.$$

Also gilt $\bar{P}_j \subset B_\rho(a) \subset W_\nu$. Damit ist Schritt 2 bewiesen.

Schritt 3. Die Behauptung von Lemma 5.2.8 gilt mit $\delta_0 = \rho/\sqrt{n}$.

Sei $P = \{P_1, \dots, P_k\}$ eine Partition von B mit $\delta(P) < \delta_0$. Definiere

$$J_A := \{j \in \{1, \dots, k\} \mid \bar{P}_j \cap A \neq \emptyset\}$$

und, für $\nu = 1, \dots, N$,

$$J_\nu := \{j \in \{1, \dots, k\} \mid \bar{P}_j \subset W_\nu\}.$$

Dann gilt nach Schritt 2 $J_A \subset \bigcup_{\nu=1}^N J_\nu$. Daraus folgt

$$\sum_{j \in J_A} \text{Vol}_n(P_j) \leq \sum_{\nu=1}^N \left(\sum_{j \in J_\nu} \text{Vol}_n(P_j) \right) \leq \sum_{\nu=1}^N \text{Vol}_n(W_\nu) < \varepsilon.$$

Hier folgt die zweite Ungleichung aus Lemma 5.1.7 und die letzte Ungleichung aus (5.2.3). Damit ist Lemma 5.2.8 bewiesen. \square

Beweis von Satz 5.2.2. Der Beweis hat zwei Schritte.

Schritt 1. Ist $f(x) = 0$ für $x \in B \setminus A$ so ist $f \in \mathcal{R}(B)$ und

$$\int_B f = 0.$$

Wir nehmen ohne Einschränkung der Allgemeinheit an, dass f nicht identisch verschwindet. Also ist $\|f\| = \sup_B |f| \neq 0$. Sei nun $\varepsilon > 0$ gegeben. Nach Lemma 5.2.8 existiert eine Partition

$$P = \{P_1, \dots, P_k\} \in \mathcal{P}(B)$$

so dass

$$\sum_{\overline{P}_i \cap A \neq \emptyset} \text{Vol}_n(P_i) < \frac{\varepsilon}{4\|f\|}. \quad (5.2.4)$$

Dann gilt

$$\overline{S}(f, P) = \sum_{\overline{P}_i \cap A \neq \emptyset} \left(\sup_{\overline{P}_i} f \right) \text{Vol}_n(P_i) \leq \|f\| \sum_{\overline{P}_i \cap A \neq \emptyset} \text{Vol}_n(P_i) \leq \frac{\varepsilon}{4}$$

und ebenso

$$\underline{S}(f, P) = \sum_{\overline{P}_i \cap A \neq \emptyset} \left(\inf_{\overline{P}_i} f \right) \text{Vol}_n(P_i) \geq -\|f\| \sum_{\overline{P}_i \cap A \neq \emptyset} \text{Vol}_n(P_i) \geq -\frac{\varepsilon}{4}.$$

Damit folgt Schritt 1 aus Satz 5.3.1.

Schritt 2. Ist f auf $B \setminus A$ stetig, so ist $f \in \mathcal{R}(B)$.

Sei $\varepsilon > 0$ gegeben. Wir nehmen an, dass f nicht identisch verschwindet und wählen eine Partition $P = \{P_1, \dots, P_k\} \in \mathcal{P}(B)$ so dass (5.2.4) gilt. Betrachte die kompakte Menge

$$K := \bigcup_{\overline{P}_i \cap A = \emptyset} \overline{P}_i.$$

Die Restriktion $f|_K : K \rightarrow \mathbb{R}$ ist stetig. Nach einem Satz aus Analysis I ist $f|_K$ also gleichmäßig stetig. Sei

$$V := \sum_{i=1}^k \text{Vol}_n(P_i).$$

Dann existiert eine Konstante $\delta_0 > 0$ mit

$$x, y \in K, \quad \max_{\nu=1, \dots, n} |x_\nu - y_\nu| < \delta_0 \quad \implies \quad |f(x) - f(y)| \leq \frac{\varepsilon}{2V}.$$

Sei $Q = \{Q_1, \dots, Q_\ell\} \in \mathcal{P}(B)$ eine Partition mit $\delta(Q) < \delta_0$. Dann gilt

$$|f(x) - f(y)| \leq \frac{\varepsilon}{2V}$$

für alle $j = 1, \dots, \ell$ und alle $x, y \in Q_j \cap K$, und daher

$$\overline{P}_i \cap A = \emptyset \quad \implies \quad \sup_{P_i \cap Q_j} f - \inf_{P_i \cap Q_j} f \leq \frac{\varepsilon}{2V}$$

für alle i, j . Daraus folgt

$$\begin{aligned} & \overline{S}(f, P \wedge Q) - \underline{S}(f, P \wedge Q) \\ &= \sum_i \sum_j \left(\sup_{P_i \cap Q_j} f - \inf_{P_i \cap Q_j} f \right) \text{Vol}_n(P_i \cap Q_j) \\ &\leq \sum_{\overline{P}_i \cap A = \emptyset} \sum_j \left(\sup_{P_i \cap Q_j} f - \inf_{P_i \cap Q_j} f \right) \text{Vol}_n(P_i \cap Q_j) \\ &\quad + 2 \|f\| \sum_{\overline{P}_i \cap A \neq \emptyset} \sum_j \text{Vol}_n(P_i \cap Q_j) \\ &\leq \frac{\varepsilon}{2V} \sum_{\overline{P}_i \cap A = \emptyset} \sum_j \text{Vol}_n(P_i \cap Q_j) \\ &\quad + 2 \|f\| \sum_{\overline{P}_i \cap A \neq \emptyset} \sum_j \text{Vol}_n(P_i \cap Q_j) \\ &= \frac{\varepsilon}{2V} \sum_{\overline{P}_i \cap A = \emptyset} \text{Vol}_n(P_i) + 2 \|f\| \sum_{\overline{P}_i \cap A \neq \emptyset} \text{Vol}_n(P_i) \\ &< \frac{\varepsilon}{2V} \sum_i \text{Vol}_n(P_i) + \frac{\varepsilon}{2} \\ &= \varepsilon. \end{aligned}$$

Hier folgt die vorletzte Ungleichung aus (5.2.4). Nach Lemma 5.1.13 ist also f Riemann-integrierbar. Damit sind Schritt 2 und Satz 5.2.2 bewiesen. \square

5.3 Riemannsche Summen

Das erste Ziel dieses Abschnitts ist der Beweis des folgenden wichtigen Kriteriums für die Riemannsche Integrierbarkeit. Der Satz zeigt, dass eine beschränkte Funktion genau dann Riemann-integrierbar ist wenn die sogenannten *Riemannschen Summen* konvergieren. Dieser Satz wird verwendet, um einige grundlegende Eigenschaften des Riemannschen Integrals herzuleiten.

Satz 5.3.1 (Riemannsche Summen). *Sei $B \subset \mathbb{R}^n$ eine kompakte Menge mit $\mathcal{P}(B) \neq \emptyset$, sei $c \in \mathbb{R}$, und sei $f : B \rightarrow \mathbb{R}$ eine beschränkte Funktion. Dann sind folgende Aussagen äquivalent.*

(i) f ist Riemann-integrierbar und

$$c = \int_B f(x) dx_1 \cdots dx_n.$$

(ii) Für jedes $\varepsilon > 0$ existiert ein $P \in \mathcal{P}(B)$ so dass

$$c - \varepsilon < \underline{S}(f, P) \leq \overline{S}(f, P) < c + \varepsilon. \quad (5.3.1)$$

(iii) Für jedes $\varepsilon > 0$ existiert ein $\delta_0 > 0$ so dass für jedes $P \in \mathcal{P}(B)$ gilt

$$\delta(P) < \delta_0 \quad \implies \quad c - \varepsilon < \underline{S}(f, P) \leq \overline{S}(f, P) < c + \varepsilon. \quad (5.3.2)$$

(iv) Für jedes $\varepsilon > 0$ existiert eine reelle Zahl $\delta_0 > 0$ so dass für jede Partition $P = \{P_1, \dots, P_k\} \in \mathcal{P}(B)$ und alle $x_1, \dots, x_k \in \mathbb{R}^n$ gilt

$$\begin{array}{l} \delta(P) < \delta_0 \\ x_i \in \overline{P}_i \forall i \end{array} \quad \implies \quad \left| c - \sum_{i=1}^k f(x_i) \text{Vol}_n(P_i) \right| < \varepsilon. \quad (5.3.3)$$

Proof. Siehe Seite 173. □

Bemerkung 5.3.2. Die Zahlen $\sum_{i=1}^k f(x_i) \text{Vol}_n(P_i)$ in Teil (iv) von Satz 5.3.1 heißen **Riemannsche Summen** von f und das Kriterium in Teil (iv) von Satz 5.3.1 kann in der Formel

$$\int_B f(x) dx_1 \cdots dx_n = \lim_{\substack{\delta(P) \rightarrow 0 \\ P=(P_1, \dots, P_k) \in \mathcal{P}(B) \\ x_1 \in \overline{P}_1, \dots, x_k \in \overline{P}_k}} \sum_{i=1}^k f(x_i) \text{Vol}_n(P_i) \quad (5.3.4)$$

zusammengefasst werden. Die Riemannschen Summen konvergieren also gegen das Integral für $\delta(P) \rightarrow 0$.

Zum Beweis von Satz 5.3.1 benötigen wir das folgende Lemma.

Lemma 5.3.3. Sei $B \subset \mathbb{R}^n$ eine kompakte Teilmenge mit $\mathcal{P}(B) \neq \emptyset$, sei $Q = (Q_1, \dots, Q_\ell) \in \mathcal{P}(B)$, sei $f : B \rightarrow \mathbb{R}$ eine beschränkte Funktion, und sei $\varepsilon > 0$. Dann existiert ein $\delta_0 > 0$, so dass für alle $P \in \mathcal{P}(B)$ gilt

$$\delta(P) < \delta_0 \quad \Longrightarrow \quad \begin{aligned} \underline{S}(f, P) &\geq \underline{S}(f, Q) - \varepsilon, \\ \overline{S}(f, P) &\leq \overline{S}(f, Q) + \varepsilon. \end{aligned} \quad (5.3.5)$$

Beweis. Sei $M := \sup_{x \in B} |f(x)|$ und $A := \bigcup_{j=1}^{\ell} \partial Q_j$. Dann ist A eine Jordansche Nullmenge (Beispiel 5.2.4). Nach Lemma 5.2.8 existiert daher ein $\delta_0 > 0$, so dass für jede Partition $P = \{P_1, \dots, P_k\}$ von B gilt

$$\delta(P) < \delta_0 \quad \Longrightarrow \quad \sum_{\overline{P}_i \cap A \neq \emptyset} \text{Vol}_n(P_i) < \frac{\varepsilon}{2M}. \quad (5.3.6)$$

Sei nun $P = \{P_1, \dots, P_k\} \in \mathcal{P}(B)$ mit $\delta(P) < \delta_0$. Dann gilt

$$\begin{aligned} \underline{S}(f, P) &\leq \underline{S}(f, P \wedge Q) \\ &= \sum_{i=1}^k \sum_{j=1}^{\ell} \left(\inf_{\overline{P}_i \cap \overline{Q}_j} f \right) \text{Vol}_n(P_i \cap Q_j) \\ &= \sum_{\overline{P}_i \cap A = \emptyset} \sum_{j=1}^{\ell} \left(\inf_{\overline{P}_i \cap \overline{Q}_j} f \right) \text{Vol}_n(P_i \cap Q_j) \\ &\quad + \sum_{\overline{P}_i \cap A \neq \emptyset} \sum_{j=1}^{\ell} \left(\inf_{\overline{P}_i \cap \overline{Q}_j} f \right) \text{Vol}_n(P_i \cap Q_j) \\ &\leq \sum_{\overline{P}_i \cap A = \emptyset} \left(\inf_{\overline{P}_i} f \right) \text{Vol}_n(P_i) \\ &\quad + M \sum_{\overline{P}_i \cap A \neq \emptyset} \left(\sum_{j=1}^{\ell} \text{Vol}_n(P_i \cap Q_j) \right) \\ &\leq \underline{S}(f, P) + 2M \sum_{\overline{P}_i \cap A \neq \emptyset} \text{Vol}_n(P_i) \\ &< \underline{S}(f, P) + \varepsilon. \end{aligned}$$

Die Ungleichung $\overline{S}(f, P) > \overline{S}(f, Q) - \varepsilon$ zeigt man genauso. Damit ist Lemma 5.3.3 bewiesen. \square

Beweis von Satz 5.3.1. Wir zeigen zunächst, dass (iii) und (iv) äquivalent sind. Die Implikation (iii) \implies (iv) folgt sofort aus der Ungleichung

$$\underline{S}(f, P) \leq \sum_{i=1}^k f(x_i) \text{Vol}_n(P_i) \leq \overline{S}(f, P)$$

für jede Partition $P = \{P_1, \dots, P_k\}$ von B und alle $x_i \in \overline{P}_i$. Um die Umkehrung (iv) \implies (iii) zu zeigen, geben wir ein $\varepsilon > 0$ vor und wählen $\delta_0 > 0$, so dass (5.3.3) gilt. Sei $P = (P_1, \dots, P_k) \in \mathcal{P}(B)$ mit $\delta(P) < \delta_0$. Dann gilt

$$\underline{S}(f, P) = \inf_{x_i \in \overline{P}_i} \sum_{i=1}^k f(x_i) \text{Vol}_n(P_i), \quad \overline{S}(f, P) = \sup_{x_i \in \overline{P}_i} \sum_{i=1}^k f(x_i) \text{Vol}_n(P_i),$$

und nach (5.3.3) erhalten wir daher die Ungleichung

$$c - \varepsilon \leq \underline{S}(f, P) \leq \overline{S}(f, P) \leq c + \varepsilon.$$

Damit habe wir gezeigt, dass (iii) und (iv) äquivalent sind.

Wir beweisen (i) \implies (iii). Sei also $f : B \rightarrow \mathbb{R}$ Riemann-integrierbar, sei

$$c = \int_B f(x) dx_1 \cdots dx_n = \sup_{P \in \mathcal{P}(B)} \underline{S}(f, P) = \inf_{Q \in \mathcal{P}(B)} \overline{S}(f, Q),$$

und sei $\varepsilon > 0$. Nach Lemma 5.1.13 existiert eine Partition $Q \in \mathcal{P}(B)$ mit

$$\overline{S}(f, Q) - \underline{S}(f, Q) < \frac{\varepsilon}{2}.$$

Da $\underline{S}(f, Q) \leq c \leq \overline{S}(f, Q)$ ist, folgen daraus die Ungleichungen

$$c - \frac{\varepsilon}{2} < \underline{S}(f, Q) \leq c \leq \overline{S}(f, Q) < c + \frac{\varepsilon}{2}.$$

Nach Lemma 5.3.3 existiert ein $\delta_0 > 0$, so dass jede Partition $P \in \mathcal{P}(B)$ mit $\delta(P) < \delta_0$ die Ungleichungen

$$\underline{S}(f, P) \geq \underline{S}(f, Q) - \frac{\varepsilon}{2} > c - \varepsilon, \quad \overline{S}(f, P) \leq \overline{S}(f, Q) + \frac{\varepsilon}{2} < c + \varepsilon$$

erfüllt. Also erfüllt f die Bedingung (iii).

Die Implikation (iii) \implies (ii) folgt aus der Existenz einer Partition P von B mit $\delta(P) < \delta_0$ (Übung 5.2.7), und die Implikation (ii) \implies (i) folgt direkt aus der Definition des Riemannsches Integrals. Damit ist Satz 5.3.1 bewiesen. \square

Der Banachraum der integrierbaren Funktionen

Sei $B \subset \mathbb{R}^n$ eine kompakte Teilmenge mit $\mathcal{P}(B) \neq \emptyset$, das heißt B besitzt eine Partition (Definition 5.1.1). Wir bezeichnen die Menge der Riemann-integrierbaren Funktion auf B (Definition 5.1.10) mit

$$\mathcal{R}(B) := \{f : B \rightarrow \mathbb{R} \mid f \text{ ist Riemann-integrierbar}\}$$

und verwenden manchmal die Abkürzung $\int_B f := \int_B f(x) dx_1 \cdots dx_n$ für das Integral. Es sei daran erinnert, dass das **Volumen der Menge B** durch

$$\text{Vol}_n(B) := \int_B 1 = \overline{S}(1, P) = \underline{S}(1, P) = \sum_{i=1}^k \text{Vol}_n(P_i) \quad (5.3.7)$$

für $P = \{P_1, \dots, P_k\} \in \mathcal{P}(B)$ definiert ist. Diese Zahl ist unabhängig von P , wie wir in Beispiel 5.1.11 gesehen haben. Die Supremumsnorm einer beschränkten Funktion $f : B \rightarrow \mathbb{R}$ wird mit

$$\|f\| := \sup_{x \in B} |f(x)| \quad (5.3.8)$$

bezeichnet. Der folgende Satz fasst einige grundlegende Eigenschaften des Riemannsches Integrals zusammen. Insbesondere ist $\mathcal{R}(B)$ nach (i) und (v) ein Banachraum und die Abbildung $\mathcal{R}(B) \rightarrow \mathbb{R} : f \mapsto \int_B f$ ist nach (i) und (vi) ein beschränktes lineares Funktional.

Satz 5.3.4 (Eigenschaften des Riemannsches Integrals). *Sei $B \subset \mathbb{R}^n$ eine kompakte Teilmenge mit $\mathcal{P}(B) \neq \emptyset$. Dann gilt folgendes.*

(i) *Sind $f, g : B \rightarrow \mathbb{R}$ Riemann-integrierbar und $\lambda \in \mathbb{R}$, so sind auch die Funktionen $f + g : B \rightarrow \mathbb{R}$ und $\lambda f : B \rightarrow \mathbb{R}$ Riemann-integrierbar und es gilt*

$$\int_B (f + g) = \int_B f + \int_B g, \quad \int_B \lambda f = \lambda \int_B f.$$

(ii) *Sind $f, g : B \rightarrow \mathbb{R}$ Riemann-integrierbar so sind auch die Funktionen fg , $\max\{f, g\}$, $\min\{f, g\}$, $|f|$ Riemann-integrierbar.*

(iii) *Sind $f, g \in \mathcal{R}(B)$ und gilt $f(x) \leq g(x)$ für alle $x \in B$, so gilt $\int_B f \leq \int_B g$*

(iv) *Für jedes $f \in \mathcal{R}(B)$ gilt $|\int_B f| \leq \int_B |f| \leq \|f\| \text{Vol}_n(B)$.*

(v) *Wenn $f_\nu \in \mathcal{R}(B)$, $\nu \in \mathbb{N}$, eine Folge Riemann-integrierbarer Funktionen ist, die gleichmäßig gegen $f : B \rightarrow \mathbb{R}$ konvergiert, dann ist auch f Riemann-integrierbar, und es gilt $\int_B f = \lim_{\nu \rightarrow \infty} \int_B f_\nu$.*

(vi) *$\mathcal{R}(B)$ ein Banachraum bezüglich der Supremumsnorm und das lineare Funktional $\mathcal{R}(B) \rightarrow \mathbb{R} : f \mapsto \int_B f$ ist stetig bezüglich dieser Norm.*

Beweis. Wir beweisen (i). Seien $f, g : B \rightarrow \mathbb{R}$ Riemann-integrierbar und

$$a := \int_B f, \quad b := \int_B g.$$

Wir beweisen, dass $h := f + g : B \rightarrow \mathbb{R}$ und $c := a + b \in \mathbb{R}$ Bedingung (iv) in Satz 5.3.1 erfüllen. Sei $\varepsilon > 0$ gegeben. Dann existiert nach Satz 5.3.1 ein $\delta_0 > 0$, so dass jede Partition $P = \{P_1, \dots, P_k\} \in \mathcal{P}(B)$ mit $\delta(P) < \delta_0$ und alle Vektoren $x_i \in \overline{P}_i$, die Ungleichungen

$$\left| a - \sum_{i=1}^k f(x_i) \text{Vol}_n(P_i) \right| < \frac{\varepsilon}{2}, \quad \left| b - \sum_{i=1}^k g(x_i) \text{Vol}_n(P_i) \right| < \frac{\varepsilon}{2}.$$

erfüllen. Daraus folgt nach der Dreiecksungleichung

$$\left| a + b - \sum_{i=1}^k (f(x_i) + g(x_i)) \text{Vol}_n(P_i) \right| < \varepsilon,$$

ebenfalls für jede Partition $P = \{P_1, \dots, P_k\} \in \mathcal{P}(B)$ mit $\delta(P) < \delta_0$ und alle Vektoren $x_i \in \overline{P}_i$. Damit haben wir gezeigt, dass $h := f + g$ und $c := a + b$ die Bedingung (iv) in Satz 5.3.1 erfüllen. Also ist $f + g$ Riemann-integrierbar und

$$\int_B (f + g) = a + b = \int_B f + \int_B g.$$

Ebenso zeigt man, dass für jedes $f \in \mathcal{R}(B)$ und jede reelle Zahl λ , die Funktion $\lambda f : B \rightarrow \mathbb{R}$ Riemann-integrierbar und $\int_B \lambda f = \lambda \int_B f$ ist. Damit ist (i) bewiesen.

Wir beweisen (ii). Sei $P = \{P_1, \dots, P_k\}$ eine Partition von B . Dann gilt für jedes i und alle $x, y \in \overline{P}_i$

$$\begin{aligned} f(x)g(x) - f(y)g(y) &= f(x)(g(x) - g(y)) + (f(x) - f(y))g(y) \\ &\leq \|f\| (g(x) - g(y)) + \|g\| (f(x) - f(y)) \\ &\leq \|f\| \left(\sup_{\overline{P}_i} g - \inf_{\overline{P}_i} g \right) + \|g\| \left(\sup_{\overline{P}_i} f - \inf_{\overline{P}_i} f \right). \end{aligned}$$

Bilden wir das Supremum über $x, y \in \overline{P}_i$, so ergibt sich für $i = 1, \dots, k$ die Ungleichung

$$\sup_{\overline{P}_i} fg - \inf_{\overline{P}_i} fg \leq \|f\| \left(\sup_{\overline{P}_i} g - \inf_{\overline{P}_i} g \right) + \|g\| \left(\sup_{\overline{P}_i} f - \inf_{\overline{P}_i} f \right).$$

Daraus folgt

$$\overline{S}(fg, P) - \underline{S}(fg, P) \leq \|f\| (\overline{S}(g, P) - \underline{S}(g, P)) + \|g\| (\overline{S}(f, P) - \underline{S}(f, P)).$$

Wir zeigen nun, dass fg Riemann-integrierbar ist. Dazu nehmen wir ohne Einschränkung der Allgemeinheit an, dass weder f noch g identisch verschwindet, so dass $\|f\| \neq 0$ und $\|g\| \neq 0$ sind. Sei nun $\varepsilon > 0$ gegeben. Da f und g Riemann-integrierbar sind, existieren nach Lemma 5.1.13 zwei Partitionen $P, Q \in \mathcal{P}(B)$ mit

$$\overline{S}(f, P) - \underline{S}(f, P) \leq \frac{\varepsilon}{2\|g\|}, \quad \overline{S}(g, Q) - \underline{S}(g, Q) \leq \frac{\varepsilon}{2\|f\|}.$$

Unter Verwendung von Lemma 5.1.9 folgt daraus die Ungleichungskette

$$\begin{aligned} & \overline{S}(fg, P \wedge Q) - \underline{S}(fg, P \wedge Q) \\ & \leq \|f\| (\overline{S}(g, P \wedge Q) - \underline{S}(g, P \wedge Q)) + \|g\| (\overline{S}(f, P \wedge Q) - \underline{S}(f, P \wedge Q)) \\ & \leq \|f\| (\overline{S}(g, Q) - \underline{S}(g, Q)) + \|g\| (\overline{S}(f, P) - \underline{S}(f, P)) \\ & < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Daraus folgt nach Lemma 5.1.13, dass fg Riemann-integrierbar ist. Wir zeigen nun, dass $|f|$ Riemann-integrierbar ist. Dazu wählen wir wieder eine Partition $P = \{P_1, \dots, P_k\}$ von B . Dann gilt für jedes i und alle $x, y \in \overline{P}_i$

$$|f(x)| - |f(y)| \leq |f(x) - f(y)| \leq \sup_{\overline{P}_i} f - \inf_{\overline{P}_i} f.$$

Daraus folgt $\sup_{\overline{P}_i} |f| - \inf_{\overline{P}_i} |f| \leq \sup_{\overline{P}_i} f - \inf_{\overline{P}_i} f$ für alle i und daher

$$\overline{S}(|f|, P) - \underline{S}(|f|, P) \leq \overline{S}(f, P) - \underline{S}(f, P).$$

Also folgt aus Lemma 5.1.13, dass $|f|$ Riemann-integrierbar ist. Nun gilt

$$\max\{f, g\} + \min\{f, g\} = f + g, \quad \max\{f, g\} - \min\{f, g\} = |f - g|.$$

Diese beiden Funktionen sind nach dem bisher gezeigten Riemann-integrierbar. Also folgt aus (i), dass auch die Funktionen $\max\{f, g\}$ und $\min\{f, g\}$ Riemann-integrierbar sind. Damit ist (ii) bewiesen.

Wir beweisen (iii). Sind $f, g \in \mathcal{R}(B)$ und $f \leq g$ so gilt $\underline{S}(f, P) \leq \underline{S}(g, P)$ für jede Partition $P \in \mathcal{P}(B)$ und daraus folgt sofort $\int_B f \leq \int_B g$.

Wir beweisen (iv). Ist $f \in \mathcal{R}(B)$, so sind die Funktionen $-|f|$ und $|f|$ nach (i) und (ii) Riemann-integrierbar. Außerdem gilt $-|f| \leq f \leq |f|$. Also folgt aus (iii), dass

$$-\int_B |f| \leq \int_B f \leq \int_B |f|.$$

Dies ist äquivalent zu $|\int_B f| \leq \int_B |f|$. Außerdem ist das Integral einer konstanten Funktion $g(x) = c$ nach Beispiel 5.1.11 durch $\int_B c = c \text{Vol}_n(B)$ gegeben. Mit $c = \|f\| \geq |f|$ folgt daraus nach (iii) die Ungleichung

$$\|f\| \text{Vol}_n(B) = \int_B \|f\| \geq \int_B |f|.$$

Damit ist (iv) bewiesen.

Wir beweisen (v). Sei also $f_\nu : B \rightarrow \mathbb{R}$ eine Folge Riemann-integrierbarer Funktionen, die gleichmäßig gegen $f : B \rightarrow \mathbb{R}$ konvergiert. Aus der Definition der gleichmäßigen Konvergenz folgt, dass f beschränkt ist und dass die Folge f_ν bezüglich der Supremumsnorm gegen f konvergiert. Insbesondere ist also $(f_\nu)_{\nu \in \mathbb{N}}$ eine Cauchy-Folge bezüglich der Supremumsnorm. Nun definieren wir eine Folge $(c_\nu)_{\nu \in \mathbb{N}}$ durch

$$c_\nu := \int_B f_\nu, \quad \nu \in \mathbb{N}.$$

Nach (iv) gilt für alle $\nu, \nu' \in \mathbb{N}$ die Ungleichung

$$|c_\nu - c_{\nu'}| \leq \int_B |f_\nu - f_{\nu'}| \leq \|f_\nu - f_{\nu'}\| \text{Vol}_n(B). \quad (5.3.9)$$

Daraus folgt, dass $(c_\nu)_{\nu \in \mathbb{N}}$ eine Cauchy-Folge in \mathbb{R} ist. Da jede Cauchy-Folge reeller Zahlen konvergiert, existiert auch der Grenzwert

$$c := \lim_{\nu \rightarrow \infty} c_\nu.$$

Bilden wir in (5.3.9) nun den Grenzübergang $\nu' \rightarrow \infty$ so erhalten wir die Ungleichung

$$|c_\nu - c| \leq \|f_\nu - f\| \text{Vol}_n(B). \quad (5.3.10)$$

Wir beweisen, dass f Riemann-integrierbar und $\int_B f = c$ ist. Sei $\varepsilon > 0$ gegeben. Da f_ν in der Supremumsnorm gegen f konvergiert, existiert eine Zahl $\nu \in \mathbb{N}$ so dass

$$\|f_\nu - f\| < \frac{\varepsilon}{3 \text{Vol}_n(B)}, \quad |c_\nu - c| < \frac{\varepsilon}{3}. \quad (5.3.11)$$

Hier folgt die zweite Ungleichung aus der ersten und (5.3.10). Da $c_\nu = \int_B f_\nu$ ist, existiert nach Satz 5.3.1 eine Partition $P \in \mathcal{P}(B)$, so dass

$$c_\nu - \frac{\varepsilon}{3} < \underline{S}(f_\nu, P) \leq \overline{S}(f_\nu, P) \leq c_\nu + \frac{\varepsilon}{3}. \quad (5.3.12)$$

Außerdem folgt aus der Definition von Ober- und Untersumme, dass

$$\begin{aligned} |\overline{S}(f_\nu, P) - \overline{S}(f, P)| &\leq \|f - f_\nu\| \operatorname{Vol}_n(B) < \frac{\varepsilon}{3}, \\ |\underline{S}(f_\nu, P) - \underline{S}(f, P)| &\leq \|f - f_\nu\| \operatorname{Vol}_n(B) < \frac{\varepsilon}{3}. \end{aligned} \quad (5.3.13)$$

Kombinieren wir die Ungleichungen (5.3.11), (5.3.12), und (5.3.13) so erhalten wir

$$c - \varepsilon < c_\nu - \frac{2\varepsilon}{3} < \underline{S}(f_\nu, P) - \frac{\varepsilon}{3} < \underline{S}(f, P)$$

und

$$\overline{S}(f, P) < \overline{S}(f_\nu, P) + \frac{\varepsilon}{3} < c_\nu + \frac{2\varepsilon}{3} < c + \varepsilon.$$

Also folgt aus Satz 5.3.1, dass f Riemann-integrierbar und $\int_B f = c$ ist. Damit ist (v) bewiesen.

Wir beweisen (vi). Sei $f_\nu \in \mathcal{R}(B)$ eine Cauchy-Folge bezüglich der Supremumsnorm. Die Ungleichung

$$|f_\nu(x) - f_{\nu'}(x)| \leq \|f_\nu - f_{\nu'}\|$$

für alle $x \in B$ und alle $\nu, \nu' \in \mathbb{N}$ zeigt, dass die Folge $(f_\nu(x))_{\nu \in \mathbb{N}}$ in \mathbb{R} für jedes $x \in B$ eine Cauchy-Folge ist. Also folgt aus dem Vollständigkeitsaxiom für die reellen Zahlen, dass die Folge $(f_\nu(x))_{\nu \in \mathbb{N}}$ für jedes $x \in B$ konvergiert. Wir bezeichnen den Grenzwert mit

$$f(x) := \lim_{\nu \rightarrow \infty} f_\nu(x).$$

Dies definiert eine Funktion $f : B \rightarrow \mathbb{R}$. Wir zeigen, dass die Funktionenfolge f_ν gleichmäßig gegen f konvergiert. Sei $\varepsilon > 0$ gegeben. Da f_ν eine Cauchy-Folge bezüglich der Supremumsnorm ist, existiert eine natürlich Zahl $\nu_0 \in \mathbb{N}$ so dass für alle $\nu, \nu' \in \mathbb{N}$ gilt, dass

$$\nu, \nu' \geq \nu_0 \quad \implies \quad \|f_\nu - f_{\nu'}\| \leq \varepsilon.$$

Daraus folgt $|f_\nu(x) - f_{\nu'}(x)| \leq \varepsilon$ für alle $x \in B$ und alle $\nu, \nu' \geq \nu_0$. Bilden wir nun den Grenzwert $\nu' \rightarrow \infty$ so ergibt sich die Ungleichung

$$|f_\nu(x) - f(x)| = \lim_{\nu' \rightarrow \infty} |f_\nu(x) - f_{\nu'}(x)| \leq \varepsilon$$

für jedes $\nu \in \mathbb{N}$ mit $\nu \geq \nu_0$ und alle $x \in B$. Bilden wir das Supremum über $x \in B$, so erhalten wir die Ungleichung

$$\|f_\nu - f\| \leq \varepsilon.$$

für alle $\nu \in \mathbb{N}$ mit $\nu \geq \nu_0$. Mit anderen Worten, die Folge f_ν konvergiert gegen f in der Supremumsnorm, und dies ist gleichbedeutend mit der Aussage, dass die Folge f_ν gleichmäßig gegen f konvergiert. Nun folgt aus (iv), dass f Riemann-integrierbar ist. Also haben wir gezeigt, dass $\mathcal{R}(B)$ mit der Supremumsnorm ein vollständiger normierter Vektorraum und damit ein Banachraum ist. Die Stetigkeit des linearen Funktionals

$$\mathcal{R}(B) \rightarrow \mathbb{R} : f \mapsto \int_B f$$

folgt sofort aus (iv). Damit ist Satz 5.3.4 bewiesen. \square

Sei $B \subset \mathbb{R}^n$ eine Menge mit $\mathcal{P}(B) \neq \emptyset$ und sei

$$\mathcal{B}(B) := \{f : B \rightarrow \mathbb{R} \mid \exists c > 0 \forall x \in B \ |f(x)| \leq c\}$$

die Menge der beschränkten Funktionen. Dies ist ein Banachraum mit der Supremumsnorm

$$\|f\| = \sup_{x \in B} |f(x)|.$$

Die Menge $\mathcal{R}(B)$ der Riemann-integrierbaren Funktionen ist, nach Teil (iv) von Satz 5.3.4, ein abgeschlossener linearer Unterraum von $\mathcal{B}(B)$ und damit selbst wieder ein Banachraum. Weiterhin ist die Menge

$$\mathcal{C}(B) = \{f : B \rightarrow \mathbb{R} \mid f \text{ ist stetig}\}$$

aller stetigen Funktionen nach Satz 5.2.2 ein abgeschlossener linearer Unterraum von $\mathcal{R}(B)$ und ist damit ebenfalls ein Banachraum mit der Supremumsnorm (denn der Grenzwert einer gleichmäßig konvergierenden Folge stetiger Funktionen ist stetig). Damit haben wir folgende Inklusionen von Banachräumen mit der Supremumsnorm

$$\mathcal{C}(B) \subset \mathcal{R}(B) \subset \mathcal{B}(B).$$

5.4 Iterierte Integrale

Der **Satz von Fubini** hat eine zentrale Bedeutung für das Riemannsche Integral. Er wird hier zunächst bewiesen und anschließend für die Berechnung des Volumens des Einheitssimplex verwendet.

Satz 5.4.1 (Fubini). *Seien $p, q \in \mathbb{N}$ und $n := p + q$. Seien $A \subset \mathbb{R}^p$, $B \subset \mathbb{R}^q$ kompakte Mengen mit $\mathcal{P}(A) \neq \emptyset$ und $\mathcal{P}(B) \neq \emptyset$. Dann ist $\mathcal{P}(A \times B) \neq \emptyset$. Sei $f : A \times B \rightarrow \mathbb{R}$ eine Riemann-integrierbare Funktion und für jedes $x \in A$ sei die Funktion $B \rightarrow \mathbb{R} : y \mapsto f_x(y) := f(x, y)$ Riemann-integrierbar. Dann ist die Funktion $A \rightarrow \mathbb{R} : x \mapsto \int_B f_x$ Riemann-integrierbar und es gilt*

$$\int_{A \times B} f = \int_A \left(\int_B f(x, y) dy_1 \cdots dy_q \right) dx_1 \cdots dx_p. \quad (5.4.1)$$

Beweis. Definiere $F : A \rightarrow \mathbb{R}$ und $c \in \mathbb{R}$ durch

$$F(x) := \int_B f_x = \int_B f(x, y) dy_1 \cdots dy_q, \quad c := \int_{A \times B} f$$

Wir beweisen in zwei Schritten, dass $F \in \mathcal{R}(A)$ und $\int_A F = c$ ist.

Schritt 1. *Seien $P = \{P_1, \dots, P_k\} \in \mathcal{P}(A)$ und $Q = \{Q_1, \dots, Q_\ell\} \in \mathcal{P}(B)$ zwei Partitionen. Dann ist die Menge*

$$P \times Q := \{P_i \times Q_j \mid i = 1, \dots, k, j = 1, \dots, \ell\}$$

eine Partition von $A \times B$ mit $\delta(P \times Q) = \max\{\delta(P), \delta(Q)\}$ und es gilt

$$\underline{S}(f, P \times Q) \leq \underline{S}(F, P) \leq \overline{S}(F, P) \leq \overline{S}(f, P \times Q).$$

Für jedes $i \in \{1, \dots, k\}$ und jedes $j \in \{1, \dots, \ell\}$ ist die Menge $P_i \times Q_j$ ein offener Quader in $\mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^q$ mit Abschluss $\overline{P_i \times Q_j} = \overline{P_i} \times \overline{Q_j}$, maximaler Seitenlänge

$$\delta(P_i \times Q_j) = \max\{\delta(P_i), \delta(Q_j)\},$$

und n -dimensionalem Volumen

$$\text{Vol}_n(P_i \times Q_j) = \text{Vol}_p(P_i) \text{Vol}_q(Q_j).$$

Daher bilden die offenen Quader $P_i \times Q_j$ eine Partition $P \times Q$ von $A \times B$ mit der Feinheit $\delta(P \times Q) = \max\{\delta(P), \delta(Q)\}$. Außerdem gilt für $i \in \{1, \dots, k\}$ und $x \in \overline{P}_i$ die Ungleichung

$$\begin{aligned} F(x) &\leq \overline{S}(f_x, Q) \\ &= \sum_{j=1}^{\ell} \left(\sup_{y \in \overline{Q}_j} f(x, y) \right) \text{Vol}_q(Q_j) \\ &\leq \sum_{j=1}^{\ell} \left(\sup_{P_i \times Q_j} f \right) \text{Vol}_q(Q_j). \end{aligned}$$

Daraus folgt

$$\left(\sup_{\overline{P}_i} F \right) \text{Vol}_p(P_i) \leq \sum_{j=1}^{\ell} \left(\sup_{P_i \times Q_j} f \right) \text{Vol}_n(P_i \times Q_j).$$

Bilden wir nun die Summe über alle i so ergibt sich

$$\overline{S}(F, P) \leq \overline{S}(f, P \times Q).$$

Ebenso zeigt man die Ungleichung

$$\underline{S}(F, P) \geq \underline{S}(f, P \times Q).$$

Damit ist Schritt 1 bewiesen.

Schritt 2. Für jedes $\varepsilon > 0$ existiert eine Partition $P \in \mathcal{P}(A)$ so dass

$$c - \varepsilon < \underline{S}(F, P) \leq \overline{S}(F, P) < c + \varepsilon.$$

Sei $\varepsilon > 0$ gegeben. Nach Satz 5.3.1 existiert ein $\delta_0 > 0$ so dass jede Partition $R \in \mathcal{P}(A \times B)$ mit $\delta(R) < \delta_0$ die Ungleichung

$$c - \varepsilon < \underline{S}(f, R) \leq \overline{S}(f, R) < c + \varepsilon$$

erfüllt. Wähle eine Partition $P \in \mathcal{P}(A)$ mit Feinheit $\delta(P) < \delta_0$ und eine Partition $Q \in \mathcal{P}(B)$ mit Feinheit $\delta(Q) < \delta_0$. Nach Schritt 1 ist $P \times Q$ eine Partition von $A \times B$ mit der Feinheit $\delta(P \times Q) < \delta_0$ und es gilt

$$c - \varepsilon < \underline{S}(f, P \times Q) \leq \underline{S}(F, P) \leq \overline{S}(F, P) \leq \overline{S}(f, P \times Q) < c + \varepsilon.$$

Damit ist Schritt 2 bewiesen.

Nach Schritt 2 und Satz 5.3.1 ist die Funktion $F : A \rightarrow \mathbb{R}$ Riemann-integrierbar und ihr Integral ist $\int_A F = c$. Damit ist Satz 5.4.1 bewiesen. \square

Beispiel 5.4.2. Dieses Beispiel zeigt, dass aus der Integrierbarkeit von f in Satz 5.4.1 nicht die Integrierbarkeit von f_x für jedes $x \in A$ folgt. Sei

$$A = B = [0, 1]$$

und sei $K \subset [0, 1]$ die Cantor-Menge. Dann ist K eine Jordansche Nullmenge in \mathbb{R} und $K \times [0, 1]$ eine Jordansche Nullmenge in \mathbb{R}^2 (Übung). Nun definieren wir die Funktion $f : A \times B \rightarrow \mathbb{R}$ durch

$$f(x, y) := \begin{cases} 1, & \text{für } x \in K, y \in [0, 1] \cap \mathbb{Q}, \\ 0, & \text{sonst.} \end{cases}$$

Diese Funktion ist stetig auf $(A \times B) \setminus (K \times B)$ und ist daher nach Satz 5.2.2 Riemann integrierbar. Jedoch ist $f_x : B \rightarrow \mathbb{R}$ nur für $x \in A \setminus K$ Riemann-integrierbar, nicht aber für $x \in K$.

Beispiel 5.4.3. Dieses Beispiel zeigt, dass aus der Integrierbarkeit von f_x für jedes $x \in A$ nicht die Integrierbarkeit von f folgt. Sei $A = B = [0, 1]$ und definiere $f : A \times B \rightarrow \mathbb{R}$ durch

$$f(x, y) := \begin{cases} 1, & \text{für } x \in [0, 1] \cap \mathbb{Q}, \\ 0, & \text{für } x \in [0, 1] \setminus \mathbb{Q}. \end{cases}$$

Dann ist $f_x : B \rightarrow \mathbb{R}$ für jedes $x \in [0, 1]$ konstant und daher Riemann-integrierbar. Jedoch ist $\underline{S}(f, P) = 0$ und $\overline{S}(f, P) = 1$ für jede Partition P von $[0, 1]^2$. Daher ist f nicht Riemann-integrierbar.

Beispiel 5.4.4. Seien $A \subset \mathbb{R}^p$ und $B \subset \mathbb{R}^q$ kompakte Mengen mit $\mathcal{P}(A) \neq \emptyset$ und $\mathcal{P}(B) \neq \emptyset$. Seien $f : A \rightarrow \mathbb{R}$ und $g : B \rightarrow \mathbb{R}$ Riemann-integrierbar. Definiere $h : A \times B \rightarrow \mathbb{R}$ durch

$$h(x, y) := f(x)g(y).$$

Dann ist $h : A \times B \rightarrow \mathbb{R}$ Riemann-integrierbar (Übung). Außerdem ist die Funktion $h_x := h(x, \cdot) = f(x)g$ für jedes $x \in A$ Riemann-integrierbar. Also folgt aus Satz 5.4.1 und Satz 5.3.4 (i) die Formel

$$\int_{A \times B} h = \left(\int_A f \right) \left(\int_B g \right).$$

Beispiel 5.4.5. Für $n \in \mathbb{N}$ und $m \in \mathbb{N}_0$ sei $f_{n,m} : [0, 1]^n \rightarrow \mathbb{R}$ die Funktion

$$f_{n,m}(x) := \begin{cases} (1 - x_1 - \cdots - x_n)^m, & \text{falls } x_1 + \cdots + x_n \leq 1, \\ 0, & \text{falls } x_1 + \cdots + x_n > 1. \end{cases} \quad (5.4.2)$$

Dann ist $f_{n,m}$ Riemann-integrierbar und

$$\int_{[0,1]^n} f_{n,m} = \frac{1}{(m+1)(m+2)\cdots(m+n)} = \frac{m!}{(m+n)!}. \quad (5.4.3)$$

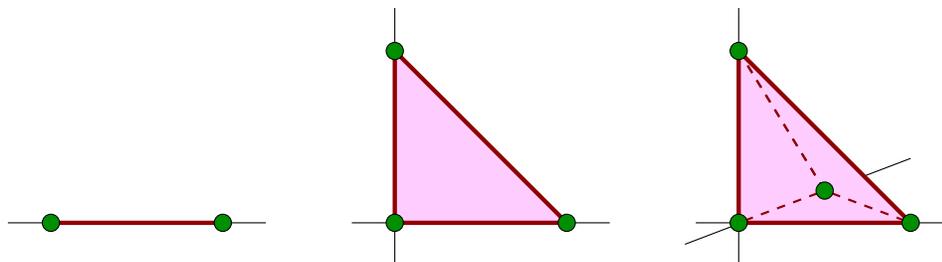
Zunächst ist $f_{n,m}$ für $m \geq 1$ stetig, und $f_{n,0}$ ist stetig auf dem Komplement der Menge $A := \{x \in [0, 1]^n \mid x_1 + \cdots + x_n = 1\}$. Dass es sich bei A um eine Jordansche Nullmenge handelt, folgt daraus, dass A der Graph seiner stetigen Funktion ist (siehe auch Beispiel 5.6.10). Nach Satz 5.2.2 sind die Funktionen $f_{n,m}$ daher Riemann-integrierbar. Wir verifizieren nun die Gleichung (5.4.3) durch vollständige Induktion über n . Zunächst gilt für jede ganze Zahl $m \geq 0$ die Gleichung

$$\int_0^1 f_{1,m}(x) = \int_0^1 (1-x)^m dx = \int_0^1 t^m dt = \frac{1}{m+1}.$$

Damit ist (5.4.3) für $n = 1$ gezeigt. Wir nehmen nun an, dass $n \geq 2$ ist und die Gleichung (5.4.3) für $n - 1$ gilt. Nach Satz 5.4.1, und mit Δ^{n-1} wie in (5.4.4) auf der nächsten Seite, erhalten wir dann die Gleichung

$$\begin{aligned} \int_{[0,1]^n} f_{n,m} &= \int_{[0,1]^{n-1}} \left(\int_0^1 f_{n,m}(x) dx_n \right) dx_1, \dots, dx_{n-1} \\ &= \int_{\Delta^{n-1}} \left(\int_0^{1-\sum_{i=1}^{n-1} x_i} \left(1 - \sum_{i=1}^n x_i \right)^m dx_n \right) dx_1 \cdots dx_{n-1} \\ &= \int_{\Delta^{n-1}} \frac{(1 - x_1 - \cdots - x_{n-1})^{m+1}}{m+1} dx_1 \cdots dx_{n-1} \\ &= \frac{1}{m+1} \int_{[0,1]^{n-1}} f_{n-1,m+1}(x) dx_1 \cdots dx_{n-1} \\ &= \frac{1}{(m+1)(m+2)\cdots(m+n)}. \end{aligned}$$

Der letzte Schritt folgt aus der Induktionsannahme. Außerdem steht $\int_{\Delta^{n-1}} f$ für das Integral über $[0, 1]^{n-1}$ von der Funktion, die auf Δ^{n-1} mit f übereinstimmt und ausserhalb von Δ^{n-1} gleich Null ist.

Abbildung 5.2: Der Einheitssimplex für $n = 1, 2, 3$.

Für $m = 0$ kann man die Formel (5.4.3) wie folgt geometrisch interpretieren. Der **Einheitssimplex** im \mathbb{R}^n ist die Menge

$$\Delta^n := \left\{ x \in \mathbb{R}^n \mid x_i \geq 0 \forall i, \sum_{i=1}^n x_i \leq 1 \right\}. \quad (5.4.4)$$

Die Funktion $f_{n,0} : [0, 1]^n \rightarrow \mathbb{R}$ ist die **charakteristische Funktion** dieser Menge, das heißt

$$f_{n,0}(x) = \begin{cases} 1, & \text{für } x \in \Delta^n, \\ 0, & \text{für } x \notin \Delta^n. \end{cases}$$

Wir könnten also das Integral

$$\int_{[0,1]^n} f_{n,0}(x) dx_1 \cdots dx_n = \frac{1}{n!}$$

auch interpretieren als das Integral der konstanten Funktion $f(x) = 1$ auf dem Einheitssimplex Δ^n und ihr Integral als das **Volumen des Einheitssimplex**. Dies könnte man in der Form

$$\text{Vol}_n(\Delta^n) = \int_{\Delta^n} 1 dx_1 \cdots dx_n = \frac{1}{n!}$$

schreiben. Der Einheitssimplex im \mathbb{R}^n hat also das Volumen $1/n!$. Jedoch ist hier darauf zu achten, dass wir bislang das Integral einer Funktion nur über Mengen definiert haben, die eine Partition besitzen. Die Menge Δ^n besitzt aber keine Partition. Es ist daher sinnvoll, den Integralbegriff etwas zu erweitern. Das soll im folgenden Abschnitt genauer ausgeführt werden.

5.5 Jordan-messbare Mengen

Bisher haben wir lediglich das Volumen von solchen Teilmengen des \mathbb{R}^n definiert, welche Partitionen besitzen (Definition 5.1.12). Dies ist eine sehr restriktive Bedingung, und in diesem Abschnitt geht es darum gehen, die Klasse der Mengen, deren Volumen wir definieren können, dramatisch zu erweitern.

Definition 5.5.1. *Eine beschränkte Teilmenge $B \subset \mathbb{R}^n$ heißt **Jordan-messbar**, wenn ihr Rand ∂B eine Jordansche Nullmenge ist.*

Definition 5.5.2 (Integration über Jordan-messbare Mengen).

Sei $B \subset \mathbb{R}^n$ eine Jordan-messbare Menge und $Q \subset \mathbb{R}^n$ ein abgeschlossener achsenparalleler Quader, der B enthält. Sei $f : Q \rightarrow \mathbb{R}$ eine Riemann-integrierbare Funktion. Die Zahl

$$\int_B f := \int_Q 1_B(x) f(x) dx_1 \cdots dx_n \quad (5.5.1)$$

*heißt **Integral von f über B** . Hier bezeichnen wir mit*

$$1_B : \mathbb{R}^n \rightarrow \mathbb{R}, \quad 1_B(x) := \begin{cases} 1, & \text{für } x \in B, \\ 0, & \text{für } x \in \mathbb{R}^n \setminus B, \end{cases} \quad (5.5.2)$$

*die **charakteristische Funktion von B** . Die reelle Zahl*

$$\mu_n(B) := \text{Vol}_n(B) := \int_B 1 = \int_Q 1_B(x) dx_1 \cdots dx_n \quad (5.5.3)$$

*wird das **Jordan-Maß von B** oder das **Volumen von B** genannt.*

Bemerkung 5.5.3. (i) Die Funktion $1_B : Q \rightarrow \mathbb{R}$ ist stetig auf $Q \setminus \partial B$. Da ∂B eine Jordansche Nullmenge ist, ist 1_B nach Satz 5.2.2 Riemann-integrierbar. Nach Satz 5.3.4 ist $1_B f : Q \rightarrow \mathbb{R}$ Riemann-integrierbar.

(ii) Das Integral $\int_Q 1_B f$ in (5.5.1) ist unabhängig von der Wahl des abgeschlossenen achsenparallelen Quaders Q , der B enthält.

(iii) Sei $B \subset \mathbb{R}^n$ eine Jordansche Nullmenge. Dann ist $\partial B = \overline{B}$ eine Jordansche Nullmenge nach Lemma 5.2.3. Daher ist B Jordan-messbar und es gilt $\text{Vol}_n(B) = 0$ nach Satz 5.2.2.

(iv) Ist $Q \subset \mathbb{R}^n$ ein abgeschlossene achsenparallele Quader, so ist ∂Q eine Jordansche Nullmenge nach Beispiel 5.2.4 (ii). Daher ist Q Jordan-messbar. Außerdem stimmt in diesem Fall die Definition des Volumens in Definition 5.5.2 mit der in (5.1.2) nach Beispiel 5.1.11 überein. Für den Beweis dieser Übereinstimmung kann auch Satz 5.4.1 verwendet werden.

(v) Sei $B \subset \mathbb{R}^n$ eine kompakte Menge mit $\mathcal{P}(B) \neq \emptyset$. Sei $P = \{P_1, \dots, P_k\}$ eine Partition von B . Dann ist $\partial B \subset \bigcup_i \partial P_i$. Daher ist B Jordan-messbar und hat das Volumen $\text{Vol}_n(B) = \int_B 1 = \sum_{i=1}^k \text{Vol}_n(P_i)$ nach Beispiel 5.1.11. Also stimmt die Definition des Volumens von B in Definition 5.5.2 mit der in Definition 5.1.12 überein.

(vi) Nach Beispiel 5.4.5 ist der standard n -Simplex $\Delta^n \subset \mathbb{R}^n$ Jordan-messbar und hat das Volumen $\text{Vol}_n(\Delta^n) = 1/n!$.

(vii) Die Menge $B := [0, 1] \cap \mathbb{Q}$ ist nicht Jordan-messbar, da $\partial B = [0, 1]$ ist.

Lemma 5.5.4. *Seien $A, B \subset \mathbb{R}^n$ zwei Jordan-messbare Mengen, $Q \subset \mathbb{R}^n$ ein abgeschlossener achsenparalleler Quader mit $A, B \subset Q$, und $f : Q \rightarrow \mathbb{R}$ eine Riemann-integrierbare Funktion. Dann gilt folgendes.*

(i) *Die Mengen $A \cap B$, $A \cup B$, $A \setminus B$, $B \setminus A$ sind Jordan-messbar.*

(ii) *Es gilt*

$$\int_{A \cup B} f + \int_{A \cap B} f = \int_A f + \int_B f \quad (5.5.4)$$

und $\text{Vol}_n(A \cup B) + \text{Vol}_n(A \cap B) = \text{Vol}_n(A) + \text{Vol}_n(B)$.

(iii) *Ist $A \subset B$ so gilt $\text{Vol}_n(A) \leq \text{Vol}_n(B)$.*

Beweis. Teil (i) folgt aus der Tatsache, dass jede der Mengen $\partial(A \cup B)$, $\partial(A \cap B)$, $\partial(A \setminus B)$, $\partial(B \setminus A)$ in der Jordanschen Nullmenge $\partial A \cup \partial B$ enthalten, und damit selbst eine Jordansche Nullmenge ist. Teil (ii) folgt aus Satz 5.3.4 (i) und der Tatsache, dass $1_{A \cup B} + 1_{A \cap B} = 1_A + 1_B$ ist. Teil (iii) folgt aus Satz 5.3.4 (iii) und der Ungleichung $1_A \leq 1_B$ für $A \subset B$. \square

Definition 5.5.5. *Sei $B \subset \mathbb{R}^n$ eine Jordan-messbare Menge. Eine Funktion $f : B \rightarrow \mathbb{R}$ heißt **Riemann-integrierbar**, wenn die erweiterte Funktion*

$$\tilde{f} : \mathbb{R}^n \rightarrow \mathbb{R}, \quad \tilde{f}(x) := \begin{cases} f(x), & \text{für } x \in B, \\ 0, & \text{für } x \in \mathbb{R}^n \setminus B, \end{cases}$$

über jedem abgeschlossenen Quader $Q \subset \mathbb{R}^n$, der B enthält, Riemann-integrierbar ist. In diesem Fall heißt die Zahl

$$\int_B f := \int_Q \tilde{f}(x) dx_1 \cdots dx_n$$

das **Integral von f über B** . Nach Lemma 5.5.4 ist das Integral unabhängig von der Wahl des abgeschlossenen achsenparallelen Quaders, auf den f erweitert wird. Die Menge der Riemann-integrierbaren Funktionen $f : B \rightarrow \mathbb{R}$ wird mit $\mathcal{R}(B)$ bezeichnet.

Bemerkung 5.5.6. Sei $B \subset \mathbb{R}^n$ eine Jordan-messbare Menge. Dann ist jede beschränkte stetige Funktion $f : B \rightarrow \mathbb{R}$ Riemann-integrierbar. Dies folgt aus Satz 5.2.2 und der Tatsache, dass die erweiterte Funktion $\tilde{f} : Q \rightarrow \mathbb{R}$ in Definition 5.5.5 auf $Q \setminus \partial B$ stetig ist.

Mit dieser Definition haben wir das Riemannsches Integral auf Funktionen, die auf beliebigen Jordan-messbaren Mengen $B \subset \mathbb{R}^n$ definiert sind, erweitert. Insbesondere ist jede beschränkte stetige Funktion $f : B \rightarrow \mathbb{R}$ nach Satz 5.2.2 und Definition 5.5.5 Riemann-integrierbar. Außerdem bleiben alle Aussagen von Satz 5.3.4 gültig für Jordan-messbare Mengen B . Darüber hinaus läßt sich der Satz von Fubini wie folgt verallgemeinern.

Korollar 5.5.7 (Fubini). Seien $p, q \in \mathbb{N}$ und $n := p + q$, und seien $P \subset \mathbb{R}^p$ und $Q \subset \mathbb{R}^q$ abgeschlossene achsenparallele Quader.

(i) Sei $B \subset P \times Q$ eine Jordan-messbare Menge, so dass, für jedes $x \in P$, die Menge $B_x := \{y \in \mathbb{R}^q \mid (x, y) \in B\} \subset Q$ Jordan-messbar ist. Dann ist die Funktion $P \rightarrow \mathbb{R} : x \mapsto \text{Vol}_q(B_x)$ Riemann-integrierbar und es gilt

$$\text{Vol}_n(B) = \int_P \text{Vol}_q(B_x) dx_1 \cdots dx_p.$$

(ii) Sei $B \subset P \times Q$ wie in (i). Sei $f : B \rightarrow \mathbb{R}$ Riemann-integrierbar und für jedes $x \in P$ sei die Funktion $B_x \rightarrow \mathbb{R} : y \mapsto f_x(y) := f(x, y)$ Riemann-integrierbar. Dann ist die Funktion $P \rightarrow \mathbb{R} : x \mapsto \int_{B_x} f_x$ Riemann-integrierbar und es gilt

$$\int_B f = \int_P \left(\int_{B_x} f_x \right) dx_1 \cdots dx_p.$$

Beweis. Wir beweisen (ii). Definiere $\tilde{f} : P \times Q \rightarrow \mathbb{R}$ durch $\tilde{f}(x, y) := f(x, y)$ für $(x, y) \in B$ und $\tilde{f}(x, y) := 0$ für $(x, y) \in (P \times Q) \setminus B$. Für $x \in P$ definiere $\tilde{f}_x : Q \rightarrow \mathbb{R}$ durch $\tilde{f}_x(y) := f_x(y)$ für $y \in B_x$ und $\tilde{f}_x(y) := 0$ für $y \in Q \setminus B_x$. Dann gilt $\tilde{f}_x(y) = \tilde{f}(x, y)$ für alle $x \in P$ und $y \in Q$. Nach Voraussetzung und Definition 5.5.5 sind die Funktionen $\tilde{f} : P \times Q \rightarrow \mathbb{R}$ und $\tilde{f}_x : Q \rightarrow \mathbb{R}$ (für $x \in P$) Riemann-integrierbar, und es gilt

$$\int_B f = \int_{P \times Q} \tilde{f}, \quad \int_{B_x} f_x = \int_Q \tilde{f}_x.$$

Also folgt (ii) aus Satz 5.4.1. Aussage (i) folgt aus (ii) mit $f \equiv 1$. Damit ist Korollar 5.5.7 bewiesen. \square

Beispiel 5.5.8. Dieses Beispiel zeigt, dass aus der Jordan-Messbarkeit von B in Korollar 5.5.7 nicht die Jordan-Messbarkeit von B_x für jedes $x \in P$ folgt. Sei $P = Q = [0, 1]$, sei $K \subset [0, 1]$ die Cantor-Menge, und sei

$$B := \{(x, y) \in [0, 1]^2 \mid x \in K \implies y \in \mathbb{Q}\}.$$

Dann ist $\partial B = K \times [0, 1]$ und daher ist B Jordan-messbar (Beispiel 5.4.2). Jedoch ist $B_x \subset [0, 1]$ nur für $x \in [0, 1] \setminus K$ Jordan-messbar, nicht für $x \in K$.

Beispiel 5.5.9. Dieses Beispiel zeigt, dass aus der Jordan-Messbarkeit der Mengen B_x für alle $x \in P$ nicht die Jordan-Messbarkeit der Menge B folgt. Sei $P = Q = [0, 1]$ und $B := ([0, 1] \cap \mathbb{Q}) \times [0, 1]$. Dann ist die Menge B_x für jedes $x \in [0, 1]$ entweder leer oder gleich $[0, 1]$, und ist daher Jordan-messbar. Jedoch ist $\partial B = [0, 1]^2$ und daher ist B nicht Jordan-messbar.

Beispiel 5.5.10. Die kompakte Menge $B := \{(s, t) \in \mathbb{R}^2 \mid 0 \leq s \leq t \leq 1\}$ ist das oberhalb der Diagonalen gelegene Dreieck im Einheitsquadrat in \mathbb{R}^2 . Diese Menge ist Jordan-messbar und jede stetige Funktion $f : B \rightarrow \mathbb{R}$ ist nach Satz 5.2.2 Riemann-integrierbar. Nach Korollar 5.5.7 erfüllt jede Riemann-integrierbare Funktion $f : B \rightarrow \mathbb{R}$ die Gleichung

$$\int_B f = \int_0^1 \left(\int_0^t f(s, t) ds \right) dt = \int_0^1 \left(\int_s^1 f(s, t) dt \right) ds.$$

Beispiel 5.5.11. Sei $B \subset \mathbb{R}^n$ Jordan-messbar und $h > 0$. Der **Zylinder über B mit Höhe h** ist die Menge $Z_B := B \times [0, h]$. Diese Menge ist Jordan-messbar und hat das Volumen $\text{Vol}_{n+1}(Z_B) = h \text{Vol}_n(B)$.

Beispiel 5.5.12. Sei $B \subset \mathbb{R}^n$ Jordan-messbar und $\lambda > 0$. Dann ist die Menge $\lambda B := \{\lambda x \mid x \in B\}$ Jordan-messbar und $\text{Vol}_n(\lambda B) = \lambda^n \text{Vol}_n(B)$. Die Jordan-Messbarkeit von λB folgt daraus, dass $\partial(\lambda B) = \lambda \partial B$ ist. Die Formel für das Volumen folgt einerseits aus der Transformationsformel in Satz 5.7.11, kann andererseits aber auch direkt aus den Definitionen hergeleitet werden.

Beispiel 5.5.13. Sei $B \subset \mathbb{R}^n$ Jordan-messbar und $h > 0$. Die Menge

$$K_B := \{(x, y) \mid 0 \leq y \leq h, x \in (1 - y/h) B\}.$$

ist der **Kegel über B mit Höhe h** . Diese Menge ist Jordan-messbar, wie wir im nächsten Abschnitt sehen werden (Beispiel 5.6.12). Akzeptiert man die Jordan-Messbarkeit des Kegels K_B , so ergibt sich nach Korollar 5.5.7 (i) und Beispiel 5.5.12 die Formel

$$\text{Vol}_{n+1}(K_B) = \int_0^h \left(1 - \frac{y}{h}\right)^n \text{Vol}_n(B) dy = \frac{h}{n+1} \text{Vol}_n(B) = \frac{\text{Vol}_{n+1}(Z_B)}{n+1}.$$

5.6 Mehr über Jordan-messbare Mengen

Wie wir bereits anhand verschiedener Beispiele gesehen haben, ist es eine wichtige Frage, wie man Jordan-messbare Mengen und Jordansche Nullmengen erkennen kann. Damit beschäftigt sich dieser Abschnitt.

Charakterisierung Jordan-messbarer Mengen

Definition 5.6.1 (Quadergebäude). Eine kompakte Teilmenge $B \subset \mathbb{R}^n$ heißt **Quadergebäude** wenn sie entweder leer ist oder sich als Vereinigung endlich vieler achsenparalleler abgeschlossener Quader (mit positivem Volumen) schreiben läßt. Das heißt, eine nichtleere Teilmenge $B \subset \mathbb{R}^n$ ist genau dann ein Quadergebäude, wenn $\mathcal{P}(B) \neq \emptyset$ ist (siehe Übung 5.1.4).

Satz 5.6.2 (Charakterisierung Jordan-messbarer Mengen).

Sei $B \subset \mathbb{R}^n$ eine beschränkte Menge und $b \geq 0$ eine reelle Zahl. Sei $Q \subset \mathbb{R}^n$ ein abgeschlossener achsenparalleler Quader, der B enthält. Dann sind folgende Aussagen äquivalent.

- (i) B ist Jordan-messbar und hat das Volumen $\text{Vol}_n(B) = b$.
- (ii) Die charakteristische Funktion $1_B : Q \rightarrow \mathbb{R}$ ist Riemann-integrierbar und es gilt $\int_Q 1_B = b$.
- (iii) Für jedes $\varepsilon > 0$ existieren Quadergebäude $B_0, B_1 \subset \mathbb{R}^n$ mit

$$B_0 \subset B \subset B_1, \quad b - \varepsilon < \text{Vol}_n(B_0) \leq \text{Vol}_n(B_1) < b + \varepsilon.$$

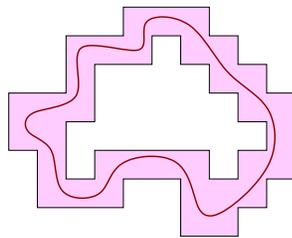


Abbildung 5.3: Jordan-messbare Mengen.

Bemerkung 5.6.3. Die Quadergebäude B_0, B_1 in Teil (iii) von Satz 5.6.2 können so gewählt werden dass alle ihre Extrempunkte (Ecken) Vektoren von der Form $x = (2^{-m}k_1, \dots, 2^{-m}k_n)$ sind mit $m \in \mathbb{N}$ und $k_1, \dots, k_n \in \mathbb{Z}$.

Beweis von Satz 5.6.2. Die Implikation (i) \implies (ii) folgt unmittelbar aus der Definition des Jordanschen Maßes und Satz 5.2.2.

Wir beweisen (ii) \implies (iii). Sei $Q \subset \mathbb{R}^n$ ein abgeschlossener achsenparalleler Quader, der B enthält. Nach Voraussetzung ist die charakteristische Funktion $1_B : Q \rightarrow \mathbb{R}$ Riemann-integrierbar und es ist $\int_Q 1_B = b$. Sei $\varepsilon > 0$. Nach Satz 5.3.1 existiert eine Partition $P = \{P_1, \dots, P_k\} \in \mathcal{P}(Q)$ mit

$$b - \varepsilon < \underline{S}(1_B, P) \leq \overline{S}(1_B, P) < b + \varepsilon.$$

Definiere

$$J_0 := \{j \in \{1, \dots, k\} \mid \overline{P}_j \subset B\}, \quad J_1 := \{j \in \{1, \dots, k\} \mid \overline{P}_j \cap B \neq \emptyset\}.$$

Dann sind die Mengen $B_0 := \bigcup_{j \in J_0} \overline{P}_j$ und $B_1 := \bigcup_{j \in J_1} \overline{P}_j$ Quadergebäude mit $B_0 \subset B \subset B_1$ und

$$b - \varepsilon < \underline{S}(1_B, P) = \text{Vol}_n(B_0) \leq \text{Vol}_n(B_1) = \overline{S}(1_B, P) < b + \varepsilon.$$

Damit ist gezeigt, dass (iii) aus (ii) folgt.

Wir beweisen (iii) \implies (i). Sei $\varepsilon > 0$ gegeben. Nach (iii) existieren zwei Quadergebäude $B_0, B_1 \subset \mathbb{R}^n$ mit $B_0 \subset B \subset B_1$ und

$$b - \frac{\varepsilon}{4} < \text{Vol}_n(B_0) \leq \text{Vol}_n(B_1) < b + \frac{\varepsilon}{4}.$$

Definiere

$$A := B_1 \setminus \overset{\circ}{B}_0 = \overline{B_1} \setminus \overline{B_0}.$$

Dann ist $\partial B \subset A$. Da B_0 ein Quadergebäude ist, stimmt der Rand des Inneren $\overset{\circ}{B}_0$ mit dem Rand von B_0 überein. Daher ist $\overset{\circ}{B}_0$ Jordan-messbar und es ist $\text{Vol}_n(\overset{\circ}{B}_0) = \text{Vol}_n(B_0) = \text{Vol}_n(\overset{\circ}{B}_0)$. Nach Lemma 5.5.4 ist A eine Jordan-messbare Menge und es gilt

$$\int_{B_1} 1_A = \text{Vol}_n(A) = \text{Vol}_n(B_1) - \text{Vol}_n(B_0) < \frac{\varepsilon}{2}.$$

Also existiert eine Partition $P = \{P_1, \dots, P_k\} \in \mathcal{P}(B_1)$ mit $\overline{S}(1_A, P) < \varepsilon/2$. Sei $J := \{j \in \{1, \dots, k\} \mid \overline{P}_j \cap A \neq \emptyset\}$. Für $j \in J$ sei $W_j \subset \mathbb{R}^n$ ein offener Quader mit $\overline{P}_j \subset W_j$ und $\text{Vol}_n(W_j) < 2\text{Vol}_n(P_j)$. Dann gilt $\partial B \subset \bigcup_{j \in J} W_j$ und

$$\sum_{j \in J} \text{Vol}_n(W_j) < 2 \sum_{j \in J} \text{Vol}_n(P_j) = 2\overline{S}(1_A, P) < \varepsilon.$$

Damit haben wir gezeigt, dass ∂B eine Jordansche Nullmenge ist. Also ist B Jordan-messbar. Die Gleichung $\text{Vol}_n(B) = b$ folgt nun unmittelbar aus (iii). Damit ist Satz 5.6.2 bewiesen. \square

Korollar 5.6.4. *Eine beschränkte Teilmenge $B \subset \mathbb{R}^n$ ist genau dann eine Jordansche Nullmenge, wenn sie Jordan-messbar ist und $\text{Vol}_n(B) = 0$ ist.*

Beweis. Sei $B \subset \mathbb{R}^n$ eine Jordan-messbare Menge mit $\text{Vol}_n(B) = 0$. Sei $\varepsilon > 0$. Dann existiert nach Satz 5.6.2 ein Quadergebäude $B_1 \subset \mathbb{R}^n$ mit $B \subset B_1$ und $\text{Vol}_n(B_1) < \varepsilon/2$. Wähle eine Partition $P = \{P_1, \dots, P_k\} \in \mathcal{P}(B_1)$. Dann gilt $B \subset B_1 = \bigcup_{i=1}^k \bar{P}_i$ und $\sum_{i=1}^k \text{Vol}_n(P_i) = \text{Vol}_n(B_1) < \varepsilon/2$. Wähle offene achsenparallelen Quader W_i mit $\bar{P}_i \subset W_i$ und $\text{Vol}_n(W_i) \leq 2\text{Vol}_n(P_i)$. Dann gilt $B \subset \bigcup_i W_i$ und $\sum_i \text{Vol}_n(W_i) < \varepsilon$. Also ist B eine Jordansche Nullmenge.

Ist B eine Jordansche Nullmenge, so ist $\dot{B} = \emptyset$ und \bar{B} eine Jordansche Nullmenge nach Lemma 5.2.3. Daraus folgt, dass $\partial B = \bar{B} \setminus \dot{B} = \bar{B}$ eine Jordansche Nullmenge ist. Außerdem folgt aus Teil (ii) von Satz 5.2.2, dass $\text{Vol}_n(B) = \int_Q 1_B = 0$ ist (für jeden abgeschlossenen Quader Q , der B enthält). Damit ist Korollar 5.6.4 bewiesen. \square

Korollar 5.6.5. *Sei $B \subset \mathbb{R}^n$ eine Jordan-messbare Menge und $f : B \rightarrow \mathbb{R}$ eine Riemann-integrierbare Funktion, so dass $f(x) \geq 0$ ist für alle $x \in B$. Dann ist die Menge*

$$C := \{(x, y) \in \mathbb{R}^n \times \mathbb{R} \mid x \in B, 0 \leq y \leq f(x)\}$$

Jordan-messbar und hat das Volumen

$$\text{Vol}_{n+1}(C) = \int_B f.$$

Beweis. Sei $Q \subset \mathbb{R}^n$ ein abgeschlossener achsenparalleler Quader, der B enthält. Definiere $\tilde{f} : Q \rightarrow \mathbb{R}$ durch $\tilde{f}(x) := f(x)$ für $x \in B$ und $\tilde{f}(x) := 0$ für $x \in Q \setminus B$. Dann ist $\tilde{f} : Q \rightarrow \mathbb{R}$ Riemann-integrierbar und $c := \int_Q \tilde{f} = \int_B f$. Nach Satz 5.3.1 existiert eine Partition $P = \{P_1, \dots, P_k\} \in \mathcal{P}(Q)$ mit

$$c - \varepsilon < \underline{S}(\tilde{f}, P) \leq \bar{S}(\tilde{f}, P) < c + \frac{\varepsilon}{2}.$$

Definiere

$$C_0 := \bigcup_{\inf_{\bar{P}_i} f > 0} \bar{P}_i \times [0, \inf_{\bar{P}_i} f], \quad C_1 := \bigcup_{i=1}^k \bar{P}_i \times \left[0, \sup_{\bar{P}_i} f + \frac{\varepsilon}{2\text{Vol}_n(Q)}\right].$$

Dann sind C_0 und C_1 zwei Quadergebäude in \mathbb{R}^{n+1} mit $C_0 \subset C \subset C_1$ und

$$c - \varepsilon < \underline{S}(\tilde{f}, P) = \text{Vol}_{n+1}(C_0) \leq \text{Vol}_{n+1}(C_1) = \bar{S}(\tilde{f}, P) + \frac{\varepsilon}{2} < c + \varepsilon.$$

Da $\varepsilon > 0$ beliebig gewählt war, folgt aus Satz 5.6.2, dass C Jordan-messbar und $\text{Vol}_{n+1}(C) = c$ ist. Damit ist Korollar 5.6.5 bewiesen. \square

Korollar 5.6.6. *Seien $p, q \in \mathbb{N}$ und $n := p + q$. Seien $A \subset \mathbb{R}^p$ und $B \subset \mathbb{R}^q$ Jordan-messbar. Dann ist $A \times B \subset \mathbb{R}^n$ Jordan-messbar und es gilt*

$$\text{Vol}_n(A \times B) = \text{Vol}_p(A)\text{Vol}_q(B).$$

Beweis. Es folgt direkt aus den Definitionen, dass das Produkt einer Jordanschen Nullmenge in \mathbb{R}^p mit einer beschränkten Teilmenge von \mathbb{R}^q eine Jordansche Nullmenge in \mathbb{R}^n ist. In diesem Fall folgt die Behauptung aus Korollar 5.6.4. Wir dürfen also annehmen, dass weder A noch B eine Jordansche Nullmenge ist. Dann gilt

$$a := \text{Vol}_p(A) > 0, \quad b := \text{Vol}_q(B) > 0.$$

Sei $\varepsilon > 0$ gegeben. Wir nehmen ohne Einschränkung der Allgemeinheit an, dass $\varepsilon < \min\{a, b, 1\}$ ist. Nach Satz 5.6.2 existieren Quadergebäude $A_0, A_1 \subset \mathbb{R}^p$ und $B_0, B_1 \subset \mathbb{R}^q$ mit

$$A_0 \subset A \subset A_1, \quad B_0 \subset B \subset B_1$$

und

$$\begin{aligned} a - \frac{\varepsilon}{2b} < \text{Vol}_p(A_0) \leq \text{Vol}_p(A_1) < a + \frac{\varepsilon}{2(b+1)}, \\ b - \frac{\varepsilon}{2a} < \text{Vol}_q(B_0) \leq \text{Vol}_q(B_1) < b + \frac{\varepsilon}{2(a+1)}. \end{aligned}$$

Dann sind $A_0 \times B_0$ und $A_1 \times B_1$ Quadergebäude in \mathbb{R}^n mit

$$A_0 \times B_0 \subset A \times B \subset A_1 \times B_1$$

und es gilt

$$\begin{aligned} ab - \varepsilon &< \left(a - \frac{\varepsilon}{2b}\right) \left(b - \frac{\varepsilon}{2a}\right) \\ &< \text{Vol}_p(A_0)\text{Vol}_q(B_0) \\ &= \text{Vol}_n(A_0 \times B_0) \\ &\leq \text{Vol}_n(A_1 \times B_1) \\ &= \text{Vol}_p(A_1)\text{Vol}_q(B_1) \\ &< \left(a + \frac{\varepsilon}{2(b+1)}\right) \left(b + \frac{\varepsilon}{2(a+1)}\right) \\ &\leq ab + \varepsilon. \end{aligned}$$

Nach Satz 5.6.2 ist also $A \times B$ Jordan-messbar mit $\text{Vol}_n(A \times B) = ab$. Damit ist Korollar 5.6.6 bewiesen. \square

Jordansche Nullmengen und Lipschitz-Stetigkeit

Teil (iii) des folgenden Satzes kann als Spezialfall des **Satzes von Sard** betrachtet werden (siehe Bemerkung 3.3.13).

Satz 5.6.7 (Jordansche Nullmengen). *Seien $p, q, n \in \mathbb{N}$.*

(i) *Sei $A \subset \mathbb{R}^p \times \mathbb{R}^q$ eine kompakte Menge und sei*

$$A^y := \{x \in \mathbb{R}^p \mid (x, y) \in A\}$$

eine Jordansche Nullmenge in \mathbb{R}^p für alle $y \in \mathbb{R}^q$. Dann ist A eine Jordansche Nullmenge.

(ii) *Sei $U \subset \mathbb{R}^n$ eine offene Menge, sei $A \subset U$ eine kompakte Jordansche Nullmenge, und sei $f : U \rightarrow \mathbb{R}^n$ eine lokal Lipschitz-stetige Abbildung. Dann ist $f(A) \subset \mathbb{R}^n$ ebenfalls eine Jordansche Nullmenge.*

(iii) *Sei $V \subset \mathbb{R}^d$ eine offene Menge, sei $B \subset V$ eine kompakte Teilmenge, und sei $g : V \rightarrow \mathbb{R}^n$ eine lokal Lipschitz-stetige Abbildung mit $d < n$. Dann ist $g(B)$ eine Jordansche Nullmenge.*

Beweis. Wir beweisen (i). Seien $E \subset \mathbb{R}^p$ und $F \subset \mathbb{R}^q$ zwei achsenparallele abgeschlossene Quader so dass

$$A \subset \overset{\circ}{E} \times \overset{\circ}{F}$$

ist und sei $V := \text{Vol}_q(F) > 0$. Sei $\varepsilon > 0$ gegeben und definiere

$$B := \{y \in F \mid A^y \neq \emptyset\}.$$

Dann ist B eine kompakte Teilmenge von $\overset{\circ}{F}$ und für jedes $y \in B$ existiert eine Partition

$$P(y) = \{P_1(y), \dots, P_{k(y)}(y)\} \in \mathcal{P}(E)$$

so dass

$$\sum_{i \in I(y)} \text{Vol}_p(P_i(y)) < \frac{\varepsilon}{V}, \quad I(y) := \left\{ i \in \{1, \dots, k(y)\} \mid \overline{P_i(y)} \cap A^y \neq \emptyset \right\}.$$

Für $y \in B$ definiere

$$J(y) := \{1, \dots, k(y)\} \setminus I(y) = \left\{ i \in \{1, \dots, k(y)\} \mid \overline{P_i(y)} \cap A^y = \emptyset \right\}$$

und

$$U(y) := \left\{ y' \in \mathbb{R}^q \mid A^{y'} \cap \overline{P_i(y)} = \emptyset \text{ für alle } i \in J(y) \right\}.$$

Dann ist $U(y)$ für jedes $y \in B$ eine offene Menge die y enthält. Daher existiert für jedes $y \in B$ ein $r(y) > 0$ mit $B_{2r(y)}(y) \subset U(y) \cap \overset{\circ}{F}$.

Mit dieser Konstruktion bilden die offenen Bälle $B_{r(y)}(y)$ mit $y \in B$ eine offene Überdeckung der kompakten Menge B . Nach Satz C.1.2 existieren daher endlich viele Elemente $y_1, \dots, y_N \in B$ mit

$$B \subset \bigcup_{\nu=1}^N B_{r(y_\nu)}(y_\nu).$$

Sei nun $Q = \{Q_1, \dots, Q_\ell\} \in \mathcal{P}(F)$ eine Partition mit der Feinheit

$$\delta(Q) < \min_{\nu=1, \dots, N} \frac{r(y_\nu)}{\sqrt{q}}.$$

Dann existiert für jedes $j \in \{1, \dots, \ell\}$ mit $\overline{Q_j} \cap B \neq \emptyset$ ein $\nu \in \{1, \dots, N\}$ mit $\overline{Q_j} \subset U(y_\nu)$. (Ist nämlich $\eta \in \overline{Q_j} \cap B$, so existiert ein $\nu \in \{1, \dots, N\}$ mit $\eta \in B_{r(y_\nu)}(y_\nu)$ und es folgt daraus, dass $\overline{Q_j} \subset B_{r(y_\nu)}(\eta) \subset B_{2r(y_\nu)}(y_\nu) \subset U(y_\nu)$ ist.) Sei nun

$$J := \{j \in \{1, \dots, \ell\} \mid \overline{Q_j} \cap B \neq \emptyset\}$$

und wähle eine Abbildung

$$J \rightarrow \{1, \dots, N\} : j \mapsto \nu(j)$$

so dass $\overline{Q_j} \subset U(y_{\nu(j)})$ ist für alle $j \in J$. Dann gilt

$$A^y \subset \bigcup_{i \in I(y_{\nu(j)})} \overline{P_i(y_{\nu(j)})} \quad \text{für alle } j \in J \text{ und alle } y \in \overline{Q_j}$$

und daher

$$A \subset \bigcup_{j \in J} \bigcup_{i \in I(y_{\nu(j)})} \overline{P_i(y_{\nu(j)})} \times \overline{Q_j}.$$

Also ist die Menge A in der Vereinigung endlich vieler abgeschlossen Quader enthalten mit dem Gesamtvolumen

$$\begin{aligned} \sum_{j \in J} \sum_{i \in I(y_{\nu(j)})} \text{Vol}_{p+q}(P_i(y_{\nu(j)}) \times Q_j) &= \sum_{j \in J} \sum_{i \in I(y_{\nu(j)})} \text{Vol}_p(P_i(y_{\nu(j)})) \text{Vol}_q(Q_j) \\ &< \frac{\varepsilon}{V} \sum_{j \in J} \text{Vol}_q(Q_j) \\ &\leq \varepsilon. \end{aligned}$$

Daher ist A eine Jordansche Nullmenge.

Wir beweisen (ii). Da f lokal Lipschitz-stetig ist, ist die Restriktion von f auf jede kompakte Teilmenge von U Lipschitz-stetig. Also ist $f|_A : A \rightarrow \mathbb{R}^n$ Lipschitz-stetig, d.h. es existiert ein $c > 0$ so dass für alle $x, y \in \mathbb{R}^n$ gilt

$$x, y \in A \quad \implies \quad \|f(x) - f(y)\| \leq c \|x - y\|.$$

Hier verwenden wir die Euklidische Norm. Ein **abgeschlossener Würfel** in \mathbb{R}^n ist eine Teilmenge der Form

$$Q = [a_1, a_1 + s] \times [a_2, a_2 + s] \times \cdots \times [a_n, a_n + s]$$

mit $a_1, \dots, a_n \in \mathbb{R}$ und $s > 0$. Die Zahl s heißt **Seitenlänge des Würfels** Q . Sei $\varepsilon > 0$ gegeben. Da A eine Jordansche Nullmenge ist, existieren endlich viele abgeschlossene Würfel $Q_1, \dots, Q_\ell \subset \mathbb{R}^n$, so dass

$$A \subset \bigcup_{j=1}^{\ell} Q_j, \quad \sum_{j=1}^{\ell} \text{Vol}_n(Q_j) \leq \frac{\varepsilon}{(3c\sqrt{n})^n}.$$

Sei $s_j > 0$ die Seitenlänge des Würfels Q_j . Dann ist $\text{Vol}_n(Q_j) = s_j^n$ und es gilt $\|x - y\| \leq \sqrt{n}s_j$ für alle $x, y \in Q_j$. Daraus folgt

$$\|f(x) - f(y)\| \leq c\sqrt{n}s_j \quad \text{für alle } x, y \in Q_j \cap A.$$

Daher existieren offene Würfel $W_j \subset \mathbb{R}^n$, $j = 1, \dots, \ell$, mit der Seitenlänge $t_j := 3c\sqrt{n}s_j$ so dass $f(Q_j \cap A) \subset W_j$ für alle j . Da $A \subset \bigcup_j Q_j$ ist, folgt daraus $f(A) \subset \bigcup_j W_j$ und

$$\sum_j \text{Vol}_n(W_j) = \sum_j (3c\sqrt{n}s_j)^n = (3c\sqrt{n})^n \sum_j \text{Vol}_n(Q_j) < \varepsilon.$$

Also ist $f(A)$ eine Jordansche Nullmenge und damit ist (ii) bewiesen.

Wir beweisen (iii). Sei $V \subset \mathbb{R}^d$ eine offene Menge, sei $B \subset V$ kompakt, und sei $g : V \rightarrow \mathbb{R}^n$ eine lokal Lipschitz-stetige Abbildung. Definiere die Mengen $A \subset U \subset \mathbb{R}^n$ und die Funktion $f : U \rightarrow \mathbb{R}^n$ durch

$$U := V \times \mathbb{R}^{n-d}, \quad A := B \times \{0\}, \quad f(x_1, \dots, x_n) := g(x_1, \dots, x_d).$$

Dann ist U offen, die Teilmenge $A \subset U$ ist eine kompakte Jordansche Nullmenge, und $f : U \rightarrow \mathbb{R}^n$ ist lokal Lipschitz-stetig. Also ist $g(B) = f(A)$ nach Teil (ii) eine Jordansche Nullmenge. Damit ist Satz 5.6.7 bewiesen. \square

Beispiele

Beispiel 5.6.8. Die Voraussetzung, dass A kompakt ist, kann in Teil (i) von Satz 5.6.7 nicht weggelassen werden. Die Menge

$$A := \{(x, y) = 2^{-m}(k, \ell) \in \mathbb{R}^2 \mid m \in \mathbb{N}, k, \ell \in \{1, 3, \dots, 2^m - 1\}\}$$

erfüllt $\bar{A} = \partial A = [0, 1]^2$ und ist daher keine Jordansche Nullmenge, obwohl die Menge $A^y \subset [0, 1]$ für jedes $y \in \mathbb{R}$ nur endlich viele Elemente enthält und daher eine Jordansche Nullmenge ist.

Beispiel 5.6.9. Die Voraussetzung, dass g lokal Lipschitz-stetig ist, kann in Teil (iii) von Satz 5.6.7 nicht weggelassen werden. Ist $g : \mathbb{R} \rightarrow \mathbb{R}^2$ eine stetige Abbildung mit $g([0, 1]) = [0, 1]^2$ (*raumfüllende Kurve*), dann ist $g([0, 1])$ keine Jordansche Nullmenge.

Beispiel 5.6.10. Sei $E \subset \mathbb{R}^n$ ein linearer Unterraum mit

$$d = \dim E < n$$

und sei $a \in \mathbb{R}^n$. Dann ist jede kompakte Teilmenge

$$K \subset a + E$$

des affinen Unterraumes $a + E$ eine Jordansche Nullmenge im \mathbb{R}^n .

Beweis: Sei e_1, \dots, e_d eine Basis von E und sei $g : \mathbb{R}^d \rightarrow \mathbb{R}^n$ die Abbildung $g(x_1, \dots, x_d) := a + \sum_{i=1}^d x_i e_i$. Diese Abbildung ist glatt und hat die Bildmenge $g(\mathbb{R}^d) = a + E$. Da $B := g^{-1}(K)$ kompakt ist, ist $K = g(B)$ nach Teil (iii) von Satz 5.6.7 eine Jordansche Nullmenge.

Beispiel 5.6.11. Der n -Simplex $\Delta^n = \{x \in \mathbb{R}^n \mid x_i \geq 0, \sum_i x_i \leq 1\}$ ist nach Beispiel 5.6.10 Jordan-messbar. (Siehe auch Beispiel 5.4.5.)

Beispiel 5.6.12. Sei $B \subset \mathbb{R}^n$ eine Jordan-messbare Menge, sei $h > 0$, und sei

$$K_B := \{(x, y) \in \mathbb{R}^n \times \mathbb{R} \mid 0 \leq y \leq h, x \in (1 - \frac{y}{h})B\}$$

der Kegel über B wie in Beispiel 5.5.13. Dann ist

$$\partial K_B = K_{\partial B} \cup (\bar{B} \times \{0\})$$

nach Teil (i) von Satz 5.6.7 eine Jordansche Nullmenge. Also ist K_B Jordan-messbar.

Beispiel 5.6.13. Sei $M \subset \mathbb{R}^n$ eine C^1 -Untermannigfaltigkeit der Dimension $d = \dim M < n$ und sei $A \subset M$ eine kompakte Teilmenge. Dann ist A eine Jordansche Nullmenge.

Beweis: Für alle $x \in A$ existieren zwei offene Mengen $U_x, V_x \subset \mathbb{R}^n$ und ein C^1 -Diffeomorphismus $\phi_x : U_x \rightarrow V_x$ mit

$$x \in U_x, \quad \phi_x(U_x \cap M) = V_x \cap (\mathbb{R}^d \times \{0\}).$$

Wähle kompakte Umgebungen $K_x \subset U_x$ von x , eine für jedes $x \in A$. Für jedes $x \in A$ ist dann die Menge $L_x := \phi_x(K_x \cap M)$ kompakt und nach Beispiel 5.6.10 eine Jordansche Nullmenge. Daher ist auch $K_x \cap M = \phi_x^{-1}(L_x)$ eine Jordansche Nullmenge, nach Teil (iii) von Satz 5.6.7. Die Mengen K_x für $x \in A$ bilden eine offene Überdeckung von A . Da A kompakt ist, existieren nach Satz C.1.2 endlich viele Punkte $x_1, \dots, x_\ell \in A$ mit $A \subset \bigcup_{j=1}^{\ell} \overset{\circ}{K}_{x_j}$. Also ist A die endliche Vereinigung der Jordanschen Nullmengen $K_{x_j} \cap A$ für $j = 1, \dots, \ell$ und ist deshalb ebenfalls eine Jordansche Nullmenge.

Beispiel 5.6.14. Eine Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ heißt **eigentlich**, wenn für jede kompakte Teilmenge $K \subset \mathbb{R}$ auch die Menge $f^{-1}(K) = \{x \in \mathbb{R}^n \mid f(x) \in K\}$ kompakt ist. Äquivalent dazu ist die Bedingung, dass für jede Folge $(x_\nu)_{\nu \in \mathbb{N}}$ von Vektoren in \mathbb{R}^n gilt

$$\sup_{\nu} |f(x_\nu)| < \infty \quad \implies \quad (x_\nu) \text{ besitzt eine konvergente Teilfolge.} \quad (5.6.1)$$

Sei nun $f : \mathbb{R}^n \rightarrow [0, \infty)$ stetig differenzierbar und eigentlich. Sei $c > 0$ ein regulärer Wert von f . Dann ist die Menge

$$B := \{x \in \mathbb{R}^n \mid f(x) \leq c\}$$

kompakt. Ihr Rand ist die kompakte Menge $\partial B = \{x \in \mathbb{R}^n \mid f(x) = c\}$. Da c ein regulärer Wert von f ist, ist ∂B eine $(n-1)$ -dimensionale C^1 -Untermannigfaltigkeit von \mathbb{R}^n (Korollar 3.3.12) und ist daher nach Beispiel 5.6.13 eine Jordansche Nullmenge. Also ist B Jordan-messbar.

Beispiel 5.6.15. Der Einheitsball

$$B^n := \{x \in \mathbb{R}^n \mid \|x\| \leq 1\}$$

ist Jordan-messbar nach Beispiel 5.6.14 mit $f(x) = \|x\|^2$, $c = 1$. Für $k \in \mathbb{N}$ beweisen wir die Formeln

$$\text{Vol}_{2k}(B^{2k}) = \frac{\pi^k}{k!}, \quad \text{Vol}_{2k+1}(B^{2k+1}) = \frac{2^{k+1}\pi^k}{3 \cdot 5 \cdots (2k+1)}. \quad (5.6.2)$$

Dazu betrachten wir die obere Halbkugel in B^{n+1} . Der Schnitt durch diese Halbkugel in der Höhe $t \in [0, 1]$ ist der Ball vom Radius $r = \sqrt{1 - t^2}$ im \mathbb{R}^n . Dieser ist nach Beispiel 5.5.12 Jordan-messbar mit $\text{Vol}_n(rB^n) = r^n \text{Vol}_n(B^n)$. Daher folgt aus dem Satz von Fubini (Korollar 5.5.7) die Formel

$$\text{Vol}_{n+1}(B^{n+1}) = 2 \int_0^1 (1 - t^2)^{n/2} dt \cdot \text{Vol}_n(B^n) \quad (5.6.3)$$

Der Faktor

$$c_n := \int_0^1 (1 - t^2)^{n/2} dt, \quad n \in \mathbb{N}_0 = \mathbb{N} \cup \{0\} \quad (5.6.4)$$

läßt sich induktiv wie folgt berechnen.

Es gilt $c_0 = 1$ und, mit der Substitution $t = \sin(\theta)$ für $0 \leq \theta \leq \pi/2$,

$$c_1 = \int_0^1 \sqrt{1 - t^2} dt = \int_0^{\pi/2} \cos^2(\theta) d\theta = \frac{\pi}{4}. \quad (5.6.5)$$

Da $\text{Vol}_1(B^1) = 2$ ist, folgt daraus $\text{Vol}_2(B^2) = 2c_1 \text{Vol}_1(B^1) = \pi$ nach (5.6.3). Außerdem gilt $\frac{d}{dt} t(1 - t^2)^{n/2} = (n + 1)(1 - t^2)^{n/2} - n(1 - t^2)^{(n-2)/2}$. Daraus folgt durch Integration die Gleichung $0 = (n + 1)c_n - nc_{n-2}$ und daher

$$c_n = \frac{n}{n + 1} c_{n-2}$$

für alle $n \in \mathbb{N}$ mit $n \geq 2$. Mit $c_0 = 1$ und $c_1 = \pi/4$ erhalten wir daher

$$c_{2k} = \frac{2k}{2k + 1} \frac{2k - 2}{2k - 1} \cdots \frac{2}{3}, \quad c_{2k+1} = \frac{2k + 1}{2k + 2} \frac{2k - 1}{2k} \cdots \frac{3}{4} \pi$$

für $k \in \mathbb{N}$. Diese Formeln lassen sich wiederum in der Form

$$c_{2k} = \frac{2^k k!}{3 \cdot 5 \cdot 7 \cdots (2k + 1)}, \quad c_{2k+1} = \frac{3 \cdot 5 \cdot 7 \cdots (2k + 1)}{2^{k+2} (k + 1)!} \pi$$

schreiben. Nach (5.6.3) und (5.6.4) folgen daraus die Gleichungen

$$\begin{aligned} \frac{\text{Vol}_{2k+2}(B^{2k+2})}{\text{Vol}_{2k}(B^{2k})} &= 4c_{2k}c_{2k+1} = \frac{\pi}{k + 1}, \\ \frac{\text{Vol}_{2k+1}(B^{2k+1})}{\text{Vol}_{2k-1}(B^{2k-1})} &= 4c_{2k-1}c_{2k} = \frac{2\pi}{2k + 1}. \end{aligned}$$

Mit $\text{Vol}_1(B^1) = 2$ und $\text{Vol}_2(B^2) = \pi$ folgt daraus (5.6.2).

5.7 Die Transformationsformel

Die Transformationsformel (Satz 5.7.11) ist die natürliche Erweiterung der Substitutionsregel auf Funktionen von mehreren Variablen. Sie ist von grundlegender Bedeutung für das Riemannsche Integral und ihr Beweis ist der aufwendigste in diesem Kapitel. Dies ist in scharfem Gegensatz zum Fall $n = 1$, wo sich die Substitutionsregel leicht aus dem Hauptsatz der Infinitesimalrechnung herleiten läßt. Der folgende vorbereitende Teilabschnitt erweitert den Begriff einer Partition, indem Quader durch Jordan-messbare Mengen ersetzt werden. Die Konvergenz der entsprechend verallgemeinerten Riemannschen Summen wird im Beweis der Transformationsformel benötigt.

Verallgemeinerte Riemannsche Summen

Definition 5.7.1 (Jordan-Partition). Sei $B \subset \mathbb{R}^n$ eine Jordan-messbare Menge. Eine **Jordan-Partition** von B ist eine Menge $Z = \{B_1, \dots, B_k\}$ von kompakten Jordan-messbaren Mengen $B_i \subset B$, so dass $B = \bigcup_{i=1}^k B_i$ ist und, für alle $i, j \in \{1, \dots, k\}$ mit $i \neq j$, $B_i \cap B_j$ eine Jordansche Nullmenge ist. Die Menge aller Jordan-Partitionen von B wird mit $\mathcal{P}_{\text{Jord}}(B)$ bezeichnet.

Bemerkung 5.7.2. Sei $B \subset \mathbb{R}^n$ eine kompakte Jordan-messbare Menge und sei $f \in \mathcal{R}(B)$. Sei $Z = \{B_1, \dots, B_k\}$ eine Jordan-Partition von B . Dann gilt

$$\int_B f = \sum_{i=1}^k \int_{B_i} f, \quad \text{Vol}_n(B) = \sum_{i=1}^k \text{Vol}_n(B_i).$$

Die erste Gleichung folgt aus Lemma 5.5.4 per vollständiger Induktion, und die zweite Gleichung folgt aus der ersten mit $f \equiv 1$.

Wir verwenden nun die Norm $\|x\|_\infty := \max_i |x_i|$ für $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ und bezeichnen mit $d_\infty(x, y) := \|x - y\|_\infty$ die zugehörige Abstandsfunktion $d_\infty : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$. Der **Durchmesser einer Teilmenge** $B \subset \mathbb{R}^n$ bezüglich dieser Metrik ist $\delta(B) := \sup_{x, y \in B} d_\infty(x, y)$. Der Durchmesser von B ist genau dann endlich wenn B beschränkt ist. Der **Abstand zweier nicht-leerer Teilmengen** $A, B \subset \mathbb{R}^n$ ist durch $d_\infty(A, B) := \inf_{x \in A, y \in B} d_\infty(x, y)$ definiert. Dieser Abstand kann durchaus Null sein, ohne dass die Mengen A und B sich auch nur schneiden, geschweige denn übereinstimmen.

Definition 5.7.3 (Feinheit). Sei $B \subset \mathbb{R}^n$ eine Jordan-messbare Menge. Die **Feinheit** einer Jordan-Partition $Z = \{B_1, \dots, B_k\} \in \mathcal{P}_{\text{Jord}}(B)$ ist die Zahl $\delta(Z) := \max_j \delta(B_j)$ (der maximale Durchmesser der B_j bezüglich d_∞).

Lemma 5.7.4. *Sei $B \subset \mathbb{R}^n$ eine Jordan-messbare Menge und $\Gamma \subset B$ eine Jordansche Nullmenge. Sei $\varepsilon > 0$ gegeben. Dann existiert eine Zahl $\delta_0 > 0$, so dass für jede Jordan-Partition $Z = \{B_1, \dots, B_k\} \in \mathcal{P}_{\text{Jord}}(B)$ gilt*

$$\delta(Z) < \delta_0 \quad \implies \quad \sum_{d_\infty(B_i, \Gamma) \leq \delta_0} \text{Vol}_n(B_i) < \varepsilon.$$

Beweis. Wähle offene Quader $W_1, \dots, W_N \subset \mathbb{R}^n$, so dass

$$\bar{\Gamma} \subset \bigcup_{\nu=1}^N W_\nu, \quad \sum_{\nu=1}^N \text{Vol}_n(W_\nu) < \varepsilon.$$

Man kann dann wie in Schritt 1 im Beweis von Lemma 5.2.8 zeigen, dass eine Zahl $\rho > 0$ existiert, so dass für jeden Punkt $x \in \bar{\Gamma}$, der Ball vom Radius ρ mit Mittelpunkt x bezüglich der Metrik d_∞ ganz in einer der Mengen W_1, \dots, W_N enthalten ist, d.h. $\exists \rho > 0 \forall x \in \bar{\Gamma} \exists \nu \in \{1, \dots, N\} \forall \xi \in \mathbb{R}^n$:

$$d_\infty(x, \xi) < \rho \quad \implies \quad \xi \in W_\nu. \quad (5.7.1)$$

Sei nun $Z = \{B_1, \dots, B_k\} \in \mathcal{P}_{\text{Jord}}(B)$ mit $\delta(Z) < \rho/2$. Definiere

$$\begin{aligned} J_\Gamma &:= \{j \in \{1, \dots, k\} \mid d_\infty(B_j, \bar{\Gamma}) \leq \rho/2\}, \\ J_\nu &:= \{j \in \{1, \dots, k\} \mid B_j \subset W_\nu\}, \quad \nu = 1, \dots, N. \end{aligned}$$

Dann gilt

$$J_\Gamma \subset \bigcup_{\nu=1}^N J_\nu. \quad (5.7.2)$$

Sei nämlich $j \in J_\Gamma$. Dann existieren zwei Elemente $x \in \bar{\Gamma}$ und $y \in B_j$, so dass $d_\infty(x, y) \leq \rho/2$ ist. (Hier verwenden wir die Tatsache, dass $\bar{\Gamma}$ und B_j kompakte Mengen sind.) Für dieses x existiert nun ein $\nu \in \{1, \dots, N\}$, so dass (5.7.1) gilt. Daraus folgt $B_j \subset W_\nu$, denn für $\xi \in B_j$ gilt die Ungleichung $d_\infty(x, \xi) \leq d_\infty(x, y) + d_\infty(y, \xi) < \rho/2 + \delta(B_j) < \rho$ und daher ist $\xi \in W_\nu$. Daraus folgt $j \in J_\nu$. Damit ist (5.7.2) bewiesen. Aus (5.7.2) folgt die Ungleichung

$$\begin{aligned} \sum_{j \in J_\Gamma} \text{Vol}_n(B_j) &\leq \sum_{\nu=1}^N \sum_{j \in J_\nu} \text{Vol}_n(B_j) = \sum_{\nu=1}^N \text{Vol}_n \left(\bigcup_{j \in J_\nu} B_j \right) \\ &\leq \sum_{\nu=1}^N \text{Vol}_n(W_\nu) < \varepsilon. \end{aligned}$$

Damit ist Lemma 5.7.4 bewiesen. \square

Sei nun $B \subset \mathbb{R}^n$ eine Jordan-messbare Menge und $Z = \{B_1, \dots, B_k\}$ eine Jordan-Partition von B . Sei $f : B \rightarrow \mathbb{R}$ eine beschränkte Funktion. Genau wie in Definition 5.1.8 definieren wir die **Obersumme** und die **Untersumme** von f und Z durch

$$\overline{S}(f, Z) := \sum_{j=1}^k (\sup_{B_j} f) \text{Vol}_n(B_j), \quad \underline{S}(f, Z) := \sum_{j=1}^k (\inf_{B_j} f) \text{Vol}_n(B_j).$$

Ebenso wie in Lemma 5.1.9 beweist man die Ungleichung

$$\sup_{Z \in \mathcal{P}_{\text{Jord}}(B)} \underline{S}(f, Z) \leq \inf_{Z \in \mathcal{P}_{\text{Jord}}(B)} \overline{S}(f, Z). \quad (5.7.3)$$

Der nächste Satz zeigt, dass eine beschränkte Funktion $f : B \rightarrow \mathbb{R}$ genau dann Riemann-integrierbar ist, wenn in (5.7.3) Gleichheit herrscht.

Satz 5.7.5. *Sei $B \subset \mathbb{R}^n$ eine Jordan-messbare Menge, $f : B \rightarrow \mathbb{R}$ eine beschränkte Funktion, und $c \in \mathbb{R}$. Dann sind folgende Aussagen äquivalent.*

- (i) f ist Riemann-integrierbar und $c = \int_B f$.
- (ii) Für jedes $\varepsilon > 0$ existiert ein $Z \in \mathcal{P}_{\text{Jord}}(B)$, so dass gilt

$$c - \varepsilon < \underline{S}(f, Z) \leq \overline{S}(f, Z) < c + \varepsilon.$$

- (iii) Für jedes $\varepsilon > 0$ existiert ein $\delta_0 > 0$, so dass für jedes $Z \in \mathcal{P}_{\text{Jord}}(B)$ gilt

$$\delta(Z) < \delta_0 \quad \implies \quad c - \varepsilon < \underline{S}(f, Z) \leq \overline{S}(f, Z) < c + \varepsilon. \quad (5.7.4)$$

- (iv) Für jedes $\varepsilon > 0$ existiert eine Zahl $\delta_0 > 0$ so dass für jede Jordan-Partition $P = \{B_1, \dots, B_k\} \in \mathcal{P}_{\text{Jord}}(B)$ und alle $x_1, \dots, x_k \in \mathbb{R}^n$ gilt

$$\delta(Z) < \delta_0 \quad \implies \quad \left| c - \sum_{i=1}^k f(x_i) \text{Vol}_n(B_i) \right| < \varepsilon. \quad (5.7.5)$$

Die Zahlen $\sum_{i=1}^k f(x_i) \text{Vol}_n(B_i)$ heißen (**verallgemeinerte**) **Riemannsche Summen** von f .

Proof. Siehe Seite 203. □

Zum Beweis von Satz 5.7.5 benötigen wir das folgende Lemma welches die Ober- und Untersummen von f und Z mit den Ober- und Untersummen aus Definition 5.1.8 vergleicht.

Lemma 5.7.6. Seien B , f , c wie in Satz 5.3.1 und sei $Q \subset \mathbb{R}^n$ ein abgeschlossener achsenparalleler Quader, der B enthält. Definiere $\tilde{f} : Q \rightarrow \mathbb{R}$ durch

$$\tilde{f}(x) := \begin{cases} f(x), & \text{für } x \in B, \\ 0, & \text{für } x \in Q \setminus B. \end{cases}$$

Sei $P = \{P_1, \dots, P_k\}$ eine Partition von Q und definiere Z_P und $\varepsilon_P > 0$ durch

$$Z_P := \{B_1, \dots, B_k\}, \quad B_i := \bar{P}_i \cap B,$$

und

$$\varepsilon_P := \sum_{\bar{P}_i \cap \partial B \neq \emptyset} \text{Vol}_n(P_i).$$

Dann ist $Z_P \in \mathcal{P}_{\text{Jord}}(B)$ und es gilt

$$\left| \bar{S}(\tilde{f}, P) - \bar{S}(f, Z) \right| \leq 2 \|f\| \varepsilon_P, \quad \left| \underline{S}(\tilde{f}, P) - \underline{S}(f, Z) \right| \leq 2 \|f\| \varepsilon_P. \quad (5.7.6)$$

Hier bezeichnet $\|f\| := \sup_B |f|$ die Supremumsnorm von f .

Beweis. Definiere die Indexmengen

$$I_0 := \{i \in \{1, \dots, k\} \mid \bar{P}_i \subset \overset{\circ}{B}\}, \quad I_1 := \{i \in \{1, \dots, k\} \mid \bar{P}_i \cap \partial B \neq \emptyset\}.$$

Dann gilt

$$\begin{aligned} \bar{S}(\tilde{f}, P) &= \sum_{i=1}^k \left(\sup_{\bar{P}_i} \tilde{f} \right) \text{Vol}_n(P_i) \\ &= \sum_{i \in I_0} \left(\sup_{B_i} f \right) \text{Vol}_n(B_i) + \sum_{i \in I_1} \left(\sup_{\bar{P}_i} \tilde{f} \right) \text{Vol}_n(P_i) \\ &= \sum_{i=1}^k \left(\sup_{B_i} f \right) \text{Vol}_n(B_i) - \sum_{i \in I_1} \left(\sup_{B_i} f \right) \text{Vol}_n(B_i) + \sum_{i \in I_1} \left(\sup_{\bar{P}_i} \tilde{f} \right) \text{Vol}_n(P_i) \\ &= \bar{S}(f, Z_P) - \sum_{i \in I_1} \left(\sup_{B_i} f \right) \text{Vol}_n(B_i) + \sum_{i \in I_1} \left(\sup_{\bar{P}_i} \tilde{f} \right) \text{Vol}_n(P_i). \end{aligned}$$

Die letzten beiden Summanden sind im Betrag kleiner als $\varepsilon_P \|f\|$. Daraus folgt die erste Ungleichung in (5.7.6). Die zweite Ungleichung beweist man genauso. Damit ist Lemma 5.7.6 bewiesen. \square

Beweis von Satz 5.7.5. Wir beweisen (i) \implies (ii). Wir nehmen ohne Einschränkung der Allgemeinheit an, dass f nicht identisch verschwindet. Seien Q und f wie in Lemma 5.7.6. Wähle eine Partition $P \in \mathcal{P}(Q)$, so dass

$$\varepsilon_P < \frac{\varepsilon}{4\|f\|}, \quad c - \frac{\varepsilon}{2} < \underline{S}(\tilde{f}, P) \leq \overline{S}(\tilde{f}, P) < c + \frac{\varepsilon}{2}.$$

Die Existenz einer solchen Partition folgt aus Lemma 5.2.8 und Satz 5.3.1. Nach Lemma 5.7.6 folgt daraus

$$\overline{S}(f, Z_P) \leq \overline{S}(\tilde{f}, P) + 2\|f\|\varepsilon_P < c + \frac{\varepsilon}{2} + 2\|f\|\varepsilon_P < c + \varepsilon,$$

$$\underline{S}(f, Z_P) \geq \underline{S}(\tilde{f}, P) - 2\|f\|\varepsilon_P > c - \frac{\varepsilon}{2} - 2\|f\|\varepsilon_P > c - \varepsilon.$$

Damit ist gezeigt, dass (ii) aus (i) folgt.

Wir beweisen (ii) \implies (iii). Sei $\varepsilon > 0$ gegeben. Nach (ii) existiert eine Jordan-Partition $Y = \{A_1, \dots, A_k\} \in \mathcal{P}_{\text{Jord}}(B)$ mit

$$c - \varepsilon/2 < \underline{S}(f, Y) \leq \overline{S}(f, Y) < c + \varepsilon/2.$$

Dann ist

$$\Gamma := \bigcup_{i=1}^k \partial A_i$$

eine Jordansche Nullmenge. Also existiert nach Lemma 5.7.4 ein $\delta_0 > 0$, so dass für alle $Z = \{B_1, \dots, B_\ell\} \in \mathcal{P}_{\text{Jord}}(B)$ gilt

$$\delta(Z) < \delta_0 \quad \implies \quad \sum_{d_\infty(B_j, \Gamma) \leq \delta_0} \text{Vol}_n(B_j) < \frac{\varepsilon}{4\|f\|}.$$

Behauptung. Sei $Z = \{B_1, \dots, B_\ell\} \in \mathcal{P}_{\text{Jord}}(B)$ mit $\delta(Z) < \delta_0$ und sei $j \in \{1, \dots, \ell\}$ mit $d_\infty(B_j, \Gamma) > \delta_0$. Dann existiert ein $i \in \{1, \dots, k\}$ mit

$$B_j \subset A_i.$$

Da $B_j \subset B = \bigcup_i A_i$ ist, existiert ein Index $i \in \{1, \dots, k\}$ mit $A_i \cap B_j \neq \emptyset$. Sei $a \in A_i \cap B_j$. Dann ist $d_\infty(a, \partial A_i) > \delta_0$ und daher

$$B_j \subset \overline{B}_{\delta_0}(a; d_\infty) = \{x \in \mathbb{R}^n \mid d_\infty(a, x) \leq \delta_0\} \subset A_i.$$

Damit ist die Behauptung bewiesen.

Sei nun $Z = \{B_1, \dots, B_\ell\} \in \mathcal{P}_{\text{Jord}}(B)$ mit $\delta(Z) < \delta_0$. Definiere

$$\begin{aligned} J_0 &:= \{j \in \{1, \dots, \ell\} \mid d_\infty(B_j, \Gamma) \leq \delta_0\}, \\ J_1 &:= \{j \in \{1, \dots, \ell\} \mid d_\infty(B_j, \Gamma) > \delta_0\}. \end{aligned}$$

Außerdem sei

$$Y \wedge Z := \{A_i \cap B_j \mid i = 1, \dots, k, j = 1, \dots, \ell\}.$$

Dann ist $Y \wedge Z \in \mathcal{P}_{\text{Jord}}(B)$ und

$$\begin{aligned} c - \frac{\varepsilon}{2} &< \underline{S}(f, Y) \\ &\leq \underline{S}(f, Y \wedge Z) \\ &= \sum_{i=1}^k \sum_{j=1}^{\ell} \left(\inf_{A_i \cap B_j} f \right) \text{Vol}_n(A_i \cap B_j) \\ &= \sum_{j \in J_0} \sum_{i=1}^k \left(\inf_{A_i \cap B_j} f \right) \text{Vol}_n(A_i \cap B_j) + \sum_{j \in J_1} \sum_{i=1}^k \left(\inf_{A_i \cap B_j} f \right) \text{Vol}_n(A_i \cap B_j) \\ &= \sum_{j \in J_0} \sum_{i=1}^k \left(\inf_{A_i \cap B_j} f \right) \text{Vol}_n(A_i \cap B_j) + \sum_{j \in J_1} \left(\inf_{B_j} f \right) \text{Vol}_n(B_j) \\ &= \sum_{j \in J_0} \sum_{i=1}^k \left(\inf_{A_i \cap B_j} f \right) \text{Vol}_n(A_i \cap B_j) - \sum_{j \in J_0} \left(\inf_{B_j} f \right) \text{Vol}_n(B_j) + \underline{S}(f, Z) \\ &\leq 2 \|f\| \sum_{j \in J_0} \text{Vol}_n(B_j) + \underline{S}(f, Z) \\ &< \frac{\varepsilon}{2} + \underline{S}(f, Z). \end{aligned}$$

Hier haben wir im fünften Schritt die Behauptung verwendet. Damit gilt

$$\underline{S}(f, Z) > c - \varepsilon$$

und ebenso

$$\overline{S}(f, Z) < c + \varepsilon$$

für jedes $Z \in \mathcal{P}_{\text{Jord}}(B)$ mit $\delta(Z) < \delta_0$. Damit haben wir gezeigt, dass (iii) aus (ii) folgt.

Wir beweisen (iii) \implies (i). Seien Z und \tilde{f} wie in Lemma 5.7.6. Wir können ohne Einschränkung der Allgemeinheit annehmen, dass f nicht identisch verschwindet. Sei $\varepsilon > 0$ gegeben. Nach (iii) existiert eine Konstante $\delta_0 > 0$ so dass für alle $Z \in \mathcal{P}_{\text{Jord}}(B)$ gilt

$$\delta(Z) < \delta_0 \quad \implies \quad c - \varepsilon/2 < \underline{S}(f, Z) \leq \overline{S}(f, Z) < c + \varepsilon/2.$$

Nach Lemma 5.2.8 können wir δ_0 so klein wählen, dass für alle $P \in \mathcal{P}(Q)$ gilt

$$\delta(P) < \delta_0 \quad \implies \quad \varepsilon_P < \frac{\varepsilon}{4\|f\|}.$$

Sei nun $P \in \mathcal{P}(Q)$ mit $\delta(P) < \delta_0$. Sei $Z_P \in \mathcal{P}_{\text{Jord}}(B)$ wie in Lemma 5.7.6. Dann gilt $\delta(Z_P) < \delta_0$ und daraus folgt

$$\begin{aligned} c - \varepsilon &< c - \varepsilon/2 - 2\|f\|\varepsilon_P \\ &< \underline{S}(f, Z_P) - 2\|f\|\varepsilon_P \\ &\leq \underline{S}(\tilde{f}, P) \\ &\leq \overline{S}(\tilde{f}, P) \\ &\leq \overline{S}(f, Z_P) + 2\|f\|\varepsilon_P \\ &< c + \varepsilon/2 + 2\|f\|\varepsilon_P \\ &< c + \varepsilon. \end{aligned}$$

Da $\varepsilon > 0$ beliebig gewählt war, folgt aus Satz 5.3.1, dass $\tilde{f} : Q \rightarrow \mathbb{R}$ Riemann-integrierbar und $\int_Q \tilde{f} = c$ ist. Daraus folgt nach Definition, dass die Funktion $f : B \rightarrow \mathbb{R}$ Riemann-integrierbar und $\int_B f = c$ ist.

Damit haben wir gezeigt, dass die Aussagen (i), (ii) und (iii) zueinander äquivalent sind. Die Äquivalenz der Aussagen (iii) und (iv) beweist man genau wie in Satz 5.3.1. Damit ist Satz 5.7.5 bewiesen. \square

Übung 5.7.7. (i) Jede beschränkte zusammenhängende Teilmenge von \mathbb{R} ist Jordan-messbar.

(ii) Es gibt eine kompakte Teilmenge $K \subset \mathbb{R}$, die nicht Jordan-messbar ist. **Hinweis:** Modifizieren Sie die Konstruktion der Cantor-Menge.

(ii) Es gibt eine kompakte zusammenhängende Teilmenge von \mathbb{R}^2 , die nicht Jordan-messbar ist.

Lineare Transformationen

Satz 5.7.8. Sei $B \subset \mathbb{R}^n$ Jordan-messbar und $\Phi \in \mathbb{R}^{n \times n}$, $v \in \mathbb{R}^n$. Dann ist die Menge $\Phi B + v := \{\Phi x + v \mid x \in B\}$ Jordan-messbar und hat das Jordan-Maß $\text{Vol}_n(\Phi B + v) = |\det(\Phi)| \text{Vol}_n(B)$.

Proof. Siehe Seite 207. □

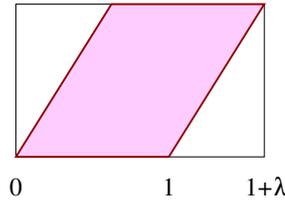


Abbildung 5.4: Die Scherung eines Quadrats.

Beispiel 5.7.9. Sei $\lambda > 0$ und

$$B := [0, 1]^2, \quad \Phi := \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}.$$

Dann ist $A := \Phi B = \{(x, y) \in \mathbb{R}^2 \mid 0 \leq y \leq 1, \lambda y \leq x \leq \lambda y + 1\}$. (Siehe Abbildung 5.4.) Für $y \in [0, 1]$ sei $A^y := \{x \in \mathbb{R} \mid (x, y) \in A\} = [\lambda y, \lambda y + 1]$. Dann ist A^y Jordan-messbar und $\text{Vol}_1(A^y) = 1$ für jedes $y \in [0, 1]$. Außerdem ist $\partial A = \Phi \partial B$ nach Satz 5.6.7 eine Jordansche Nullmenge und daher ist A Jordan-messbar. Also gilt nach Korollar 5.5.7

$$\text{Vol}_2(B\Phi) = \int_0^1 \text{Vol}_1(A^y) dy = 1 = |\det(\Phi)| \text{Vol}_2(B).$$

Beispiel 5.7.10. Seien a_1, \dots, a_n positive reelle Zahlen. Der Ellipsoid

$$E(a_1, \dots, a_n) := \left\{ x \in \mathbb{R}^n \mid \sum_{i=1}^n \frac{x_i^2}{a_i^2} \leq 1 \right\}$$

ist das Bild des Einheitsballes $B := \{x \in \mathbb{R}^n \mid \sum_{i=1}^n x_i^2 \leq 1\}$ unter der Diagonalmatrix $\Phi := \text{diag}(a_1, \dots, a_n)$. Nach Satz 5.7.8 ist $E(a_1, \dots, a_n)$ also Jordan-messbar und hat das Volumen

$$\text{Vol}_n(E(a_1, \dots, a_n)) = a_1 \cdots a_n \text{Vol}_n(B).$$

Hier haben wir die Formel $\det(\Phi) = a_1 \cdots a_n$ verwendet. Wir werden später eine Formel für das Volumen des Einheitsballes herleiten (Beispiel 5.6.15).

Beweis von Satz 5.7.8. Der Beweis hat sieben Schritte.

Schritt 1. Sei $\Phi \in \mathbb{R}^{n \times n}$ mit $\det(\Phi) = 0$ und $v \in \mathbb{R}^n$. Sei $B \subset \mathbb{R}^n$ eine beschränkte Teilmenge. Dann ist $\Phi B + v$ eine Jordansche Nullmenge.

Die Bildmenge $E := \text{im } \Phi = \{\Phi x \mid x \in \mathbb{R}^n\}$ von Φ ist ein linearer Unterraum von \mathbb{R}^n der Dimension $d := \dim E < n$ und $\Phi(B)$ ist eine beschränkte Teilmenge von E . Also ist $\Phi(B)$ nach Beispiel 5.6.10 eine Jordansche Nullmenge.

Schritt 2. Sei $\Phi \in \mathbb{R}^{n \times n}$ und $v \in \mathbb{R}^n$. Ist $B \subset \mathbb{R}^n$ Jordan-messbar, so ist auch $\Phi B + v$ Jordan-messbar.

Ist $\det(\Phi) = 0$, so folgt die Jordan-Messbarkeit von $\Phi B + v$ aus Schritt 1. Ist $\det(\Phi) \neq 0$ so ist die durch $\phi(x) := \Phi x + v$ für $x \in \mathbb{R}^n$ definierte Abbildung $\phi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ein Diffeomorphismus. Daher ist der Rand der Menge $\Phi B + v = \phi(B)$ die Menge $\partial(\Phi B + v) = \phi(\partial B) = \Phi \partial B + v$. Dies ist nach Satz 5.6.7 eine Jordansche Nullmenge. Daher ist $\Phi B + v$ Jordan-messbar.

Schritt 3. Sei $B \subset \mathbb{R}^n$ eine Jordan-messbare Menge und $v \in \mathbb{R}^n$ und $s \geq 0$. Dann gilt $\text{Vol}_n(sB + v) = s^n \text{Vol}_n(B)$.

Ist B ein offener oder abgeschlossener Quader, so folgt die Behauptung unmittelbar aus den Definitionen. Ist B ein Quadergebäude, wähle eine Partition $P = \{P_1, \dots, P_k\} \in \mathcal{P}(B)$. Dann ist $\text{Vol}_n(B) = \sum_{i=1}^k \text{Vol}_n(P_i)$, nach Bemerkung 5.5.3 (v). Außerdem ist

$$sP + v := \{sP_1 + v, \dots, sP_k + v\}$$

eine Partition von $sB + v$. Daraus folgt die Behauptung für Quadergebäude.

Sei nun $B \subset \mathbb{R}^n$ eine beliebige Jordan-messbare Menge und $b := \text{Vol}_n(B)$. Ist $s = 0$ so ist $sB + v = \{v\}$ eine Jordansche Nullmenge und erfüllt daher die Behauptung von Schritt 3. Sei also $s > 0$ und wähle eine Konstante $\varepsilon > 0$. Dann existieren nach Satz 5.6.2 zwei Quadergebäude $B_0, B_1 \subset \mathbb{R}^n$, so dass

$$B_0 \subset B \subset B_1, \quad b - \frac{\varepsilon}{s^n} < \text{Vol}_n(B_0) \leq \text{Vol}_n(B_1) < b + \frac{\varepsilon}{s^n}.$$

Daraus folgt $sB_0 + v \subset sB + v \subset sB_1 + v$ und

$$s^n b - \varepsilon < \text{Vol}_n(sB_0 + v) \leq \text{Vol}_n(sB + v) \leq \text{Vol}_n(sB_1 + v) < s^n b + \varepsilon.$$

Da $\varepsilon > 0$ beliebig gewählt war, folgt die Behauptung aus Satz 5.6.2.

Schritt 4. Sei $\Phi \in \mathbb{R}^{n \times n}$ und $v \in \mathbb{R}^n$. Sei $W_0 := [0, 1]^n$ der abgeschlossene Einheitswürfel. Definiere die Zahl $\lambda(\Phi) \geq 0$ durch

$$\lambda(\Phi) := \text{Vol}_n(\Phi W_0).$$

Dann gilt

$$\text{Vol}_n(\Phi B + v) = \lambda(\Phi) \text{Vol}_n(B) \quad (5.7.7)$$

für jede Jordan-messbare Menge $B \subset \mathbb{R}^n$.

Daß ΦB für jede Jordan-messbare Sei $a = (a_1, \dots, a_n) \in \mathbb{R}^n$ und $s > 0$ und

$$B := [a_1, a_1 + s] \times \cdots \times [a_n, a_n + s] = sW_0 + a.$$

Dann ist $\text{Vol}_n(B) = s^n$ und daher folgt aus Schritt 3, dass

$$\text{Vol}_n(\Phi B) = \text{Vol}_n(s\Phi W_0 + sa) = s^n \text{Vol}_n(\Phi W_0) = \lambda(\Phi) \text{Vol}_n(B).$$

Damit ist Schritt 4 für abgeschlossene Würfel bewiesen.

Sei nun $B \subset \mathbb{R}^n$ ein Quadergebäude, das eine Partition $P = \{P_1, \dots, P_k\}$ besitzt, so dass \bar{P}_i für jedes i ein abgeschlossener Würfel ist. Dann bilden die Mengen $\Phi \bar{P}_i + v$, $i = 1, \dots, k$, eine Jordan-Partition von $\Phi B + v$. Nach dem bisher bewiesenen und Bemerkung 5.7.2 gilt daher

$$\text{Vol}_n(\Phi B + v) = \sum_{i=1}^k \text{Vol}_n(\Phi \bar{P}_i + v) = \lambda(\Phi) \sum_{i=1}^k \text{Vol}_n(P_i) = \lambda(\Phi) \text{Vol}_n(B).$$

Damit ist Schritt 4 für jedes Quadergebäude, das eine Partition aus Würfeln besitzt, bewiesen.

Sei nun $B \subset \mathbb{R}^n$ eine beliebige Jordan-messbare Menge und $b := \text{Vol}_n(B)$. Sei $\varepsilon > 0$. Dann existieren nach Satz 5.6.2 zwei Quadergebäude $B_0, B_1 \subset \mathbb{R}^n$, so dass $B_0 \subset B \subset B_1$ ist und

$$b - \varepsilon \leq \text{Vol}_n(B_0) \leq \text{Vol}_n(B_1) \leq b + \varepsilon.$$

Nach Bemerkung 5.6.3 können B_0 und B_1 so gewählt werden, dass sie Partitionen aus Würfeln besitzen. Damit folgt aus dem bisher bewiesenen, dass

$$\lambda(\Phi)(b - \varepsilon) \leq \text{Vol}_n(\Phi B_0 + v) \leq \text{Vol}_n(\Phi B_1 + v) \leq \lambda(\Phi)(b + \varepsilon).$$

Daraus folgt $\lambda(\Phi)(b - \varepsilon) \leq \text{Vol}_n(\Phi B + v) \leq \lambda(\Phi)(b + \varepsilon)$ für alle $\varepsilon > 0$ und daher $\text{Vol}_n(\Phi B + v) = \lambda(\Phi)b$. Damit ist Schritt 4 bewiesen.

Schritt 5. Die in Schritt 4 definierte Funktion $\lambda : \mathbb{R}^{n \times n} \rightarrow [0, \infty)$ hat folgende Eigenschaften.

(a) Ist $\Phi \in \mathbb{R}^{n \times n}$ mit $\det(\Phi) = 0$ so gilt $\lambda(\Phi) = 0$.

(b) Für alle $s \geq 0$ gilt $\lambda(s\mathbb{1}) = s^n$.

(c) Für alle $\Phi, \Psi \in \mathbb{R}^{n \times n}$ gilt $\lambda(\Phi\Psi) = \lambda(\Phi)\lambda(\Psi)$.

Ist $\det(\Phi) = 0$, so ist ΦW_0 nach Schritt 1 eine Jordansche Nullmenge. Also gilt nach Korollar 5.6.4

$$\lambda(\Phi) = \text{Vol}_n(\Phi W_0) = 0$$

Ist $s \geq 0$, so gilt nach Schritt 3

$$\lambda(s\mathbb{1}) = \text{Vol}_n(sW_0) = s^n.$$

Sind $\Phi, \Psi \in \mathbb{R}^{n \times n}$, so gilt nach Schritt 4

$$\lambda(\Phi\Psi) = \text{Vol}_n(\Phi\Psi W_0) = \lambda(\Phi)\text{Vol}_n(\Psi W_0) = \lambda(\Phi)\lambda(\Psi).$$

Damit ist Schritt 5 bewiesen.

Schritt 6. Für jede Elementarmatrix $\Phi \in \mathbb{R}^{n \times n}$ gilt $\lambda(\Phi) = |\det(\Phi)|$.

Eine Elementarmatrix ist entweder eine Permutationsmatrix, oder eine Diagonalmatrix, oder eine Scherung. Eine Permutationsmatrix $\Phi \in \mathbb{R}^{n \times n}$ hat Einträge Null und Eins, mit genau einer Eins in jeder Zeile und jeder Spalte. Eine solche Matrix repräsentiert eine Abbildung $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ der Form

$$\Phi x = (x_{\sigma(1)}, \dots, x_{\sigma(n)}), \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n,$$

wobei $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ eine Permutation (also eine bijektive Abbildung) ist. Sie hat die Determinante $\det(\Phi) = \pm 1$ und bildet den Einheitswürfel W_0 bijektiv auf sich selbst ab. Daher gilt

$$\lambda(\Phi) = \text{Vol}_n(\Phi W_0) = \text{Vol}_n(W_0) = 1 = |\det(\Phi)|.$$

Ist $\Phi = \text{diag}(\lambda_1, \dots, \lambda_n)$ eine Diagonalmatrix, so gilt $\det(\Phi) = \lambda_1 \cdots \lambda_n$ und

$$\Phi W_0 = I_1 \times \cdots \times I_n, \quad I_i := \begin{cases} [0, \lambda_i], & \text{falls } \lambda_i \geq 0, \\ [\lambda_i, 0], & \text{falls } \lambda_i < 0. \end{cases}$$

Das Intervall I_i hat also die Länge $|\lambda_i|$ und daraus folgt

$$\lambda(\Phi) = \text{Vol}_n(\Phi W_0) = \prod_{i=1}^n |\lambda_i| = |\det(\Phi)|.$$

Eine Scherung ist eine Matrix der Form

$$\Phi = \begin{pmatrix} 1 & \lambda & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

mit $\lambda > 0$. Diese Matrix hat Determinante Eins und das Bild des Einheitswürfels ist die Menge

$$\Phi W_0 = \{(x_1 + \lambda x_2, x_2, \dots, x_n) \mid 0 \leq x_i \leq 1\} = P \times Q$$

mit

$$P := \{(x_1 + \lambda x_2, x_2) \mid 0 \leq x_i \leq 1\}, \quad Q := [0, 1]^{n-2}.$$

Nach Beispiel 5.7.9 gilt $\text{Vol}_2(P) = 1$. Also folgt aus Korollar 5.6.6 die Gleichung

$$\lambda(\Phi) = \text{Vol}_n(P \times Q) = \text{Vol}_2(P) \text{Vol}_{n-2}(Q) = 1 = |\det(\Phi)|.$$

Damit ist Schritt 6 bewiesen.

Schritt 7. Für jede Matrix $\Phi \in \mathbb{R}^{n \times n}$ gilt $\lambda(\Phi) = |\det(\Phi)|$.

Nach einem Satz aus der Linearen Algebra läßt sich jede Matrix $\Phi \in \mathbb{R}^{n \times n}$ als Produkt

$$\Phi = \Phi_1 \Phi_2 \cdots \Phi_\ell$$

von Elementarmatrizen schreiben. Da die Determinante eines Produktes quadratischer Matrizen gleich dem Produkt der Determinanten ist (ebenfalls nach einem Satz aus der Linearen Algebra), folgt aus Schritt 5 und Schritt 6 die Gleichung

$$\lambda(\Phi) = \prod_{i=1}^{\ell} \lambda(\Phi_i) = \prod_{i=1}^{\ell} |\det(\Phi_i)| = |\det(\Phi)|.$$

Damit ist Schritt 7 bewiesen.

Die Behauptung von Satz 5.7.8 folgt direkt aus Schritt 2, Schritt 4 und Schritt 7. \square

Nichtlineare Transformationen

Es folgt der Hauptsatz dieses Abschnitts.

Satz 5.7.11 (Transformationsformel). *Sei $U \subset \mathbb{R}^n$ eine offene Menge und $\phi : U \rightarrow \mathbb{R}^n$ eine stetig differenzierbare Abbildung. Sei $A \subset U$ eine kompakte Jordan-messbare Menge und sei $N \subset A$ eine kompakte Jordansche Nullmenge. Wir setzen voraus, dass die Restriktion $\phi|_{A \setminus N} : A \setminus N \rightarrow \mathbb{R}^n$ injektiv ist und die Bedingung*

$$\det(d\phi(x)) \neq 0 \quad \text{für alle } x \in A \setminus N \quad (5.7.8)$$

erfüllt. Dann ist die Menge $B := \phi(A) \subset \mathbb{R}^n$ Jordan-messbar. Ist $f : B \rightarrow \mathbb{R}$ Riemann-integrierbar, so ist auch $f \circ \phi : A \rightarrow \mathbb{R}$ Riemann-integrierbar und es gilt die Gleichung

$$\int_B f(y) dy_1 \cdots dy_n = \int_A f(\phi(x)) |\det(d\phi(x))| dx_1 \cdots dx_n. \quad (5.7.9)$$

Proof. Siehe Seite 214. □

Lemma 5.7.12. *Sei $U \subset \mathbb{R}^n$ eine offene Menge, sei $A \subset U$ eine kompakte Jordan-messbare Menge, sei $\phi : U \rightarrow \mathbb{R}^n$ eine stetig differenzierbare Abbildung, und sei $N := \{x \in A \mid \det(d\phi(x)) = 0\}$ eine Jordansche Nullmenge. Dann ist $B := \phi(A)$ Jordan-messbar und es gilt $\partial B \subset \phi(\partial A \cup N)$.*

Beweis. Zunächst ist B das stetige Bild einer kompakten Menge und ist daher kompakt und damit auch abgeschlossen. Außerdem ist

$$A \setminus (\partial A \cup N) = \overset{\circ}{A} \setminus N$$

eine offene Teilmenge von U und nach Voraussetzung gilt $\det(d\phi(x)) \neq 0$ für alle $x \in A \setminus (\partial A \cup N)$. Nach Korollar 3.1.13 ist daher $\phi(A \setminus (\partial A \cup N))$ eine offene Menge und daraus folgt $\phi(A \setminus (\partial A \cup N)) \subset \overset{\circ}{B}$. Damit erhalten wir

$$\partial B = B \setminus \overset{\circ}{B} \subset \phi(A) \setminus \phi(A \setminus (\partial A \cup N)) \subset \phi(\partial A \cup N).$$

Nach Satz 5.6.7 ist $\phi(\partial A \cup N)$ eine Jordansche Nullmenge. Also ist ∂B eine Jordansche Nullmenge und damit ist Lemma 5.7.12 bewiesen. □

Bemerkung 5.7.13. Die Behauptung von Lemma 5.7.12 gilt auch dann, wenn $N := \{x \in A \mid \det(d\phi(x)) = 0\}$ keine Jordansche Nullmenge ist. Ohne diese Voraussetzung wird der Beweis jedoch aufwendiger. Er beruht darauf, dass $\phi(N)$ auch dann eine Jordansche Nullmenge ist, wenn N selbst keine Jordansche Nullmenge ist. Dies ist ein Spezialfall des **Satzes von Sard** (siehe Bemerkung 3.3.13), und daraus folgt die Behauptung wie in Lemma 5.7.12.

Lemma 5.7.14. *Sei $U \subset \mathbb{R}^n$ eine offenen Menge und $K \subset U$ eine kompakte Teilmenge. Sei $\phi : U \rightarrow \mathbb{R}^n$ eine stetig differenzierbare Abbildung, so dass $\det(d\phi(x)) \neq 0$ is für alle $x \in K$. Dann existiert für jedes $\varepsilon > 0$ eine Konstante $\delta > 0$, so dass folgendes gilt. Ist $0 < s < \delta$ und $a \in \mathbb{R}^n$ mit*

$$W := [a_1 - s, a_1 + s] \times \cdots \times [a_n - s, a_n + s] \subset K, \quad (5.7.10)$$

dann gilt

$$\left| \text{Vol}_n(\phi(W)) - |\det(d\phi(a))| \text{Vol}_n(W) \right| \leq \varepsilon \text{Vol}_n(W). \quad (5.7.11)$$

Beweis. Wir erinnern zunächst an die Definition der Normen

$$\|x\|_\infty := \max_{i=1, \dots, n} |x_i|, \quad \|\Phi\|_\infty := \sup_{0 \neq x \in \mathbb{R}^n} \frac{\|\Phi x\|_\infty}{\|x\|_\infty} \quad (5.7.12)$$

für $x \in \mathbb{R}^n$ und $\Phi \in \mathbb{R}^{n \times n}$. Das ist die Menge $W \subset \mathbb{R}^n$ in (5.7.10) der abgeschlossene Ball $W = \overline{B}_s(a; d_\infty)$ vom Radius s mit Mittelpunkt a bezüglich der Abstandsfunktion $\mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R} : (x, y) \mapsto d_\infty(x, y) := \|x - y\|_\infty$. Nach Voraussetzung sind die Abbildungen

$$K \rightarrow \mathbb{R} : x \mapsto \|d\phi(x)^{-1}\|_\infty, \quad K \rightarrow \mathbb{R} : x \mapsto |\det(d\phi(x))|$$

stetig. Da K kompakt ist, sind diese Funktionen beschränkt. Daher gibt es also eine Konstante $c > 0$, so dass

$$\|d\phi(x)^{-1}\|_\infty \leq c, \quad |\det(d\phi(x))| \leq c \quad \text{für alle } x \in K. \quad (5.7.13)$$

Sei $\varepsilon > 0$ gegeben. Wähle $\rho > 0$ so klein, dass

$$1 - \frac{\varepsilon}{c} < (1 - \rho)^n < (1 + \rho)^n < 1 + \frac{\varepsilon}{c}. \quad (5.7.14)$$

Wähle $\delta > 0$ so klein, dass für alle $x, y \in \mathbb{R}^n$ gilt

$$x, y \in K, \quad \|x - y\|_\infty < \delta \quad \implies \quad \|d\phi(x) - d\phi(y)\| < \frac{\rho}{c}. \quad (5.7.15)$$

Eine solche Konstant existiert, da die Abbildung $d\phi : U \rightarrow \mathbb{R}^{n \times n}$ auf der kompakten Menge $K \subset U$ gleichmäßig stetig ist. Wir beweisen, dass die Behauptung von Lemma 5.7.14 mit dieser Konstanten δ erfüllt ist.

Seien also $a \in \mathbb{R}^n$ und $0 < s < \delta$ gegeben, so dass

$$W := \overline{B}_s(a; d_\infty) \subset K$$

ist. Dann gilt $\|a - x\| < \delta$ für alle $x \in W$ und daraus folgt nach (5.7.15) mit $\Phi := d\phi(a) \in \mathbb{R}^{n \times n}$ die Ungleichung

$$\|d\phi(x) - \Phi\| < \frac{\rho}{c} \leq \frac{\rho}{\|\Phi^{-1}\|_\infty} \quad \text{für alle } x \in W.$$

Hier folgt die zweite Ungleichung aus (5.7.13) mit $x = a$. Betrachten wir nun die Abbildung

$$\psi : U \rightarrow \mathbb{R}^n, \quad \psi(x) := \Phi^{-1}\phi(x).$$

Dann gilt $d\psi(x) = \Phi^{-1}d\phi(x)$ und

$$\|d\psi(x) - \mathbb{1}\|_\infty = \|\Phi^{-1}(d\phi(x) - \Phi)\|_\infty \leq \|\Phi^{-1}\|_\infty \|d\phi(x) - \Phi\|_\infty < \rho$$

für alle $x \in W$. Nach Lemma 3.1.12 folgt daraus

$$(1 - \rho)\Phi(W - a) \subset \phi(W) - \phi(a) \subset (1 + \rho)\Phi(W + a).$$

Nach Lemma 5.7.12 ist $\phi(W) - \phi(a)$ Jordan-messbar und nach Satz 5.7.8 gilt

$$\begin{aligned} (1 - \rho)^n |\det(\Phi)| \text{Vol}_n(W) &= \text{Vol}_n((1 - \rho)\Phi(W - a)) \\ &\leq \text{Vol}_n(\phi(W)) \\ &\leq \text{Vol}_n((1 + \rho)\Phi(W + a)) \\ &= (1 + \rho)^n |\det(\Phi)| \text{Vol}_n(W). \end{aligned}$$

Daraus folgt die Ungleichung

$$\begin{aligned} &\left| \text{Vol}_n(\phi(W)) - |\det(\Phi)| \text{Vol}_n(W) \right| \\ &\leq \max \{1 - (1 - \rho)^n, (1 + \rho)^n - 1\} |\det(\Phi)| \text{Vol}_n(W) \\ &\leq \max \{1 - (1 - \rho)^n, (1 + \rho)^n - 1\} c \text{Vol}_n(W) \\ &\leq \varepsilon \text{Vol}_n(W). \end{aligned}$$

Hier folgt die zweite Ungleichung aus (5.7.13) und die dritte aus (5.7.14). Damit ist Lemma 5.7.14 bewiesen. \square

Beweis von Satz 5.7.11. Der Beweis hat drei Schritte.

Schritt 1. Ist $N = \emptyset$ und $A \subset U$ ein Quadergebäude, das eine Partition aus Würfeln besitzt, so gilt

$$\text{Vol}_n(\phi(A)) = \int_A |\det(d\phi(x))| dx_1 \cdots dx_n. \quad (5.7.16)$$

Sei $\varepsilon > 0$ gegeben und wähle $\delta > 0$ so, dass die Behauptung von Lemma 5.7.14 für $K = A$ gilt. Die Funktion $A \rightarrow \mathbb{R} : x \mapsto |\det(d\phi(x))|$ ist stetig und daher Riemann-integrierbar (siehe Bemerkung 5.5.6). Nach Satz 5.3.1 existiert also eine Partition $P = \{P_1, \dots, P_k\}$ von A so dass \bar{P}_i für jedes i ein abgeschlossener Würfel ist und die für alle $x_i \in \bar{P}_i$ die Ungleichung

$$\left| \int_A |\det(d\phi(x))| dx_1 \cdots dx_n - \sum_{i=1}^k |\det(d\phi(x_i))| \text{Vol}_n(P_i) \right| < \varepsilon \quad (5.7.17)$$

erfüllt. Darüber hinaus können wir P so wählen, dass $\delta(P) < \delta$ ist. Sei nun $x_i \in P_i$ der Mittelpunkt von P_i . Dann gilt nach Wahl von δ die Ungleichung

$$\left| \text{Vol}_n(\phi(P_i)) - |\det(d\phi(x_i))| \text{Vol}_n(P_i) \right| < \varepsilon \text{Vol}_n(P_i) \quad (5.7.18)$$

für $i = 1, \dots, k$. Nun ist $\bigcup_i \partial P_i$ eine Jordansche Nullmenge. Nach Satz 5.6.7 ist damit auch $\phi(\bigcup_i \partial P_i)$ eine Jordansche Nullmenge. Nach Lemma 5.7.12 mit $N = \emptyset$ gilt außerdem $\partial\phi(P_i) \subset \phi(\partial P_i)$ für alle i . Damit ist $\bigcup_i \partial\phi(P_i)$ eine Jordansche Nullmenge und daher gilt $\text{Vol}_n(\phi(A)) = \sum_{i=1}^k \text{Vol}_n(\phi(P_i))$ nach Lemma 5.5.4. Kombinieren wir dies mit den Ungleichungen (5.7.17) und (5.7.18), so ergibt sich

$$\begin{aligned} \left| \text{Vol}_n(\phi(A)) - \int_A |\det(d\phi)| \right| &= \left| \sum_{i=1}^k \text{Vol}_n(\phi(P_i)) - \int_A |\det(d\phi)| \right| \\ &\leq \sum_{i=1}^k \left| \text{Vol}_n(\phi(P_i)) - |\det(d\phi(x_i))| \text{Vol}_n(P_i) \right| \\ &\quad + \left| \sum_{i=1}^k |\det(d\phi(x_i))| \text{Vol}_n(P_i) - \int_A |\det(d\phi)| \right| \\ &\leq (\text{Vol}_n(A) + 1) \varepsilon. \end{aligned}$$

Da $\varepsilon > 0$ beliebig gewählt war, folgt daraus Schritt 1.

Schritt 2. *Unter den Voraussetzungen von Satz 5.7.11 gilt (5.7.16).*

Gegeben sei $\varepsilon > 0$. Wähle $c > 0$ mit

$$\|\phi(x) - \phi(x')\|_\infty \leq c \|x - x'\|_\infty \quad \text{für alle } x, x' \in A. \quad (5.7.19)$$

Eine solche Konstante $c > 0$ existiert nach dem Schrankensatz (Satz 2.4.1).

Wähle eine Konstante $C > 0$ mit

$$|\det(d\phi(x))| \leq C \quad \text{für alle } x \in A. \quad (5.7.20)$$

Wähle zwei Quadergebäude $A_0, A_1 \subset \mathbb{R}^n$, die Partitionen aus Würfeln besitzen und folgende Bedingungen erfüllen

$$\begin{aligned} A_0 &\subset A \setminus N, & \text{Vol}_n(A) - \varepsilon &< \text{Vol}_n(A_0), \\ A \setminus A_0 &\subset A_1, & \text{Vol}_n(A_1) &< \text{Vol}_n(A \setminus A_0) + \varepsilon < 2\varepsilon. \end{aligned} \quad (5.7.21)$$

Solche Quadergebäude existieren nach Satz 5.6.2.

Wähle eine Partition $P = \{P_1, \dots, P_k\}$ von A_1 . Dann ist \bar{P}_i für jedes i ein abgeschlossener Würfel. Ist s_i die Seitenlänge von P_i , so gilt $\|x - x'\|_\infty \leq s_i$ für alle $x, x' \in \bar{P}_i$. Daraus folgt nach (5.7.19) die Ungleichung

$$\|\phi(x) - \phi(x')\|_\infty \leq cs_i \quad \text{für alle } x, x' \in \bar{P}_i \cap A.$$

Daher existiert für jedes i ein abgeschlossener Würfel $Q_i \subset \mathbb{R}^n$ mit Seitenlänge $2cs_i$, so dass $\phi(\bar{P}_i \cap A) \subset Q_i$ und $\text{Vol}_n(Q_i) = (2cs_i)^n$ ist. Außerdem gilt

$$\phi(A) \setminus \phi(A_0) \subset \phi(A \setminus A_0) \subset \phi\left(\bigcup_i \bar{P}_i \cap A\right) \subset \bigcup_i Q_i$$

und daher

$$\text{Vol}_n(\phi(A) \setminus \phi(A_0)) \leq \sum_{i=1}^k \text{Vol}_n(Q_i) = (2c)^n \sum_{i=1}^k s_i^n = (2c)^n \text{Vol}_n(A_1) < 2(2c)^n \varepsilon.$$

Außerdem gilt nach Schritt 1, sowie (5.7.20) und (5.7.21), dass

$$\text{Vol}_n(\phi(A_0)) = \int_{A_0} |\det(d\phi)|, \quad \int_{A \setminus A_0} |\det(d\phi)| \leq C \text{Vol}_n(A \setminus A_0) \leq C\varepsilon.$$

Daraus folgt

$$\begin{aligned} \left| \text{Vol}_n(\phi(A)) - \int_A |\det(d\phi)| \right| &\leq \text{Vol}_n(\phi(A) \setminus \phi(A_0)) + \int_{A \setminus A_0} |\det(d\phi)| \\ &\leq (2(2c)^n + C)\varepsilon. \end{aligned}$$

Da $\varepsilon > 0$ beliebig gewählt war, folgt daraus die Behauptung von Schritt 2.

Schritt 3. Wir beweisen Satz 5.7.11.

Seien $N \subset A \subset U$, $\phi : U \rightarrow \mathbb{R}^n$ und $f : B = \phi(A) \rightarrow \mathbb{R}$ wie in den Voraussetzungen von Satz 5.7.11. Die Supremumsnorm von f wird mit

$$\|f\| := \sup_{y \in B} |f(y)|$$

bezeichnet. Außerdem verwenden wir die Notation (5.7.12).

Sei $\varepsilon > 0$ gegeben. Da $A \subset U$ kompakt und $\phi : U \rightarrow \mathbb{R}^n$ stetig differenzierbar ist, ist die Restriktion $\phi|_A : A \rightarrow \mathbb{R}^n$ nach dem Schrankensatz (Satz 2.4.1) Lipschitz-stetig. Daher existiert eine Konstante $c > 0$ mit

$$\|\phi(x) - \phi(x')\|_\infty \leq c \|x - x'\|_\infty \quad \text{für alle } x, x' \in A. \quad (5.7.22)$$

Da $f : B \rightarrow \mathbb{R}$ Riemann-integrierbar ist, existiert nach Satz 5.7.5 ein $\delta_0 > 0$, so dass für jede Jordan-Partition $Z = \{B_1, \dots, B_k\} \in \mathcal{P}_{\text{Jord}}(B)$ und alle Vektoren $y_1, \dots, y_k \in \mathbb{R}^n$ folgendes gilt:

$$\delta(Z) < c\delta_0, \quad y_i \in B_i \quad \forall i \quad \implies \quad \left| \int_B f - \sum_{i=1}^k f(y_i) \text{Vol}_n(B_i) \right| < \frac{\varepsilon}{1 + \|f\| \text{Vol}_n(A)}. \quad (5.7.23)$$

Da die Funktion $A \rightarrow \mathbb{R} : x \mapsto \det(d\phi(x))$ gleichmäßig stetig ist können wir δ_0 so klein wählen, dass alle $x, x' \in A$ mit $\|x - x'\|_\infty < \delta_0$ die Ungleichung

$$|\det(d\phi(x)) - \det(d\phi(x'))| < \frac{\varepsilon}{1 + \|f\| \text{Vol}_n(A)} \quad (5.7.24)$$

erfüllen. Sei nun $Y = \{A_1, \dots, A_k\} \in \mathcal{P}_{\text{Jord}}(A)$ eine Jordan-Partition von A mit

$$\delta(Y) = \max_{i=1, \dots, k} \delta(A_i) < \delta_0.$$

Da $\phi|_{A \setminus N}$ injektiv ist, gilt $\phi(A_i) \cap \phi(A_j) \subset \phi((A_i \cap A_j) \cup N)$. Nach Satz 5.6.7 folgt daraus, dass $\phi(A_i) \cap \phi(A_j)$ für $i \neq j$ eine Jordansche Nullmenge ist. Damit ist die Menge

$$Z := \{B_1, \dots, B_k\}, \quad B_i := \phi(A_i),$$

eine Jordan-Partition von $B = \phi(A)$. Außerdem gilt $\|x - x'\|_\infty < \delta_0$ für alle i und alle $x, x' \in A_i$. Nach (5.7.22) folgt daraus $\|y - y'\|_\infty < c\delta_0$ für alle i und alle $y, y' \in B_i$. Das heißt, die Jordan-Partition $Z \in \mathcal{P}_{\text{Jord}}(B)$ erfüllt die Ungleichung $\delta(Z) < c\delta_0$. Damit erfüllt Z die Voraussetzung von (5.7.23).

Mit $x_i \in A_i$ und $y_i := \phi(x_i)$ ergibt sich

$$\begin{aligned}
& \left| \int_B f - \sum_{i=1}^k f(\phi(x_i)) |\det(d\phi(x_i))| \text{Vol}_n(A_i) \right| \\
& \leq \left| \int_B f - \sum_{i=1}^k f(y_i) \text{Vol}_n(B_i) \right| \\
& \quad + \|f\| \sum_{i=1}^k |\text{Vol}_n(\phi(A_i)) - |\det(d\phi(x_i))| \text{Vol}_n(A_i)| \\
& < \frac{\varepsilon}{1 + \|f\| \text{Vol}_n(A)} + \|f\| \sum_{i=1}^k \int_{A_i} \|d\phi(x) - d\phi(x_i)\| dx \\
& \leq \frac{\varepsilon}{1 + \|f\| \text{Vol}_n(A)} + \sum_{i=1}^k \frac{\varepsilon \|f\| \text{Vol}_n(A_i)}{1 + \|f\| \text{Vol}_n(A)} = \varepsilon.
\end{aligned}$$

Hier folgt die zweite Ungleichung aus (5.7.23) und die dritte aus (5.7.24). Da $\varepsilon > 0$ beliebig gewählt war, folgt die Gleichung (5.7.9) aus Satz 5.7.5. Damit ist Satz 5.7.11 bewiesen. \square

Beispiel 5.7.15. Die **Jacobi-Transformation** $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ist durch

$$\phi(u, v) := (u - uv, uv)$$

für $(u, v) \in \mathbb{R}^2$ definiert. Bezeichnen wir $x := u - uv$ und $y := uv$, so ergibt sich für die Umkehrabbildung die Formel $u = x + y$ und $v = y/(x + y)$. Das heißt, ϕ bildet die offene Menge $U := \{(u, v) \in \mathbb{R}^2 \mid u \neq 0\}$ diffeomorph auf $V := \{(x, y) \in \mathbb{R}^2 \mid x + y \neq 0\}$ ab und hat die Ableitung

$$d\phi(u, v) = \begin{pmatrix} 1 - v & -u \\ v & u \end{pmatrix}, \quad \det(d\phi(u, v)) = u.$$

Daher sind die Voraussetzungen von Satz 5.7.11 mit $A_R := [0, R] \times [0, 1]$, $N := \{0\} \times [0, 1]$, und $\Delta_R := \phi(A_R) = \{(x, y) \in \mathbb{R}^2 \mid x, y \geq 0, x + y \leq R\}$ erfüllt, und daraus folgt die Gleichung

$$\int_{\Delta_R} f(x, y) dx dy = \int_0^R \int_0^1 f(u - uv, uv) u dv du \quad (5.7.25)$$

für jede stetige Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. In Gleichung (5.7.25) ist es auch interessant den Grenzwert für $R \rightarrow \infty$ zu betrachten, falls dieser existiert. Solche Grenzwerte sind das Thema des nächsten Abschnitts

5.8 Uneigentliche Integrale

Definition 5.8.1. Sei $U \subset \mathbb{R}^n$ offen. Eine Funktion $f : U \rightarrow \mathbb{R}$ heißt **lokal Riemann-integrierbar**, wenn die Restriktion $f|_B : B \rightarrow \mathbb{R}$ für jede kompakte Jordan-messbare Teilmenge $B \subset U$ Riemann-integrierbar ist. Wir verwenden die Bezeichnungen

$$\begin{aligned} \mathcal{B}(U) &:= \{B \subset U \mid B \text{ ist kompakt und Jordan-messbar}\}, \\ \mathcal{R}_{\text{loc}}(U) &:= \{f : U \rightarrow \mathbb{R} \mid f \text{ ist lokal Riemann-integrierbar}\}. \end{aligned}$$

Definition 5.8.2. Sei $U \subset \mathbb{R}^n$ offen. Eine Funktion $f : U \rightarrow \mathbb{R}$ heißt (**uneigentlich**) **Riemann-integrierbar**, wenn sie lokal Riemann-integrierbar ist und

$$\sup_{B \in \mathcal{B}(U)} \int_B |f(x)| dx < \infty. \quad (5.8.1)$$

Übung 5.8.3. Sei $U \subset \mathbb{R}^n$ offen. Dann existiert eine Folge $B_i \in \mathcal{B}(U)$ mit

$$B_i \subset \overset{\circ}{B}_{i+1} \text{ für alle } i \in \mathbb{N}, \quad U = \bigcup_{i=1}^{\infty} B_i. \quad (5.8.2)$$

Satz 5.8.4 (Das uneigentliche Integral).

Sei $U \subset \mathbb{R}^n$ offen und sei $f : U \rightarrow \mathbb{R}$ eine uneigentlich Riemann-integrierbare Funktion. Dann existiert genau eine reelle Zahl $c \in \mathbb{R}$, welche die folgenden beiden Bedingungen erfüllt.

- (i) Für jede Folge $B_i \in \mathcal{B}(U)$, die (5.8.2) erfüllt, gilt $c = \lim_{i \rightarrow \infty} \int_{B_i} f(x) dx$.
- (ii) Für jedes $\varepsilon > 0$ existiert ein $B_0 \in \mathcal{B}(U)$, so dass jede kompakte Menge $B \in \mathcal{B}(U)$ mit $B_0 \subset B \subset U$ die Ungleichung $|c - \int_B f(x) dx| < \varepsilon$ erfüllt.

Beweis. Siehe Seite 219. □

Definition 5.8.5 (Das uneigentliche Integral). Seien U und f wie in Satz 5.8.4. Dann heißt die Zahl $c \in \mathbb{R}$ in Satz 5.8.4 das (**uneigentliche**) **Riemannsche Integral von f über U** und wird mit $\int_U f$ bezeichnet. Das heißt, es gilt

$$\int_U f := \int_U f(x) dx := \lim_{i \rightarrow \infty} \int_{B_i} f(x) dx \quad (5.8.3)$$

für jede Folge $B_i \in \mathcal{B}(U)$ die die Bedingungen in (5.8.2) erfüllt.

Beweis von Satz 5.8.4. Der Beweis hat fünf Schritte.

Schritt 1. Sei $B_i \in \mathcal{B}(U)$ eine Folge kompakter Jordan-messbarer Teilmengen von U , die (5.8.2) erfüllt. Sei $B \subset U$ eine beliebige kompakte Teilmenge. Dann existiert eine Zahl $i \in \mathbb{N}$ so dass $B \subset B_i$ ist.

Wir argumentieren indirekt und nehmen an, dass $B \not\subset B_i$ ist für alle $i \in \mathbb{N}$. Dann existiert nach dem abzählbaren Auswahlaxiom eine Folge $(x_i)_{i \in \mathbb{N}}$ mit

$$x_i \in B \setminus B_i \quad \text{für alle } i \in \mathbb{N}. \quad (5.8.4)$$

Da B kompakt ist existiert eine konvergente Teilfolge $(x_{i_\nu})_{\nu \in \mathbb{N}}$ mit Grenzwert

$$x_0 := \lim_{\nu \rightarrow \infty} x_{i_\nu} \in B.$$

Da B eine Teilmenge von U ist, gilt $x_0 \in U$ und, da $U = \bigcup_{i=1}^{\infty} B_i$ ist nach (5.8.2), existiert folglich eine Zahl $i_0 \in \mathbb{N}$ mit

$$x_0 \in B_{i_0} \subset \overset{\circ}{B}_{i_0+1}.$$

Da die Folge $(x_{i_\nu})_{\nu \in \mathbb{N}}$ gegen x_0 konvergiert und die Menge $\overset{\circ}{B}_{i_0+1}$ offen ist, existiert eine Zahl $\nu_0 \in \mathbb{N}$, so dass für alle $\nu \in \mathbb{N}$ folgendes gilt:

$$\nu \geq \nu_0 \quad \implies \quad x_\nu \in \overset{\circ}{B}_{i_0+1}.$$

Wählen wir die Zahl $\nu \in \mathbb{N}$ so gross, dass $\nu \geq \nu_0$ ist und $i_\nu > i_0$ ist, so gilt

$$x_{i_\nu} \in \overset{\circ}{B}_{i_0+1} \subset B_{i_0+1} \subset B_{i_\nu}.$$

Diese steht im Widerspruch zu (5.8.4) und damit ist Schritt 1 bewiesen.

Schritt 2. Sei $B_i \in \mathcal{B}(U)$ eine Folge kompakter Jordan-messbarer Teilmengen von U , die (5.8.2) erfüllt. Dann gilt $\lim_{i \rightarrow \infty} \int_{B_i} |f| = \sup_{B \in \mathcal{B}(U)} \int_K |f|$.

Sei $C := \sup_{B \in \mathcal{B}(U)} \int_K |f|$ und sei eine Konstante $\varepsilon > 0$ gegeben. Dann existiert eine Menge $B \in \mathcal{B}(U)$ mit

$$\int_B |f| > C - \varepsilon.$$

Nach Schritt 1 existiert eine Zahl $i_0 \in \mathbb{N}$ mit $B \subset B_{i_0}$. Ist $i \in \mathbb{N}$ mit $i \geq i_0$ so gilt $B \subset B_{i_0} \subset B_i$ und daher

$$C - \varepsilon < \int_B |f| \leq \int_{B_i} |f| \leq C.$$

Damit ist Schritt 2 bewiesen.

Schritt 3. Sei $B_i \in \mathcal{B}(U)$ eine Folge kompakter Jordan-messbarer Teilmengen von U , die (5.8.2) erfüllt. Dann ist $c_i := \int_{B_i} f(x) dx$ eine Cauchy-Folge. Sei $\varepsilon > 0$ gegeben. Dann existiert nach Schritt 2 eine natürliche Zahl $i_0 \in \mathbb{N}$ mit $\int_{B_{i_0}} |f| > C - \varepsilon$. Nun seien $i, j \in \mathbb{N}$ gegeben mit $j \geq i \geq i_0$. Dann gilt

$$|c_i - c_j| = \left| \int_{B_j \setminus B_i} f \right| \leq \int_{B_j \setminus B_i} |f| = \int_{B_j} |f| - \int_{B_i} |f| \leq C - \int_{B_{i_0}} |f| < \varepsilon.$$

Damit ist Schritt 3 bewiesen.

Schritt 4. Sei $B_i \in \mathcal{B}(U)$ eine Folge kompakter Jordan-messbarer Teilmengen von U , die (5.8.2) erfüllt. Dann existiert der Grenzwert

$$c := \lim_{i \rightarrow \infty} \int_{B_i} f(x) dx. \quad (5.8.5)$$

und erfüllt die Bedingung (ii) in Satz 5.8.4.

Der Grenzwert existiert nach Schritt 3. Sei $C := \sup_{B \in \mathcal{B}(U)} \int_B |f|$ und $\varepsilon > 0$ gegeben. Dann existiert nach (5.8.5) und Schritt 2 eine Zahl $i_0 \in \mathbb{N}$ mit

$$\left| c - \int_{B_{i_0}} f \right| < \frac{\varepsilon}{2}, \quad \int_{B_{i_0}} |f| > C - \frac{\varepsilon}{2}.$$

Sei $B \in \mathcal{B}(U)$ mit $B_{i_0} \subset B$. Dann gilt

$$\left| \int_{B \setminus B_{i_0}} f \right| \leq \int_{B \setminus B_{i_0}} |f| = \int_B |f| - \int_{B_{i_0}} |f| \leq C - \int_{B_{i_0}} |f| < \frac{\varepsilon}{2}$$

und daher

$$\left| c - \int_B f \right| = \left| c - \int_{B_{i_0}} f - \int_{B \setminus B_{i_0}} f \right| \leq \left| c - \int_{B_{i_0}} f \right| + \left| \int_{B \setminus B_{i_0}} f \right| < \varepsilon.$$

Damit ist Schritt 4 bewiesen.

Schritt 5. Sei $c \in \mathbb{R}$ eine Konstante, die Bedingung (ii) in Satz 5.8.4 erfüllt. Dann erfüllt c auch Bedingung (i) in Satz 5.8.4.

Sei $B_i \in \mathcal{B}(U)$ eine Folge kompakter Jordan-messbarer Teilmengen von U , die (5.8.2) erfüllt, und sei $\varepsilon > 0$. Dann existiert nach (ii) ein $B_0 \in \mathcal{B}(U)$, so dass jedes $B \in \mathcal{B}(U)$ mit $B_0 \subset B$ die Ungleichung $|c - \int_B f| < \varepsilon$ erfüllt. Nun existiert nach Schritt 1 ein $i_0 \in \mathbb{N}$ mit $B_0 \subset B_{i_0}$. Ist $i \in \mathbb{N}$ mit $i \geq i_0$, so gilt $B_0 \subset B_{i_0} \subset B_i$, und daher $|c - \int_{B_i} f| < \varepsilon$. Damit haben wir gezeigt, dass $c = \lim_{i \rightarrow \infty} \int_{B_i} f$ ist. Damit sind Schritt 5 und Satz 5.8.4 bewiesen. \square

Beispiele

Beispiel 5.8.6. Seien $p, q > 0$ und sei die Funktion $f : (0, \infty)^2 \rightarrow \mathbb{R}$ durch

$$f(x, y) := x^{p-1}y^{q-1}e^{-x-y}$$

für $x, y > 0$ gegeben. Für $R > 1$ ist $K_R := [R^{-1}, R]^2$ eine kompakte Teilmenge von $U = (0, \infty)^2$ und nach dem Satz von Fubini (Satz 5.4.1) gilt

$$\int_{K_R} f(x, y) dx dy = \left(\int_{1/R}^{\infty} x^{p-1} e^{-x} dx \right) \left(\int_{1/R}^{\infty} y^{q-1} e^{-y} dy \right).$$

Daraus folgt, mit dem Grenzübergang $R \rightarrow \infty$, dass die Funktion $f : U \rightarrow \mathbb{R}$ uneigentlich Riemann-integrierbar ist und ihr Integral durch

$$\int_U f(x, y) dx dy = \Gamma(p)\Gamma(q) \quad (5.8.6)$$

gegeben ist. (Hier ist $\Gamma(p) := \int_0^{\infty} x^{p-1} e^{-x} dx$ Gamma-Funktion.) Andererseits läßt sich das Integral von f nach Satz 5.8.4 auch als Grenzwert der Integrale über die Mengen $\Delta_R = \{(x, y) \in \mathbb{R}^2 \mid x > 0, y > 0, x + y < R\}$ mit $R \rightarrow \infty$ schreiben. Nach Beispiel 5.7.15 folgt daraus

$$\begin{aligned} \int_U f(x, y) dx dy &= \lim_{R \rightarrow \infty} \int_{\Delta_R} f(x, y) dx dy \\ &= \lim_{R \rightarrow \infty} \int_0^R \int_0^1 f(u - uv, uv) u dv du \\ &= \lim_{R \rightarrow \infty} \int_0^R \int_0^1 (u - uv)^{p-1} (uv)^{q-1} u e^{-u} dv du \\ &= \lim_{R \rightarrow \infty} \int_0^R \int_0^1 (1 - v)^{p-1} v^{q-1} u^{p+q-1} e^{-u} dv du \\ &= \int_0^1 (1 - v)^{p-1} v^{q-1} dv \lim_{R \rightarrow \infty} \int_0^R u^{p+q-1} e^{-u} du \\ &= \int_0^1 (1 - v)^{p-1} v^{q-1} dv \Gamma(p + q). \end{aligned}$$

Daraus ergibt sich nach (5.8.6) die Formel

$$\int_0^1 (1 - t)^{p-1} t^{q-1} dt = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p + q)}. \quad (5.8.7)$$

Beispiel 5.8.7. Wiederholen wir das Argument aus Beispiel 5.7.15 für $R = 1$ mit den offenen Mengen $(0, 1)^2$ und

$$\Delta := \{(x, y) \in \mathbb{R}^2 \mid x > 0, y > 0, x + y < 1\} \quad (5.8.8)$$

und der uneigentlich Riemann-integrierbaren Funktion $f : \Delta \rightarrow \mathbb{R}$, die durch $f(x, y) = x^{p-1}y^{q-1}$ mit $p, q > 0$ gegeben ist, so ergibt sich für das uneigentliche Integral die Formel

$$\begin{aligned} \int_{\Delta} x^{p-1}y^{q-1} dx dy &= \int_0^1 \int_0^1 (u - uv)^{p-1} (uv)^{q-1} u dv du \\ &= \left(\int_0^1 (1 - v)^{p-1} v^{q-1} dv \right) \left(\int_0^1 u^{p+q-1} du \right) \\ &= \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)} \frac{1}{p+q} = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q+1)}. \end{aligned} \quad (5.8.9)$$

Hier folgt der erste Schritt aus Gleichung (5.7.25) in Beispiel 5.7.15 und der dritte Schritt aus der Gleichung (5.8.7) in Beispiel 5.8.6.

Beispiel 5.8.8. Seien $a, b, \alpha, \beta, p, q > 0$. Wir beweisen die Formel

$$\begin{aligned} \int_{\Delta_{a,b}^{\alpha,\beta}} x^{p-1}y^{q-1} dx dy &= \frac{a^p b^q}{\alpha\beta} \frac{\Gamma\left(\frac{p}{\alpha}\right)\Gamma\left(\frac{q}{\beta}\right)}{\Gamma\left(\frac{p}{\alpha} + \frac{q}{\beta} + 1\right)}, \\ \Delta_{a,b}^{\alpha,\beta} &:= \left\{ (x, y) \in \mathbb{R}^2 \mid x > 0, y > 0, \left(\frac{x}{a}\right)^{\alpha} + \left(\frac{y}{b}\right)^{\beta} < 1 \right\}. \end{aligned} \quad (5.8.10)$$

Die Abbildung $\Delta \rightarrow \Delta_{a,b}^{\alpha,\beta} : (\xi, \eta) \mapsto \phi(\xi, \eta) := (a\xi^{1/\alpha}, b\eta^{1/\beta})$ ist ein Diffeomorphismus mit der Ableitung

$$d\phi(\xi, \eta) = \begin{pmatrix} \frac{a}{\alpha} \xi^{\frac{1}{\alpha}-1} & 0 \\ 0 & \frac{b}{\beta} \eta^{\frac{1}{\beta}-1} \end{pmatrix}.$$

Daher folgt aus der Transformationsformel in Satz 5.7.11 mit Hilfe der entsprechenden Grenzübergänge für uneigentliche Integrale die Gleichung

$$\int_{\Delta_{a,b}^{\alpha,\beta}} f(x, y) dx = \frac{ab}{\alpha\beta} \int_{\Delta} f(a\xi^{1/\alpha}, b\eta^{1/\beta}) \xi^{\frac{1}{\alpha}-1} \eta^{\frac{1}{\beta}-1} d\xi d\eta.$$

Mit $f(x, y) = x^{p-1}y^{q-1}$ erhalten wir

$$\int_{\Delta_{a,b}^{\alpha,\beta}} x^{p-1}y^{q-1} dx = \frac{a^p b^q}{\alpha\beta} \int_{\Delta} \xi^{\frac{p}{\alpha}-1} \eta^{\frac{q}{\beta}-1} d\xi d\eta$$

Damit folgt (5.8.10) aus Gleichung (5.8.9) in Beispiel 5.8.7.

Beispiel 5.8.9. Sei $\|\cdot\|$ die Euklidische Norm auf dem \mathbb{R}^n und sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ die Funktion $f(x) := e^{-\|x\|^2} = e^{-x_1^2 - \dots - x_n^2}$ für $x \in \mathbb{R}^n$. Diese Funktion ist uneigentlich Riemann-integrierbar. Wir beweisen die Formel

$$\int_{\mathbb{R}^n} e^{-\|x\|^2} dx = \pi^{n/2}. \quad (5.8.11)$$

Zunächst gilt nach dem Satz von Fubini für jedes $n \in \mathbb{N}$ die Gleichung

$$\int_{\mathbb{R}^n} e^{-\|x\|^2} dx_1 \cdots dx_n = \left(\int_{-\infty}^{\infty} e^{-x^2} dx \right)^n. \quad (5.8.12)$$

(Man wende Satz 5.4.1 auf das Integral über $[-R, R]^n$ an und betrachte den Grenzwert $R \rightarrow \infty$.) Sei $\phi : (0, \infty) \times (-\pi, \pi) \rightarrow \mathbb{R}^2 \setminus ((-\infty, 0] \times \{0\})$ der durch $\phi(r, \theta) := (r \cos(\theta), r \sin(\theta))$ definierte Diffeomorphismus (Polarkoordinaten für $n = 2$). Dann gilt $\det(d\phi(r, \theta)) = r$ und daher

$$\int_{\mathbb{R}^2} f(x, y) dx dy = \int_0^{\infty} \int_{-\pi}^{\pi} f(r \cos(\theta), r \sin(\theta)) r d\theta dr$$

für jede uneigentlich Riemann-integrierbare Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. Für die Funktion $f(x, y) := e^{-x^2 - y^2}$ erhalten wir mit Hilfe von (5.8.12) die Formel

$$\left(\int_{-\infty}^{\infty} e^{-x^2} dx \right)^2 = \int_0^{\infty} \int_{-\pi}^{\pi} e^{-r^2} r d\theta dr = 2\pi \int_0^{\infty} e^{-r^2} r dr = \pi.$$

Hier folgt der letzte Schritt aus der Substitutionsformel für Integrale in einer Variablen mit $s := r^2$. Daraus ergibt sich die Formel $\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$ und daher folgt die Gleichung (5.8.11) aus (5.8.12).

Beispiel 5.8.10. Sei $A = A^T \in \mathbb{R}^{n \times n}$ eine symmetrische positiv definite Matrix. Dann gilt

$$\int_{\mathbb{R}^n} e^{-x^T A x} dx = \frac{\pi^{n/2}}{\sqrt{\det(A)}}. \quad (5.8.13)$$

Zum Beweis verwenden wir die aus der Linearen Algebra bekannte Tatsache, dass eine (eindeutige) symmetrische positiv definite Matrix $B = B^T \in \mathbb{R}^{n \times n}$ existiert mit $B^2 = A$. Sei $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ der Diffeomorphismus $\phi(x) := Bx$. Dann ergibt sich aus Beispiel 5.8.9 zusammen mit der Transformationsformel in Satz 5.7.11 die Gleichung

$$\pi^{n/2} = \int_{\mathbb{R}^n} e^{-\|y\|^2} dy = \int_{\mathbb{R}^n} e^{-\|Bx\|^2} \det(B) dx = \int_{\mathbb{R}^n} e^{-x^T A x} dx \sqrt{\det(A)}.$$

5.9 Divergenz

In dem letzten Abschnitt dieses Kapitels verbinden wir die Transformationsformel aus Satz 5.7.11 mit dem Fluss eines stetig differenzierbaren Vektorfeldes in Kapitel 4. Hierzu beginnen wir mit der folgenden Beobachtung. Seien $U, V \subset \mathbb{R}^n$ zwei offene Mengen und sei $\phi : U \rightarrow V$ ein C^1 -Diffeomorphismus. Ist B eine kompakte Jordan-messbare Teilmenge von U , so ist $\phi(B)$ eine kompakte Jordan-messbare Teilmenge von V . Daß $\phi(B)$ kompakt ist folgt aus der Stetigkeit von ϕ und dass der Rand $\partial\phi(B) = \phi(\partial B)$ eine Jordansche Nullmenge ist, folgt aus Teil (ii) von Satz 5.6.7. Damit ist also $\phi(B)$ Jordan-messbar, wie behauptet, und hat nach der Transformationsformel (5.7.9) in Satz 5.7.11 mit $f \equiv 1$ das Volumen

$$\text{Vol}_n(\phi(B)) = \int_B |\det(d\phi(x))| dx. \quad (5.9.1)$$

Definition 5.9.1. Seien $U, V \subset \mathbb{R}^n$ offene Mengen. Ein C^1 -Diffeomorphismus $\phi : U \rightarrow V$ heißt **volumentreu** wenn jede kompakte Jordan-messbare Teilmenge $B \subset U$ die Gleichung $\text{Vol}_n(\phi(B)) = \text{Vol}_n(B)$ erfüllt.

Lemma 5.9.2. Seien $U, V \subset \mathbb{R}^n$ offene Mengen und sei $\phi : U \rightarrow V$ ein C^1 -Diffeomorphismus. Dann sind folgende Aussagen äquivalent.

- (i) ϕ ist volumentreu.
- (ii) Es gilt $|\det(d\phi(x))| = 1$ für alle $x \in U$.

Beweis. Die Implikation (ii) \implies (i) folgt aus der Gleichung (5.9.1), die wiederum direkt aus der Transformationsformel (5.7.9) in Satz 5.7.11 folgt. Zum Beweis der Umkehrung nehmen wir an, dass ϕ volumentreu ist, und wählen ein Element $a = (a_1, \dots, a_n) \in U$. Für $\varepsilon > 0$ sei

$$Q_\varepsilon := [a_1 - \varepsilon, a_1 + \varepsilon] \times \cdots \times [a_n - \varepsilon, a_n + \varepsilon].$$

Diese Menge ist Jordan-messbar und hat das Volumen $(2\varepsilon)^n$. Für $\varepsilon > 0$ hinreichend klein gilt $Q_\varepsilon \subset U$ und daher, nach (i) und (5.9.1),

$$1 = \frac{\text{Vol}_n(\phi(Q_\varepsilon))}{\text{Vol}_n(Q_\varepsilon)} = \frac{\text{Vol}_n(\phi(Q_\varepsilon))}{(2\varepsilon)^n} = \frac{1}{(2\varepsilon)^n} \int_{Q_\varepsilon} |\det(d\phi(x))| dx$$

Mit dem Grenzübergang $\varepsilon \rightarrow 0$ folgt daraus

$$|\det(d\phi(a))| = \lim_{\varepsilon \rightarrow 0} \frac{1}{(2\varepsilon)^n} \int_{Q_\varepsilon} |\det(d\phi(x))| dx = 1.$$

Also erfüllt ϕ die Bedingung (ii) und damit ist Lemma 5.9.2 bewiesen. \square

Beispiel 5.9.3. Sei $\phi : \mathbb{R}^2 \rightarrow \mathbb{R} \times (0, \infty)$ der durch

$$\phi(x, y) := (e^{-y}x, e^y)$$

für $(x, y) \in \mathbb{R}^2$ definierte Diffeomorphismus. Die Umkehrabbildung ist durch

$$\phi^{-1}(\xi, \eta) = (\xi\eta, \log(\eta))$$

für $\xi \in \mathbb{R}$ und $\eta > 0$ gegeben. Die Jacobi-Matrix

$$d\phi(x, y) = \begin{pmatrix} e^{-y} & -e^{-y}x \\ 0 & e^y \end{pmatrix}$$

hat für alle $(x, y) \in \mathbb{R}^2$ die Determinante 1 und daher ist ϕ volumentreu.

Wir wenden uns nun der Frage zu, unter welchen Voraussetzungen der Fluss eines stetig differenzierbaren Vektorfeldes aus volumentreue Diffeomorphismen besteht. Sei also $U \subset \mathbb{R}^n$ eine offene Menge, sei $f : U \rightarrow \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld, und sei $\phi : \Omega \rightarrow U$ der Fluss dieses Vektorfeldes. Das bedeutet, dass $\Omega \subset \mathbb{R} \times U$ die Menge aller Paare (t, x_0) mit $x_0 \in U$ und $t \in I(x_0)$ ist, wobei $I(x_0) \subset \mathbb{R}$ das maximale Existenzintervall für die Lösungen des Anfangswertproblems

$$\dot{x} = f(x), \quad x(0) = x_0 \tag{5.9.2}$$

ist, und dass die Funktion $I(x_0) \rightarrow U : t \mapsto \phi(t, x_0)$ für jedes $x_0 \in U$ genau die eindeutige Lösung dieses Anfangswertproblems ist (siehe Definition 4.2.2). Nach Satz 4.3.1 ist Ω eine offene Teilmenge von $\mathbb{R} \times U$ und nach Satz 4.4.1 ist die Abbildung $\phi : \Omega \rightarrow U$ stetig differenzierbar. Mit

$$U_t := \{x \in U \mid t \in I(x)\}$$

und $\phi_t(x) := \phi(t, x)$ für $x \in U_t$ folgt daraus, dass die Abbildung

$$\phi_t : U_t \rightarrow U_{-t}$$

für jedes $t \in \mathbb{R}$ ein C^1 -Diffeomorphismus zwischen offenen Teilmengen des \mathbb{R}^n ist (Bemerkung 4.3.3). Die Frage ist nun, unter welcher Voraussetzung ϕ_t volumentreu ist. Dazu erinnern wir uns daran, dass für $x_0 \in U$ und $t \in I(x_0)$ die Jacobi-Matrix $\Phi(t) := d\phi_t(x_0)$ die Lösung des Anfangswertproblems

$$\dot{\Phi}(t) = A(t)\Phi(t), \quad \Phi(0) = \mathbb{1}, \tag{5.9.3}$$

mit $A(t) := df(\phi_t(x_0))$ ist (Lemma 4.4.2).

Lemma 5.9.4. Sei $I \subset \mathbb{R}$ ein offenes Intervall mit $0 \in I$, sei $A : I \rightarrow \mathbb{R}^{n \times n}$ eine stetige Funktion, und sei $\Phi : I \rightarrow \text{GL}(n, \mathbb{R})$ die Lösung von (5.9.3). Dann gilt

$$\frac{d}{dt} \det(\Phi(t)) = \text{spur}(A(t)) \det(\Phi(t)) \quad (5.9.4)$$

für alle $t \in I$.

Beweis. Die Funktion $\det : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ ist ein homogenes Polynom vom Grade n in n^2 Variablen und ist daher glatt (Satz D.3.3). Sei

$$\rho := \det|_{\text{GL}(n, \mathbb{R})} : \text{GL}(n, \mathbb{R}) \rightarrow \mathbb{R}$$

die Einschränkung der Determinante auf die allgemeine lineare Gruppe (eine offene Teilmenge von $\mathbb{R}^{n \times n}$). Wir beweisen, dass die Ableitung von ρ an der Stelle $\Phi \in \text{GL}(n, \mathbb{R})$ die durch

$$d\rho(\Phi)\widehat{\Phi} = \text{spur}(\Phi^{-1}\widehat{\Phi}) \det(\Phi) \quad (5.9.5)$$

für $\widehat{\Phi} \in \mathbb{R}^{n \times n}$ gegebene lineare Abbildung $d\rho(\Phi) : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ ist. Dazu betrachten wir zunächst den Fall $\Phi = \mathbb{1}$. Dann gilt für $A = (a_{ij})_{i,j=1}^n \in \mathbb{R}^{n \times n}$ die Gleichung $\frac{d}{dt}|_{t=0} e^{tA} = A$ (Satz 1.5.2) und daher

$$\begin{aligned} d\rho(\mathbb{1})A &= \left. \frac{d}{dt} \right|_{t=0} \rho(e^{tA}) \\ &= \left. \frac{d}{dt} \right|_{t=0} \det(e^{tA}e_1, \dots, e^{tA}e_n) \\ &= \sum_{i=1}^n \det(e_1, \dots, e_{i-1}, Ae_i, e_{i+1}, \dots, e_n) \\ &= \sum_{i=1}^n \det(e_1, \dots, e_{i-1}, \sum_{j=1}^n a_{ji}e_j, e_{i+1}, \dots, e_n) \\ &= \sum_{i=1}^n a_{ii} \det(e_1, \dots, e_n) \\ &= \text{spur}(A). \end{aligned}$$

Hier bezeichnet e_1, \dots, e_n die Standardbasis des \mathbb{R}^n und wir haben die Tatsache verwendet, dass die Determinanten-Abbildung $\Phi \mapsto \det(\Phi)$ linear in jeder Spalte von Φ ist und verschwindet, sobald zwei Spalten von Φ übereinstimmen. Damit ist die Gleichung (5.9.5) für $\Phi = \mathbb{1}$ bewiesen.

Sei nun $\Phi \in \text{GL}(n, \mathbb{R})$ und $\widehat{\Phi} \in \mathbb{R}^{n \times n}$ und definiere

$$A := \Phi^{-1}\widehat{\Phi}.$$

Dann gilt

$$\left. \frac{d}{dt} \right|_{t=0} \Phi e^{tA} = \Phi A = \widehat{\Phi}$$

und daher

$$\begin{aligned} d\rho(\Phi)\widehat{\Phi} &= \left. \frac{d}{dt} \right|_{t=0} \rho(\Phi e^{tA}) \\ &= \left. \frac{d}{dt} \right|_{t=0} \rho(\Phi)\rho(e^{tA}) \\ &= \rho(\Phi)d\rho(\mathbb{1})A \\ &= \det(\Phi)\text{spur}(A). \end{aligned}$$

Damit ist die Gleichung (5.9.5) bewiesen, und daraus folgt für die Lösung

$$I \rightarrow \text{GL}(n, \mathbb{R}) : t \mapsto \Phi(t)$$

von Gleichung (5.9.3) nach der Kettenregel die Formel

$$\begin{aligned} \frac{d}{dt} \det(\Phi(t)) &= d\rho(\Phi(t))\dot{\Phi}(t) \\ &= \text{spur}(\Phi(t)^{-1}\dot{\Phi}(t)) \det(\Phi(t)) \\ &= \text{spur}(A(t)) \det(\Phi(t)). \end{aligned}$$

Damit ist Lemma 5.9.4 bewiesen. □

Definition 5.9.5 (Divergenz). Sei $U \subset \mathbb{R}^n$ eine offene Teilmenge und sei $f = (f_1, \dots, f_n) : U \rightarrow \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld. Die **Divergenz** von f ist die Funktion $\text{div}(f) : U \rightarrow \mathbb{R}$, die durch

$$(\text{div}(f))(x) := \sum_{i=1}^n \frac{\partial f_i}{\partial x_i}(x) = \text{spur}(df(x)) \quad (5.9.6)$$

für $x \in U$ definiert ist. Das Vektorfeld f heißt **divergenzfrei** wenn die Divergenz von f auf ganz U verschwindet, das heißt, wenn $\text{spur}(df(x)) = 0$ ist für alle $x \in U$.

Satz 5.9.6. Sei $U \subset \mathbb{R}^n$ eine offene Menge, sei $f : U \rightarrow \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld, und sei $\phi_t : U_t \rightarrow U_{-t}$ der Fluss von f . Dann sind folgende Aussagen äquivalent.

(i) Das Vektorfeld f ist divergenzfrei.

(ii) Der C^1 -Diffeomorphismus $\phi_t : U_t \rightarrow U_{-t}$ ist volumentreu für alle $t \in \mathbb{R}$.

Beweis. Lemma 5.9.4 zeigt, dass die Determinante der Jacobi-Matrix von ϕ_t die Differentialgleichung

$$\begin{aligned} \frac{d}{dt} \det(d\phi_t(x_0)) &= \operatorname{spur}(df(\phi_t(x_0))) \det(d\phi_t(x_0)), \\ \det(d\phi_0(x_0)) &= 1, \end{aligned} \quad (5.9.7)$$

erfüllt. Daraus folgt dass das Vektorfeld f genau dann divergenzfrei ist wenn der Fluss die Bedingung $\det(d\phi_t(x_0)) = 1$ für alle $x_0 \in U$ und alle $t \in I(x_0)$ erfüllt. Da die Determinante von $d\phi_t(x_0)$ niemals negativ sein kann, ist diese Bedingung nach Lemma 5.9.2 dazu äquivalent, dass $\phi_t : U_t \rightarrow U_{-t}$ für jedes $t \in \mathbb{R}$ volumentreu ist. Damit ist Satz 5.9.6 bewiesen. \square

Beispiel 5.9.7. Einfache Beispiele von divergenzfreien Vektorfeldern auf dem \mathbb{R}^2 sind das Vektorfeld

$$f(x, y) := (x, -y)$$

mit einer hyperbolischen Nullstelle im Ursprung und dem Fluss

$$\phi_t(x, y) = (e^t x, e^{-t} y),$$

und das Vektorfeld

$$f(x, y) := (-y, x)$$

mit einer elliptischen Nullstelle im Ursprung und dem Fluss

$$\phi_t(x, y) = (\cos(t)x - \sin(t)y, \sin(t)x + \cos(t)y).$$

Kapitel 6

Integration über Untermannigfaltigkeiten

Das Ziel dieses Kapitels ist es, das Riemannsche Integral einer stetigen Funktion über einer d -dimensionalen Untermannigfaltigkeit $M \subset \mathbb{R}^n$ zu definieren und den Satz von Gauß zu beweisen. Das Kapitel beginnt mit einer Einführung in das d -dimensionale Volumen und das Integral über ein Kartengebiet (Abschnitt 6.1). Anschliessend wird das Integral über ganz M definiert (Abschnitt 6.2), einige Beispiele diskutiert (Abschnitt 6.3), und der Satz von Gauß bewiesen (Abschnitt 6.4).

6.1 Das d -dimensionale Volumen

Die Länge einer Kurve

Dieser Abschnitt beginnt mit der Einführung des Begriffs der *Länge einer Kurve* im \mathbb{R}^n . Wir arbeiten stets mit dem Euklidischen inneren Produkt

$$\langle x, y \rangle := \sum_{i=1}^n x_i y_i, \quad \|x\| := \sqrt{\langle x, x \rangle}$$

für $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ und $y = (y_1, \dots, y_n) \in \mathbb{R}^n$. Sei $I = [a, b]$ ein kompaktes Intervall. Eine **Partition** von I ist ein $(N+1)$ -Tupel reeller Zahlen $P = (t_0, t_1, \dots, t_N)$ mit $N \in \mathbb{N}$ und

$$a = t_0 < t_1 < \dots < t_{N-1} < t_N = b. \quad (6.1.1)$$

(Siehe [5].) Die Menge aller Partitionen von I wird mit $\mathcal{P}(I)$ bezeichnet.

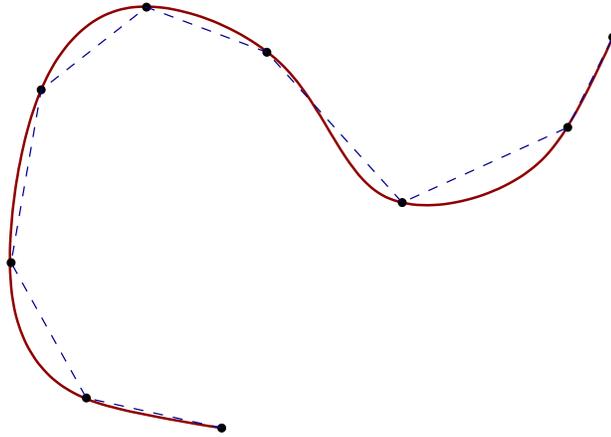


Abbildung 6.1: Die Länge einer Kurve.

Definition 6.1.1 (Die Länge einer Kurve).

Seien $a < b$ reelle Zahlen, sei $I := [a, b]$, und sei $\gamma : I \rightarrow \mathbb{R}^n$ eine stetige Abbildung. Für eine Partition $P = (t_0, t_1, \dots, t_N) \in \mathcal{P}(I)$ definieren wir

$$L(\gamma, P) := \sum_{i=1}^N \|\gamma(t_i) - \gamma(t_{i-1})\|. \quad (6.1.2)$$

Die **Länge von γ** ist die Zahl

$$L(\gamma) := \sup_{P \in \mathcal{P}(I)} L(\gamma, P) \in [0, \infty].$$

Die Kurve γ heißt **rektifizierbar** wenn $L(\gamma) < \infty$ ist.

Beispiel 6.1.2. Sei $\gamma : [0, 1] \rightarrow \mathbb{R}$ die Funktion, die auf der Cantor-Menge

$$K := \bigcap_{n=1}^{\infty} K_n, \quad K_n := \bigcup_{a_1, \dots, a_n \in \{0,1\}} \left[2 \sum_{i=1}^n \frac{a_i}{3^i}, 2 \sum_{i=1}^n \frac{a_i}{3^i} + \frac{1}{3^n} \right],$$

verschwindet und, für jedes $n \in \mathbb{N}$ und alle $a_1, \dots, a_n \in \{0, 1\}$ mit $a_n = 1$, durch die Formel

$$\gamma(t) := \frac{3^n}{2^{n-2}} \min \left\{ \left(t + \frac{1}{3^n} - 2 \sum_{i=1}^n \frac{a_i}{3^i} \right), \left(2 \sum_{i=1}^n \frac{a_i}{3^i} - t \right) \right\}$$

für $2 \sum_{i=1}^n \frac{a_i}{3^i} - \frac{1}{3^n} \leq t \leq 2 \sum_{i=1}^n \frac{a_i}{3^i}$ gegeben ist. Diese Kurve $\gamma : [0, 1] \rightarrow \mathbb{R}$ ist stetig, aber nicht rektifizierbar (siehe Abbildung 6.2).

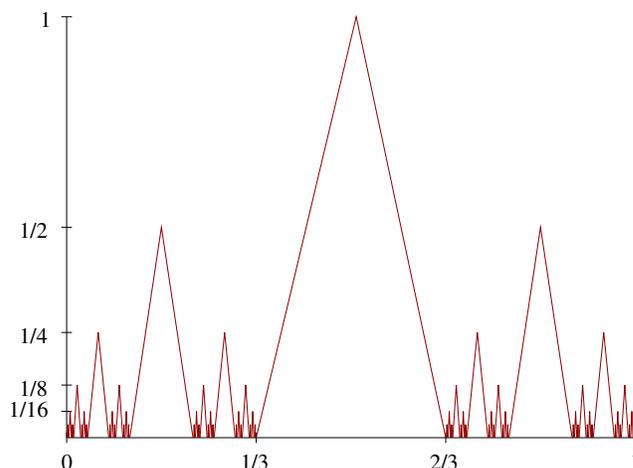


Abbildung 6.2: Eine nicht rektifizierbare Kurve.

Lemma 6.1.3. Sei $I = [a, b]$ ein kompaktes Intervall und sei $\gamma : I \rightarrow \mathbb{R}^n$ eine stetig differenzierbare Abbildung. Dann ist γ rektifizierbar und es gilt

$$L(\gamma) = \int_a^b \|\dot{\gamma}(t)\| dt. \quad (6.1.3)$$

Beweis. Der Beweis hat zwei Schritte.

Schritt 1. Die Kurve γ ist rektifizierbar und es gilt $L(\gamma) \leq \int_a^b \|\dot{\gamma}(t)\| dt$.

Sei $P = (t_0, t_1, \dots, t_N) \in \mathcal{P}(I)$ eine Partition. Dann gilt nach Teil (iii) von Lemma 2.4.4 für jedes $i \in \{1, \dots, N\}$ die Ungleichung

$$\|\gamma(t_i) - \gamma(t_{i-1})\| = \left\| \int_{t_{i-1}}^{t_i} \dot{\gamma}(t) dt \right\| \leq \int_{t_{i-1}}^{t_i} \|\dot{\gamma}(t)\| dt$$

und daraus folgt

$$L(\gamma, P) = \sum_{i=1}^N \|\gamma(t_i) - \gamma(t_{i-1})\| \leq \sum_{i=1}^N \int_{t_{i-1}}^{t_i} \|\dot{\gamma}(t)\| dt = \int_a^b \|\dot{\gamma}(t)\| dt.$$

Betrachten wir nun das Supremum über alle $P \in \mathcal{P}(I)$, so ergibt sich

$$L(\gamma) = \sup_{P \in \mathcal{P}(I)} L(\gamma, P) \leq \int_a^b \|\dot{\gamma}(t)\| dt < \infty.$$

Damit ist Schritt 1 bewiesen.

Schritt 2. Die Länge von γ ist durch (6.1.3) gegeben.

Sei $\varepsilon > 0$ gegeben. Da $\dot{\gamma} : I \rightarrow \mathbb{R}^n$ gleichmäßig stetig ist, existiert eine Konstante $\delta > 0$, so dass für alle $s, t \in I$ gilt:

$$|s - t| < \delta \quad \implies \quad \|\dot{\gamma}(s) - \dot{\gamma}(t)\| < \frac{\varepsilon}{b - a}.$$

Nun sei $P = (t_0, t_1, \dots, t_N) \in \mathcal{P}(I)$ eine Partition, welche die Ungleichung

$$0 < t_{i+1} - t_i < \delta$$

für $i = 1, \dots, N$ erfüllt. Dann erhalten wir für alle $i \in \{1, \dots, N\}$ und alle $t \in [t_{i-1}, t_i]$ die Ungleichung

$$\begin{aligned} \left\| \dot{\gamma}(t) - \frac{\gamma(t_i) - \gamma(t_{i-1})}{t_i - t_{i-1}} \right\| &= \left\| \frac{1}{t_i - t_{i-1}} \int_{t_{i-1}}^{t_i} (\dot{\gamma}(t) - \dot{\gamma}(s)) \, ds \right\| \\ &\leq \frac{1}{t_i - t_{i-1}} \int_{t_{i-1}}^{t_i} \|\dot{\gamma}(t) - \dot{\gamma}(s)\| \, ds \\ &\leq \frac{\varepsilon}{b - a}. \end{aligned}$$

Daraus folgt die Ungleichung

$$\|\dot{\gamma}(t)\| \leq \frac{\|\gamma(t_i) - \gamma(t_{i-1})\|}{t_i - t_{i-1}} + \frac{\varepsilon}{b - a}$$

für $i = 1, \dots, N$ und $t_{i-1} \leq t \leq t_i$. Integrieren wir diese Ungleichung über das Intervall $[t_{i-1}, t_i]$ und bilden die Summe über $i = 1, \dots, N$ so erhalten wir

$$\begin{aligned} \int_a^b \|\dot{\gamma}(t)\| \, dt &= \sum_{i=1}^N \int_{t_{i-1}}^{t_i} \|\dot{\gamma}(t)\| \, dt \\ &\leq \sum_{i=1}^N \left(\|\gamma(t_i) - \gamma(t_{i-1})\| + \varepsilon \frac{t_i - t_{i-1}}{b - a} \right) \\ &= L(\gamma, P) + \varepsilon \\ &\leq L(\gamma) + \varepsilon. \end{aligned}$$

Also gilt $\int_a^b \|\dot{\gamma}(t)\| \, dt \leq L(\gamma) + \varepsilon$ für alle $\varepsilon > 0$ und daher

$$\int_a^b \|\dot{\gamma}(t)\| \, dt \leq L(\gamma).$$

Nach Schritt 1 folgt daraus $\int_a^b \|\dot{\gamma}(t)\| \, dt = L(\gamma)$. Damit sind Schritt 2 und Lemma 6.1.3 bewiesen. \square

Lemma 6.1.4. *Seien $I = [a, b]$ und $I' = [a', b']$ zwei kompakte Intervalle, sei $\phi : I' \rightarrow I$ ein Homöomorphismus, und sei $\gamma : I \rightarrow \mathbb{R}^n$ eine stetige Abbildung. Dann gilt*

$$L(\gamma \circ \phi) = L(\gamma). \quad (6.1.4)$$

Beweis. Sei $P = (t_0, t_1, \dots, t_N) \in \mathcal{P}(I)$ eine Partition, die (6.1.1) erfüllt. Wir definieren $\phi^*P := (\phi^{-1}(t_0), \phi^{-1}(t_1), \dots, \phi^{-1}(t_N))$ falls ϕ strikt monoton wachsend ist und $\phi^*P := (\phi^{-1}(t_N), \phi^{-1}(t_{N-1}), \dots, \phi^{-1}(t_0))$ falls ϕ strikt monoton fallend ist. Damit ist $\phi^*P \in \mathcal{P}(I')$ und $L(\gamma \circ \phi, \phi^*P) = L(\gamma, P)$. Da jede Partition $P' \in \mathcal{P}(I')$ die Form $P' = \phi^*P$ für ein $P \in \mathcal{P}(I)$ hat, folgt daraus

$$L(\gamma \circ \phi) = \sup_{P \in \mathcal{P}(I)} L(\gamma \circ \phi, \phi^*P) = \sup_{P \in \mathcal{P}(I)} L(\gamma, P) = L(\gamma).$$

Damit ist Lemma 6.1.4 bewiesen. \square

Das d -dimensionale Volumen eines Parallelepipeds

Seien $a = (a_1, \dots, a_n)$ und $b = (b_1, \dots, b_n)$ zwei linear unabhängige Vektoren im \mathbb{R}^n . Das von ihnen aufgespannte Parallelogramm ist die Menge

$$P(a, b) := \{sa + tb \mid s, t \in \mathbb{R}, 0 \leq s, t \leq 1\}$$

und ihr Flächeninhalt ist die Zahl

$$v_2(P(a, b)) := \sqrt{\|a\|^2 \|b\|^2 - \langle a, b \rangle^2}. \quad (6.1.5)$$

Diese Formel behält ihre Gültigkeit wenn a und b linear abhängig sind. In diesem Fall gilt $\langle a, b \rangle = \pm \|a\| \|b\|$ und der Flächeninhalt ist $v_2(P(a, b)) = 0$. Im Allgemeinen betrachten wir, für eine natürliche Zahl $0 \leq d \leq n$ und d Vektoren $a_1, \dots, a_d \in \mathbb{R}^n$ das von diesen Vektoren aufgespannte Parallelepipeds

$$P(a_1, \dots, a_d) := \left\{ \sum_{i=1}^d t_i a_i \mid t_i \in \mathbb{R}, 0 \leq t_i \leq 1 \text{ für } i = 1, \dots, d \right\} \quad (6.1.6)$$

und suchen nach einer Formel für dessen d -dimensionales Volumen. Mit anderen Worten, wir suchen nach einer Funktion $v_d : (\mathbb{R}^n)^d \rightarrow [0, \infty)$, welche jedem d -Tupel von Vektoren $(a_1, \dots, a_d) \in (\mathbb{R}^n)^d$ das d -dimensionale Volumen des Parallelepipeds $P(a_1, \dots, a_d)$ zuordnet. Diese Funktion wird durch die Axiome (V1-3) im folgenden Lemma charakterisiert.

Lemma 6.1.5. Seien $d, n \in \mathbb{N}$ mit $1 \leq d \leq n$. Dann existiert genau eine Funktion $v_d : (\mathbb{R}^n)^d \rightarrow [0, \infty)$, die folgende drei Bedingungen erfüllt.

(V1) Für alle $a_1, \dots, a_d \in \mathbb{R}^n$ und alle $\lambda_1, \dots, \lambda_d \in \mathbb{R}$ gilt

$$v_d(\lambda_1 a_1, \dots, \lambda_d a_d) = \left(\prod_{i=1}^d |\lambda_i| \right) v_d(a_1, \dots, a_d). \quad (6.1.7)$$

(V2) Für alle $a_1, \dots, a_d \in \mathbb{R}^n$ und alle $i, j \in \{1, \dots, d\}$ mit $i \neq j$ gilt

$$v_d(a_1, \dots, a_{i-1}, a_i + a_j, a_{i+1}, \dots, a_d) = v_d(a_1, \dots, a_d). \quad (6.1.8)$$

(V3) Für alle $a_1, \dots, a_d \in \mathbb{R}^n$ gilt

$$\langle a_i, a_j \rangle = \delta_{ij} = \begin{cases} 1, & \text{falls } i = j, \\ 0, & \text{falls } i \neq j \end{cases} \implies v_d(a_1, \dots, a_d) = 1. \quad (6.1.9)$$

Die Abbildung v_d ist durch die Formel

$$v_d(a_1, \dots, a_d) = \sqrt{\det(A^T A)} \quad (6.1.10)$$

gegeben, wobei $A := (a_1 \cdots a_d) \in \mathbb{R}^{n \times d}$ die Matrix mit den Spalten a_1, \dots, a_d ist, das heißt $A^T A = (\langle a_i, a_j \rangle)_{i,j=1}^d \in \mathbb{R}^{d \times d}$.

Beweis. Für $d = 1$ erfüllt die durch $v_1(a) := \|a\|$ für $a \in \mathbb{R}^n$ definierte Funktion $v_1 : \mathbb{R}^n \rightarrow [0, \infty)$ die Axiome (V1) und (V3) und ist durch diese eindeutig bestimmt. Das Axiom (V2) ist in diesem Fall gegenstandslos. Für $d = 2$ stimmt die Formel (6.1.10) mit dem Flächeninhalt (6.1.5) überein. Im Allgemeinen folgt es direkt aus den Eigenschaften der Determinante, dass die durch (6.1.10) definierte Funktion $v_d : (\mathbb{R}^n)^d \rightarrow [0, \infty)$ das Skalierungsaxiom (V1), das Scherungsaxiom (V2) und das Normalisierungsaxiom (V3) erfüllt und durch diese Axiome auch eindeutig bestimmt ist. \square

Definition 6.1.6. Für $a_1, \dots, a_d \in \mathbb{R}^n$ wird die Zahl $v_d(a_1, \dots, a_d)$ in Lemma 6.1.5 das **d -dimensionale Volumen** des Parallelepipeds $P(a_1, \dots, a_d)$ in (6.1.6) genannt und mit

$$\text{Vol}_d(P(a_1, \dots, a_d)) := v_d(a_1, \dots, a_d) \quad (6.1.11)$$

bezeichnet.

Lemma 6.1.7. Sei $n \geq 2$ eine ganze Zahl und seien $a_1, \dots, a_{n-1} \in \mathbb{R}^n$. Definiere $A = (a_1 \cdots a_{n-1}) \in \mathbb{R}^{n \times (n-1)}$ und für $k = 1, \dots, n$ sei $A_k \in \mathbb{R}^{(n-1) \times (n-1)}$ die quadratische Matrix, die aus A durch Entfernung der k -ten Zeile hervorgeht. Dann gilt

$$v_d(a_1, \dots, a_{n-1}) = \sqrt{\sum_{k=1}^n \alpha_k^2}, \quad \alpha_k := (-1)^{k-1} \det(A_k). \quad (6.1.12)$$

Beweis. Sei $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$. Dann gilt für alle $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ nach Teil (iv) von Satz D.3.5 die Gleichung

$$\det(x, a_1, \dots, a_{n-1}) = \sum_{k=1}^n (-1)^{k-1} x_k \det(A_k) = \sum_{k=1}^n x_k \alpha_k = \langle x, \alpha \rangle.$$

Mit $x = a_i$ folgt daraus $\langle a_i, \alpha \rangle = 0$ für $i = 1, \dots, n$. Daraus wiederum folgt, dass die Matrix $B := (\alpha \ a_1 \cdots a_{n-1}) \in \mathbb{R}^{n \times n}$ die Gleichungen

$$B^T B = \begin{pmatrix} \|\alpha\|^2 & 0 \\ 0 & A^T A \end{pmatrix}, \quad \det(B) = \|\alpha\|^2$$

erfüllt und daher gilt $\|\alpha\|^4 = \det(B)^2 = \det(B^T B) = \|\alpha\|^2 \det(A^T A)$. Daraus folgt $\det(A^T A) = \|\alpha\|^2$ und daher, nach (6.1.9) in Lemma 6.1.5,

$$v_d(a_1, \dots, a_{n-1}) = \sqrt{\det(A^T A)} = \|\alpha\| = \sqrt{\alpha_1^2 + \cdots + \alpha_n^2}.$$

Damit ist Lemma 6.1.7 bewiesen. \square

Lemma 6.1.8. Sei $A \in \mathbb{R}^{n \times d}$ und sei $Q = [0, \lambda_1] \times \cdots \times [0, \lambda_d]$ mit $\lambda_i > 0$. Dann hat das Parallelepiped $AQ := \{Ax \mid x \in Q\}$ das d -dimensionale Volumen

$$\text{Vol}_d(AQ) = \sqrt{\det(A^T A)} \cdot \text{Vol}_d(Q). \quad (6.1.13)$$

Beweis. Seien $a_1, \dots, a_d \in \mathbb{R}^n$ die Spalten der Matrix $A \in \mathbb{R}^{n \times d}$. Dann ist

$$AQ = \left\{ \sum_{i=1}^d t_i a_i \mid 0 \leq t_i \leq \lambda_i \text{ für } i = 1, \dots, d \right\} = P(\lambda_1 a_1, \dots, \lambda_d a_d)$$

und daher gilt

$$\text{Vol}_d(AQ) = v_d(\lambda_1 a_1, \dots, \lambda_d a_d) = \lambda_1 \cdots \lambda_d \cdot v_d(a_1, \dots, a_d) = \text{Vol}_d(Q) \sqrt{\det(A^T A)}.$$

Damit ist Lemma 6.1.8 bewiesen. \square

Integration über Kartengebiete

Das nächste Ziel ist es, eine reellwertige stetige Funktion auf einer Untermannigfaltigkeit eines Euklidischen Raumes über ein Kartengebiet zu integrieren. Hierzu benötigen wir die folgenden Begriffe.

Definition 6.1.9 (Kartengebiet). Sei $M \subset \mathbb{R}^n$ eine d -dimensionale C^1 -Untermannigfaltigkeit, seien $U, W \subset \mathbb{R}^n$ offene Teilmengen, und sei

$$\phi : U \rightarrow W$$

ein C^1 -Diffeomorphismus mit

$$\phi(U \cap M) = W \cap (\mathbb{R}^d \times \{0\}),$$

das heißt ϕ ist eine Karte auf M . In dieser Situation wird die Menge $U \cap M$ **Kartengebiet** genannt. Jedes solche Kartengebiet ist also eine offene Teilmenge von M bezüglich der Relativtopologie (Abschnitt B.1). Die durch

$$\begin{aligned} V &:= \{x = (x_1, \dots, x_d) \in \mathbb{R}^d \mid (x_1, \dots, x_d, 0, \dots, 0) \in W\}, \\ \psi(x_1, \dots, x_d) &:= \phi^{-1}(x_1, \dots, x_d, 0, \dots, 0), \end{aligned} \quad (6.1.14)$$

für $(x_1, \dots, x_d) \in \Omega$ definierte Abbildung

$$\psi : V \rightarrow U \cap M$$

wird **C^1 -Parametrisierung** des Kartengebiets $U \cap M$ genannt. Jede solche C^1 -Parametrisierung ist also ein Homöomorphismus von einer offenen Teilmenge $V \subset \mathbb{R}^d$ auf ein Kartengebiet $U \cap M$ (mit $U \subset \mathbb{R}^n$ offen), ist stetig differenzierbar als Abbildung von V nach \mathbb{R}^n , und die Umkehrabbildung $\psi^{-1} : U \cap M \rightarrow V$ läßt sich zu einer stetig differenzierbaren Abbildung von U nach \mathbb{R}^d fortsetzen.

Sei $\psi : V \rightarrow U \cap M$ eine C^1 -Parametrisierung eines Kartengebiets in einer d -dimensionalen C^1 -Untermannigfaltigkeit $M \subset \mathbb{R}^n$ wie in Definition 6.1.9. Für geeignete Funktionen $f : M \rightarrow \mathbb{R}$ und geeignete Teilmengen $B \subset U \cap M$ wollen wir nun das Integral von f über B definieren. Wir werden es mit

$$\int_B f \, dS$$

bezeichnen.

Wir betrachten zunächst eine Menge der Form $B = \psi(A) \subset U \cap M$, wobei $A \subset \Omega$ ein Quadergebäude ist (Definition 5.6.1). Dann besitzt A eine Partition nach Übung 5.1.4. Laut Definition 5.1.1 existieren also achsenparallele abgeschlossene Quader $Q_1, \dots, Q_\ell \subset \Omega$ mit

$$A = \bigcup_{i=1}^{\ell} Q_i, \quad \overset{\circ}{Q}_i \cap \overset{\circ}{Q}_j = \emptyset \text{ für } i \neq j.$$

Die folgende heuristische Betrachtung motiviert die Definition des Integrals einer stetigen Funktion $f : M \rightarrow \mathbb{R}$ über B . Werden die Q_i mit sehr kleinen Durchmessern gewählt, und ist x_i der Mittelpunkt des Quaders Q_i , so weicht die Bildmenge $\psi(Q_i)$ nur wenig von der Bildmenge $\psi(x_i) + d\psi(x_i)(Q_i - x_i)$ desselben Quaders unter der Abbildung $x \mapsto \psi(x_i) + d\psi(x_i)(x - x_i)$ ab, denn diese affine Abbildung ist nach Definition der Ableitung eine *gute* Näherung an ψ in der Nähe von x_i . Damit ist das Volumen des Parallelepipeds $d\psi(x_i)Q_i$ eine gute Näherung für das noch zu definierende Volumen der tatsächlichen Bildmenge $\psi(Q_i)$. Dieser Näherungswert hat die Form

$$\begin{aligned} \text{Vol}_d(\psi(Q_i)) &\sim \text{Vol}_d(\psi(x_i) + d\psi(x_i)(Q_i - x_i)) \\ &= \text{Vol}_d(d\psi(x_i)Q_i) \\ &= \sqrt{\det(d\psi(x_i)^T d\psi(x_i))} \text{Vol}_d(Q_i). \end{aligned}$$

Hier folgt die letzte Gleichheit aus Lemma 6.1.8. Ersetzen wir nun das noch zu definierende Volumen $\text{Vol}_d(\psi(Q_i))$ durch diesen Näherungswert, dann ergibt sich für das gesuchte Integral nach dem Prinzip *“Das Integral ist angenähert gleich der Summe der Produkte von Grundfläche und Höhe über die Teilgebiete einer feinen Partition”* die heuristische Formel

$$\begin{aligned} \int_B f \, dS &\sim \sum_{i=1}^{\ell} f(\psi(x_i)) \text{Vol}_d(\psi(Q_i)) \\ &\sim \sum_{i=1}^{\ell} f(\psi(x_i)) \sqrt{\det(d\psi(x_i)^T d\psi(x_i))} \text{Vol}_d(Q_i) \\ &\sim \int_A f(\psi(x)) \sqrt{\det(d\psi(x)^T d\psi(x))} \, dx_1 \cdots dx_d. \end{aligned}$$

Hier ist der letzte Term über die Konvergenz der Riemannschen Summen in Satz 5.3.1 zu verstehen. Genau diesen letzten Term werden wir nun für die Definition des Integrals verwenden.

Definition 6.1.10 (Integral über ein Kartengebiet). Sei $M \subset \mathbb{R}^n$ eine d -dimensionale C^1 -Untermannigfaltigkeit, sei $U \subset \mathbb{R}^n$ eine offene Menge so dass $U \cap M$ ein Kartengebiet ist, sei $V \subset \mathbb{R}^d$ eine offenen Menge, sei

$$\psi : V \rightarrow U \cap M$$

eine C^1 -Parametrisierung des Kartengebiets $U \cap M$, sei $A \subset \Omega$ eine Jordanmessbare Menge (Definition 5.5.1), sei

$$B := \psi(A),$$

und sei $f : M \rightarrow \mathbb{R}$ eine Funktion mit der Eigenschaft, dass die Komposition

$$f \circ \psi : A \rightarrow \mathbb{R}$$

Riemann-integrierbar ist. Sei $g^\psi : \Omega \rightarrow \mathbb{R}$ die stetige Funktion

$$g^\psi(x) := \sqrt{\det(d\psi(x)^T d\psi(x))} \quad \text{für } x \in \Omega. \quad (6.1.15)$$

Nach Satz 5.2.2 und Satz 5.3.4 ist die Funktion

$$(f \circ \psi)g^\psi : A \rightarrow \mathbb{R}$$

Riemann-integrierbar (Definition 5.5.2). Unter diesen Voraussetzungen definieren wir das **Integral von f über B** durch

$$\int_B f \, dS := \int_A f(\psi(x))g^\psi(x) \, dx_1 \cdots dx_d \quad (6.1.16)$$

und das **d -dimensionale Volumen von B** durch

$$\text{Vol}_d(B) := \int_B 1 \, dS = \int_A g^\psi(x) \, dx_1 \cdots dx_d. \quad (6.1.17)$$

Das folgende Lemma rechtfertigt die Bezeichnung $\int_B f \, dS$ für das Integral der Funktion $f : M \rightarrow \mathbb{R}$ über der kompakten Teilmenge $B \subset M$. Denn diese Bezeichnung suggeriert, dass das Integral tatsächlich nur von B und f abhängt. Allein aus der Definition heraus ist diese Tatsache jedoch nicht unmittelbar ersichtlich. Denn für die Definition wurde ein Kartengebiet $U \cap M$ mit $B \subset U \cap M$ sowie eine Parametrisierung $\psi : V \rightarrow U \cap M$ dieses Kartengebiets gewählt. Es ist nun zu zeigen, dass die rechte Seite der Gleichung (6.1.16) nicht von diesen *Zusatzdaten* V, U, ψ abhängt, die für die Definition verwendet wurden.

Lemma 6.1.11. *Das Integral (6.1.16) in Definition 6.1.10 ist unabhängig von der Wahl von V, U, ψ mit $B \subset U \cap M$.*

Beweis. Sei $U' \subset \mathbb{R}^n$ eine offene Menge, so dass $U' \cap M$ ein Kartengebiet ist mit $B \subset U' \cap M$. Weiterhin sei $V' \subset \mathbb{R}^d$ eine offene Menge und

$$\psi' : V' \rightarrow U' \cap M$$

eine C^1 -Parametrisierung des Kartengebiets $U' \cap M$. Dann sind

$$\tilde{V} := \psi^{-1}(U \cap U' \cap M), \quad \tilde{V}' := (\psi')^{-1}(U \cap U' \cap M)$$

offene Teilmengen von \mathbb{R}^d und die Abbildungen

$$\psi : \tilde{V} \rightarrow U \cap U' \cap M, \quad \psi' : \tilde{V}' \rightarrow U \cap U' \cap M$$

sind Homöomorphismen. Außerdem ist die Abbildung

$$\phi := \psi^{-1} \circ \psi' : \tilde{V}' \rightarrow \tilde{V}$$

ein C^1 -Diffeomorphismus, nach Definition 6.1.9.

Nun folgt aus Lemma 5.7.12, dass die Menge

$$A' := (\psi')^{-1}(B) = (\psi \circ \phi)^{-1}(B) = \phi^{-1}(A) \subset \tilde{\Omega}'$$

Jordan-messbar ist. Da $\psi \circ \phi = \psi'$ ist, folgt aus der Kettenregel die Formel

$$d\psi'(x') = d\psi(\phi(x'))d\phi(x') \quad \text{für alle } x' \in \tilde{V}'$$

und daraus folgt

$$g^{\psi'}(x') = g^{\psi}(\phi(x')) |\det(d\phi(x'))| \quad \text{für alle } x' \in \tilde{V}'.$$

Also ergibt sich aus der Transformationsformel in Satz 5.7.11 die Gleichung

$$\begin{aligned} \int_{A'} (f \circ \psi') g^{\psi'} &= \int_{A'} (f \circ \psi \circ \phi)(g^{\psi \circ \phi}) |\det(d\phi)| \\ &= \int_{\phi(A')} (f \circ \psi) g^{\psi} \\ &= \int_A (f \circ \psi) g^{\psi}. \end{aligned}$$

Damit ist Lemma 6.1.11 bewiesen. \square

6.2 Definition des Integrals

Sei $M \subset \mathbb{R}^n$ eine d -dimensionale C^1 -Untermannigfaltigkeit. Unser nächstes Ziel ist es, das Integral einer stetigen Funktion $f : M \rightarrow \mathbb{R}$ über einer kompakten Teilmenge von M zu definieren, die nicht in einem Kartengebiet enthalten ist. Dies erfordert das Konzept einer Jordan-messbaren Teilmenge $B \subset M$, das heißt, einer kompakten Teilmenge, die *relativ zu M* Jordan-messbar ist, was wiederum bedeutet, dass ihr Rand $\partial_M B$ bezüglich der Relativtopologie von M eine d -dimensionale Jordansche Nullmenge ist. Des weiteren benötigen wir das Konzept einer *Partition der Eins*, welches es uns erlaubt, die Funktion f als eine Summe von Funktionen darzustellen, die jeweils auf nur einem einzigen Kartengebiet von Null verschieden sind.

Jordan-messbare Teilmengen von M

Definition 6.2.1. Sei $M \subset \mathbb{R}^n$ eine d -dimensionale C^1 -Untermannigfaltigkeit. Eine kompakte Teilmenge $N \subset M$ wird **d -dimensionale Jordansche Nullmenge**, oder auch **Jordansche Nullmenge bezüglich M** genannt, wenn für jede C^1 -Parametrisierung

$$\psi : V \rightarrow U \cap M$$

eines Kartengebiets $U \cap M$ (mit $U \subset \mathbb{R}^n$ und $V \subset \mathbb{R}^d$ offen) und jede kompakte Menge $K \subset V$, die Menge

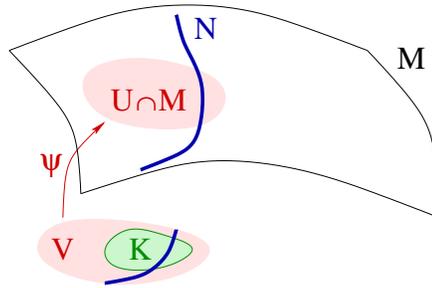
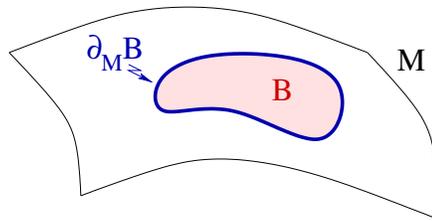
$$K \cap \psi^{-1}(N) = \{x \in K \mid \psi(x) \in N\} \subset \mathbb{R}^d$$

eine Jordansche Nullmenge ist (siehe Abbildung 6.3 und Definition 5.2.1). Eine kompakte Teilmenge $B \subset M$ wird **Jordan-messbar bezüglich M** genannt, wenn die Menge

$$\partial_M B := \left\{ x \in M \left| \begin{array}{l} \text{für alle } \varepsilon > 0 \text{ gilt} \\ B_\varepsilon(x) \cap B \neq \emptyset \text{ und} \\ B_\varepsilon(x) \cap (M \setminus B) \neq \emptyset \end{array} \right. \right\} \quad (6.2.1)$$

eine d -dimensionale Jordansche Nullmenge ist. Die Menge $\partial_M B \subset M$ wird **der Rand von B bezüglich M** genannt (siehe Abbildung 6.4).

Bemerkung 6.2.2. Sei $\|\cdot\|$ die Euklidische Norm auf dem \mathbb{R}^n sei die Abstandsfunktion $d_M : M \times M \rightarrow \mathbb{R}$ durch $d_M(x, y) := \|x - y\|$ für $x, y \in M$ definiert. Dann ist der Rand $\partial_M B$ in (6.2.1) einer Teilmenge $B \subset M$ der Rand von B als Teilmenge des metrischen Raumes (M, d_M) .

Abbildung 6.3: Eine d -dimensionale Jordansche Nullmenge.Abbildung 6.4: Der Rand von B bezüglich M .

Lemma 6.2.3. Sei $M \subset \mathbb{R}^n$ eine d -dimensionale C^1 -Untermannigfaltigkeit und sei $B \subset M$ eine kompakte und bezüglich M Jordan-messbare Menge. Seien $U \subset \mathbb{R}^n$ und $V \subset \mathbb{R}^d$ offene Mengen so dass $U \cap M$ ein Kartengebiet ist, sei $\psi : V \rightarrow U \cap M$ eine C^1 -Parametrisierung dieses Kartengebiets, und sei $K \subset V$ eine kompakte Jordan-messbare Menge. Dann ist die Menge

$$A := K \cap \psi^{-1}(B)$$

kompakt und Jordan-messbar.

Beweis. Wir zeigen zunächst, dass A eine kompakte Teilmenge von \mathbb{R}^d ist. Sei $(x_i)_{i \in \mathbb{N}}$ eine Folge in A . Da $A \subset K$ und K kompakt ist, existiert eine Teilfolge $(x_{i_\nu})_{\nu \in \mathbb{N}}$, die gegen ein Element $x := \lim_{\nu \rightarrow \infty} x_{i_\nu} \in K$ konvergiert. Da B kompakt und $(\psi(x_{i_\nu}))_{\nu \in \mathbb{N}}$ eine Folge in B ist, können wir durch Übergang zu einer weiteren Teilfolge (immer noch mit $(x_{i_\nu})_{\nu \in \mathbb{N}}$ bezeichnet) annehmen, dass die Folge $(\psi(x_{i_\nu}))_{\nu \in \mathbb{N}}$ gegen ein Element $p = \lim_{\nu \rightarrow \infty} \psi(x_{i_\nu}) \in B$ konvergiert. Da ψ stetig ist, folgt daraus die Gleichung $p = \psi(\lim_{\nu \rightarrow \infty} x_{i_\nu}) = \psi(x)$. Daher gilt $\psi(x) \in B$ und somit ist $x \in K \cap \psi^{-1}(B) = A$. Wir haben also gezeigt, dass jede Folge in A eine Teilfolge besitzt, die gegen ein Element von A konvergiert. Also ist A kompakt.

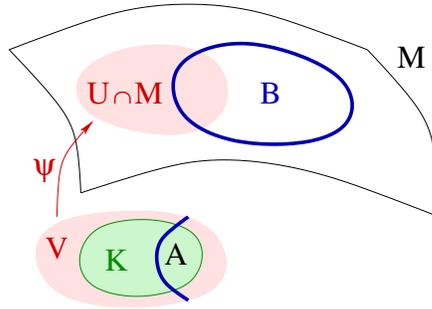


Abbildung 6.5: Eine bezüglich M Jordan-messbare Menge.

Als nächstes zeigen wir die Inklusion

$$\partial A \subset \partial K \cup (K \cap \psi^{-1}(\partial_M B)). \tag{6.2.2}$$

(Siehe Abbildung 6.5.) Da K Jordan-messbar ist, ist ∂K eine Jordansche Nullmenge und, da B bezüglich M Jordan-messbar ist, ist $\partial_M B$ eine d -dimensionale Jordansche Nullmenge, und damit ist $K \cap \psi^{-1}(\partial_M B)$ ebenfalls eine Jordansche Nullmenge. Also ist die Vereinigung dieser beiden Mengen eine Jordansche Nullmenge. Nach (6.2.2) ist daher auch ∂A eine Jordansche Nullmenge, und damit ist A Jordan-messbar.

Es bleibt also zu zeigen, dass (6.2.2) gilt. Sei $x \in \partial A \setminus \partial K$. Da wir bereits gezeigt haben, dass A eine kompakte Teilmenge von \mathbb{R}^d ist, wissen wir auch, dass A abgeschlossen ist und daraus folgt $\partial A \subset A \subset K$. Daher gilt

$$x \in K \setminus \partial K = \overset{\circ}{K}.$$

Außerdem existiert eine Folge $x_i \in \mathbb{R}^d \setminus A$, die gegen x konvergiert. Da ihr Grenzwert x ein Element der offenen Menge $\overset{\circ}{K}$ ist, gilt $x_i \in \overset{\circ}{K}$ für jeden hinreichend großen Index $i \in \mathbb{N}$. Wir können also ohne Beschränkung der Allgemeinheit annehmen, dass $x_i \in \overset{\circ}{K} \setminus A$ ist für alle $i \in \mathbb{N}$. Dann gilt

$$\psi(x_i) \notin B$$

für alle $i \in \mathbb{N}$ nach Definition der Menge A . Ferner gilt

$$\lim_{i \rightarrow \infty} \psi(x_i) = \psi(x) \in B,$$

da $x \in \partial A \subset A$ ist. Da $\psi(x_i) \in M \setminus B$ ist, folgt daraus $\psi(x) \in \partial_M B$ und daher $x \in \psi^{-1}(\partial_M B) \cap K$ ist. Wir haben also die Inklusion

$$\partial A \setminus \partial K \subset K \cap \psi^{-1}(\partial_M B)$$

bewiesen. Daraus folgt (6.2.2) und damit ist Lemma 6.2.3 bewiesen. □

Partition der Eins

Der **Träger** einer Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ist die Menge

$$\text{supp}(f) := \overline{\{x \in \mathbb{R}^n \mid f(x) \neq 0\}}.$$

Der Träger ist also der Abschluss der Menge aller Punkte in \mathbb{R}^n an denen die Funktion f nicht verschwindet.

Satz 6.2.4 (Partition der Eins). *Sei $K \subset \mathbb{R}^n$ eine kompakte Menge und seien $U_1, U_2, \dots, U_\ell \subset \mathbb{R}^n$ offene Mengen, die K überdecken, das heißt*

$$K \subset U_1 \cup U_2 \cup \dots \cup U_\ell. \quad (6.2.3)$$

Dann existieren glatte Funktionen $\rho_1, \rho_2, \dots, \rho_\ell : \mathbb{R}^n \rightarrow [0, 1]$, so dass folgendes gilt:

$$\text{supp}(\rho_i) \subset U_i, \quad \text{supp}(\rho_i) \text{ ist kompakt} \quad \text{für } i = 1, 2, \dots, \ell, \quad (6.2.4)$$

$$\sum_{i=1}^{\ell} \rho_i(x) = 1 \quad \text{für alle } x \in K. \quad (6.2.5)$$

Beweis. Für $K = \emptyset$ gilt die Behauptung trivialerweise mit $\rho_i \equiv 0$ für alle i . Sei also $K \neq \emptyset$. Dann beweisen wir die Existenz der ρ_i in vier Schritten.

Schritt 1. *Es gibt eine glatte Funktion $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$, so dass $\phi(x) > 0$ ist für alle $x \in \mathbb{R}^n$ mit $\|x\| < 1$, und $\phi(x) = 0$ ist für alle $x \in \mathbb{R}^n$ mit $\|x\| \geq 1$.*

Eine explizite Formel für eine solche Funktion ist

$$\phi(x) := \begin{cases} e^{-\frac{1}{1-\|x\|^2}}, & \text{falls } \|x\| < 1, \\ 0, & \text{falls } \|x\| \geq 1 \end{cases} \quad (6.2.6)$$

für $x \in \mathbb{R}^n$. Diese Funktion kann in der Form $\phi(x) = g(1 - \|x\|^2)$ geschrieben werden, wobei die Funktion $g : \mathbb{R} \rightarrow \mathbb{R}$ durch die Formel

$$g(s) := \begin{cases} e^{-\frac{1}{s}}, & \text{falls } s > 0, \\ 0, & \text{falls } s \leq 0, \end{cases}$$

für $s \in \mathbb{R}$ definiert ist. Daß g glatt ist, wurde in der Analysis I Vorlesung gezeigt, und damit folgt aus der Kettenregel, dass ϕ ebenfalls eine glatte Funktion ist.

Schritt 2. Es gibt eine natürliche Zahl $m \in \mathbb{N}$, Elemente $x_1, \dots, x_m \in K$, positive reelle Zahlen $\delta_1, \dots, \delta_m$, und Indizes $i_1, \dots, i_m \in \{1, \dots, \ell\}$, so dass

$$K \subset \bigcup_{j=1}^m B_{\delta_j}(x_j), \tag{6.2.7}$$

$$\overline{B_{3\delta_j}(x_j)} \subset U_{i_j} \quad \text{für } j = 1, \dots, m. \tag{6.2.8}$$

Für jedes Element $x \in K$ sei $i(x) \in \{1, \dots, \ell\}$ der kleinste Index mit $x \in U_{i(x)}$ und sei $\delta(x) > 0$ die grösste Zahl in dem Intervall $(0, 1]$ mit $B_{4\delta(x)}(x) \subset U_{i(x)}$. Dann gilt $\overline{B_{3\delta(x)}(x)} \subset U_{i(x)}$ für alle $x \in K$ und die Mengen $B_{\delta(x)}(x)$ für $x \in K$ bilden eine offene Überdeckung von K . Da K kompakt ist, existieren nach Satz C.1.2 endlich viele Elemente $x_1, \dots, x_m \in K$ mit

$$K \subset B_{\delta(x_1)}(x_1) \cup B_{\delta(x_2)}(x_2) \cup \dots \cup B_{\delta(x_m)}(x_m).$$

Mit $\delta_j := \delta(x_j)$ und $i_j := i(x_j)$ für $j = 1, \dots, m$ folgen daraus die Behauptungen (6.2.7) und (6.2.8) in Schritt 2.

Schritt 3. Es existiert eine glatte Funktion $\phi_0 : \mathbb{R}^n \rightarrow [0, \infty)$ mit

$$\begin{aligned} \phi_0(x) &> 0 \quad \text{für alle } x \in \mathbb{R}^n \setminus \overline{\bigcup_{j=1}^m B_{2\delta_j}(x_j)}, \\ \phi_0(x) &= 0 \quad \text{für alle } x \in \bigcup_{j=1}^m B_{\delta_j}(x_j). \end{aligned} \tag{6.2.9}$$

Sei $\delta := \min\{\delta_1, \dots, \delta_m\} > 0$ und definiere $\sigma_0 : \mathbb{R}^n \rightarrow [0, \infty)$ durch

$$\sigma_0(x) := \inf \left\{ \|x - y\| \mid y \in \bigcup_{j=1}^m B_{2\delta_j}(x_j) \right\} \quad \text{für } x \in \mathbb{R}^n.$$

Diese Funktion ist Lipschitz-stetig nach der Dreiecks-Ungleichung, verschwindet auf der Menge $\overline{\bigcup_{j=1}^m B_{2\delta_j}(x_j)}$, und nimmt auf deren Komplement positive Werte an. Sei $\phi : \mathbb{R}^n \rightarrow [0, 1]$ die glatte Funktion in Schritt 1 und definiere

$$\phi_0(x) := \int_{\mathbb{R}^n} \sigma_0(\xi) \phi(\delta^{-1}(x - \xi)) d\xi_1 \cdots d\xi_n \quad \text{für } x \in \mathbb{R}^n.$$

Dann ist ϕ_0 glatt, was man durch partielle Differentiation *unter dem Integral* beweisen kann. Außerdem erfüllt ϕ_0 die Bedingungen (6.2.9) in Schritt 3.

Schritt 4. Wir beweisen Satz 6.2.4.

Für $j = 1, \dots, m$ definieren wir die Funktion $\phi_j : \mathbb{R}^n \rightarrow [0, 1]$ durch

$$\phi_j(x) := \phi\left(\frac{x - x_j}{3\delta_j}\right) \quad \text{für } x \in \mathbb{R}^n.$$

Hier ist $\phi : \mathbb{R}^n \rightarrow [0, 1]$ die glatte Funktion aus Schritt 1. Dann ist ϕ_j eine glatte Funktion mit dem kompakten Träger

$$\text{supp}(\phi_j) = \overline{B_{3\delta_j}(x_j)} \subset U_{i_j}. \quad (6.2.10)$$

Außerdem gilt

$$\sum_{j=0}^m \phi_j(x) > 0 \quad \text{für alle } x \in \mathbb{R}^n. \quad (6.2.11)$$

Denn für $x \in B_{3\delta_j}(x_j)$ mit $j > 0$ gilt $\phi_j(x) > 0$ und für $x \notin \bigcup_{j=1}^m B_{3\delta_j}(x_j)$ gilt

$$x \notin \bigcup_{j=1}^m \overline{B_{2\delta_j}(x_j)} = \overline{\bigcup_{j=1}^m B_{2\delta_j}(x_j)}$$

und in diesem Fall ist $\phi_0(x) > 0$ nach Schritt 3. Damit ist (6.2.11) bewiesen. Nach Schritt 2 gilt $K \subset \bigcup_{j=1}^m B_{\delta_j}(x_j)$ und daraus folgt nach Schritt 3 dass ϕ_0 auf K verschwindet, das heißt

$$\phi_0(x) = 0 \quad \text{für alle } x \in K. \quad (6.2.12)$$

Für $i = 1, \dots, \ell$ sei

$$\begin{aligned} J_i &:= \{j \in \{1, \dots, m\} \mid i_j = i\}, \\ \rho_i(x) &:= \frac{\sum_{j \in J_i} \phi_j(x)}{\sum_{j=0}^m \phi_j(x)} \quad \text{für } x \in \mathbb{R}^n. \end{aligned} \quad (6.2.13)$$

Die Funktion $\rho_i : \mathbb{R}^n \rightarrow [0, 1]$ ist wohldefiniert nach (6.2.11). Außerdem ist sie glatt und hat den Träger

$$\text{supp}(\rho_i) \subset \bigcup_{j \in J_i} \text{supp}(\phi_j) = \bigcup_{j \in J_i} \overline{B_{3\delta_j}(x_j)} \subset U_i.$$

Hier folgen die letzten beiden Schritte aus (6.2.10). Für alle $x \in K$ gilt nach (6.2.12) und (6.2.13) die Gleichung

$$\sum_{i=1}^{\ell} \rho_i(x) = \frac{\sum_{j=1}^m \phi_j(x)}{\sum_{j=0}^m \phi_j(x)} = 1.$$

Damit ist Satz 6.2.4 bewiesen. \square

Definition des Integrals**Definition 6.2.5 (Integral über einer Mannigfaltigkeit).**

Sei $M \subset \mathbb{R}^n$ eine d -dimensionale C^1 -Untermannigfaltigkeit, sei $B \subset M$ eine kompakte bezüglich M Jordan-messbare Menge, sei $N \subset M$ eine Jordansche Nullmenge bezüglich M , und sei $f : M \rightarrow \mathbb{R}$ eine beschränkte Funktion, die auf $M \setminus N$ stetig ist. Dann sind das Riemannsche Integral von f über B und das d -dimensionale Volumen von B wie folgt definiert.

Wähle offene Teilmengen

$$U_1, U_2, \dots, U_\ell \subset \mathbb{R}^n$$

und stetige Funktionen

$$\rho_1, \rho_2, \dots, \rho_\ell : \mathbb{R}^n \rightarrow [0, 1],$$

welche folgende Bedingungen erfüllen.

- (a) Für $i = 1, \dots, \ell$ ist $U_i \cap M$ ein Kartengebiet.
- (b) Für $i = 1, \dots, \ell$ ist $\text{supp}(\rho_i)$ eine kompakte Teilmenge von U_i .
- (c) $B \subset \bigcup_{i=1}^{\ell} U_i$ und $\sum_{i=1}^{\ell} \rho_i(x) = 1$ für alle $x \in B$.

Dann ist das **Riemannsche Integral von f über B** die Zahl

$$\int_B f \, dS := \sum_{i=1}^{\ell} \int_{B \cap U_i} \rho_i f \, dS \quad (6.2.14)$$

und das **d -dimensionale Volumen von B** ist die Zahl

$$\text{Vol}_d(B) := \int_B 1 \, dS = \sum_{i=1}^{\ell} \int_{B \cap U_i} \rho_i \, dS. \quad (6.2.15)$$

In Definition 6.2.5 ist peinlich genau darauf zu achten, dass die Worte “*bezüglich M* ” nicht weggelassen werden, dass es sich also bei N um eine Jordansche Nullmenge **bezüglich M** und bei B um eine **bezüglich M** Jordan-messbare Menge handelt. Ist zum Beispiel M eine kompakte d -dimensionale C^1 -Untermannigfaltigkeit mit $d < n$, so ist M selbst, und damit auch jede Teilmenge von M , nach Beispiel 5.6.13 eine Jordansche Nullmenge in \mathbb{R}^n . Wir können aber im Allgemeinen eine stetige Funktion $f : M \rightarrow \mathbb{R}$ nicht über einer beliebigen Teilmenge von M integrieren.

Bemerkung 6.2.6 (Zur Definition des Integrals). Die Summanden auf der rechten Seite von (6.2.14) sind nach Definition 6.1.10 wie folgt zu verstehen. Da $U_i \cap M$ ein Kartengebiet ist, existiert eine offene Menge $V_i \subset \mathbb{R}^d$ und eine C^1 -Parametrisierung $\psi_i : V_i \rightarrow U_i \cap M$. Da $\text{supp}(\rho_i) \cap B$ eine kompakte Teilmenge von $U_i \cap M$ ist, ist $\psi_i^{-1}(\text{supp}(\rho_i) \cap B)$ eine kompakte Teilmenge von V_i . Daher existiert nach Übung 5.1.5 eine kompakte Jordan-messbare Menge $K_i \subset \mathbb{R}^d$ mit $\psi_i^{-1}(\text{supp}(\rho_i) \cap B) \subset K_i \subset V_i$. Nach Definition 6.2.1 ist die Menge $N_i := K_i \cap \psi_i^{-1}(N)$ eine Jordansche Nullmenge und nach Lemma 6.2.3 ist die Menge $A_i := K_i \cap \psi_i^{-1}(B)$ Jordan-messbar. Daher ist die Menge $B_i := \psi_i(A_i) = \psi_i(K_i) \cap B$ bezüglich M Jordan-messbar. Außerdem gilt $\text{supp}(\rho_i) \cap B \subset B_i$. Die Funktion $((\rho_i f) \circ \psi_i) g^{\psi_i} : A_i \rightarrow \mathbb{R}$ ist beschränkt und ist stetig auf $A_i \setminus N_i$, und ist daher Riemann integrierbar. Der i -te Summand in (6.2.14) ist zu verstehen als das Integral dieser Funktion über A_i :

$$\int_{B \cap U_i} \rho_i f dS = \int_{B_i} \rho_i f dS = \int_{A_i} \rho_i(\psi_i(x)) f(\psi_i(x)) \sqrt{\det(d\psi_i(x)^T d\psi_i(x))} dx.$$

Daß dieses Integral unabhängig von der Wahl der C^1 -Parametrisierung ψ_i und der Menge K_i ist, wurde in Lemma 6.1.11 bewiesen.

Bemerkung 6.2.7 (Das Integral ist wohldefiniert). Wir zeigen, dass die rechte Seite der Gleichung (6.2.14) unabhängig von der Wahl der offenen Mengen U_i und der stetigen Funktionen ρ_i ist. Seien also $V_1, \dots, V_m \subset \mathbb{R}^n$ offene Mengen und $\sigma_1, \dots, \sigma_m : \mathbb{R}^n \rightarrow [0, 1]$ stetige Funktionen, die ebenfalls die Bedingungen (a), (b), und (c) in Definition 6.2.5 erfüllen. Das heißt für jedes $j \in \{1, \dots, m\}$ ist $V_j \cap M$ ein Kartengebiet und $\text{supp}(\rho_j)$ eine kompakte Teilmenge von V_j , die Mengen V_j überdecken B , und es gilt $\sum_{j=1}^m \sigma_j(x) = 1$ für alle $x \in B$. Dann erhalten wir

$$\begin{aligned} \sum_{j=1}^m \int_{B \cap V_j} \sigma_j f dS &= \sum_{j=1}^m \int_{B \cap V_j} \left(\sum_{i=1}^{\ell} \rho_i \right) \sigma_j f dS \\ &= \sum_{i=1}^{\ell} \sum_{j=1}^m \int_{B \cap U_i \cap V_j} \rho_i \sigma_j f dS \\ &= \sum_{i=1}^{\ell} \int_{B \cap U_i} \sum_{j=1}^m \rho_i \sigma_j f dS \\ &= \sum_{i=1}^{\ell} \int_{B \cap U_i} \rho_i f dS. \end{aligned}$$

Bemerkung 6.2.8 (Rechenregeln). Das Integral in Definition 6.2.5 erfüllt die üblichen Rechenregeln für das Integral, wie sie zum Beispiel in Satz 5.3.4 formuliert sind. Hierzu ist es nützlich, die Bezeichnungen

$$\begin{aligned} \mathcal{F}(M) &:= \left\{ f : M \rightarrow \mathbb{R} \left| \begin{array}{l} \text{es existiert eine } d\text{-dimensionale} \\ \text{Jordansche Nullmenge } N \subset M \\ \text{so dass } f \text{ auf } M \setminus N \text{ stetig ist} \end{array} \right. \right\}, \\ \mathcal{B}(M) &:= \left\{ B \subset M \left| \begin{array}{l} B \text{ ist kompakt und} \\ \text{bezüglich } M \text{ Jordan-messbar} \end{array} \right. \right\} \end{aligned} \quad (6.2.16)$$

einzuführen. Dann ist $\mathcal{F}(M)$ ein Vektorraum und die Abbildung

$$\mathcal{F}(M) \rightarrow \mathbb{R} : f \mapsto \int_B f dS$$

ist linear für alle $B \in \mathcal{B}(M)$. Das heißt, für alle $f, g \in \mathcal{F}(M)$, alle $\lambda \in \mathbb{R}$, und alle $B \in \mathcal{B}(M)$ gilt

$$\int_B (f + g) dS = \int_B f dS + \int_B g dS$$

und

$$\int_B \lambda f dS = \lambda \int_B f dS.$$

Außerdem gilt für alle $f \in \mathcal{F}(M)$ und alle $B, C \in \mathcal{B}(M)$ mit

$$\text{Vol}_d(B \cap C) = 0$$

die Gleichung

$$\int_{B \cup C} f dS = \int_B f dS + \int_C f dS.$$

Weiterhin gilt für alle $f \in \mathcal{F}(M)$

$$f \geq 0 \quad \implies \quad \int_B f dS \geq 0$$

und

$$\left| \int_B f dS \right| \leq \int_B |f| dS.$$

Übung: Beweisen Sie diese Aussagen unter Zuhilfenahme von Satz 5.3.4.

Bemerkung 6.2.9 (Uneigentliche Integrale). Sei $M \subset \mathbb{R}^n$ eine nicht kompakte d -dimensionale C^1 -Untermannigfaltigkeit, sei $\mathcal{B}(M) \subset 2^M$ wie in (6.2.16), und sei $f \in \mathcal{F}(M)$ eine Funktion, die auf dem Komplement einer Jordanschen Nullmenge bezüglich M stetig ist. Die Funktion f heißt **(uneigentlich) Riemann-integrierbar** wenn sie der Bedingung

$$\sup_{B \in \mathcal{B}(M)} \int_B |f| dS < \infty \quad (6.2.17)$$

genügt. Dann existiert genau eine reelle Zahl c mit folgenden Eigenschaften.

(i) Für jede Folge $B_i \in \mathcal{B}(M)$ mit $B_i \subset \overset{\circ}{B}_{i+1}$ für alle i und $M = \bigcup_{i=1}^{\infty} B_i$ gilt

$$\lim_{i \rightarrow \infty} \int_{B_i} f dS = c.$$

(ii) Für jedes $\varepsilon > 0$ existiert ein $B_0 \in \mathcal{B}(M)$, so dass für alle $B \in \mathcal{B}(M)$ gilt

$$B_0 \subset B \quad \Longrightarrow \quad \left| c - \int_B f dS \right| < \varepsilon.$$

Die Existenz einer solchen Zahl $c \in \mathbb{R}$ beweist man genauso wie in Satz 5.8.4. Diese Zahl c wird das **(uneigentliche) Integral von f über M** genannt und mit $\int_M f dS := c$ bezeichnet. Das heißt

$$\int_M f dS := \lim_{i \rightarrow \infty} \int_{B_i} f dS \quad (6.2.18)$$

für jede Folge $B_i \in \mathcal{B}(M)$ mit $B_i \subset \overset{\circ}{B}_{i+1}$ für alle i und $M = \bigcup_{i=1}^{\infty} B_i$.

Bemerkung 6.2.10. Sei $M \subset \mathbb{R}^n$ eine kompakte d -dimensional Untermannigfaltigkeit, sei $N \subset M$ eine kompakte Jordansche Nullmenge bezüglich M , und sei $f : M \rightarrow \mathbb{R}$ eine stetige Funktion. Dann gilt

$$\int_M f dS = \int_{M \setminus N} f dS = \lim_{i \rightarrow \infty} \int_{B_i} f dS$$

für jede Folge $B_i \in \mathcal{B}(M)$ mit $B_i \subset \overset{\circ}{B}_{i+1}$ für alle i und $M = \bigcup_{i=1}^{\infty} B_i$. Das heißt, das uneigentliche Integral von f über $M \setminus N$ stimmt mit dem Integral von f über ganz M überein.

6.3 Beispiele

Beispiel 6.3.1 (Summierbare Familien). Sei $M \subset \mathbb{R}^n$ eine 0-dimensionale Untermannigfaltigkeit. Dann ist M **diskret** (für jedes $p \in M$ gibt es ein $\varepsilon > 0$ mit $B_\varepsilon(p) \cap M = \{p\}$), und daher endlich oder abzählbar unendlich. Die Relativtopologie auf M ist die **diskrete Topologie** (jede Teilmenge von M ist offen), jede Teilmenge von M ist Jordan-messbar bezüglich M , und eine Teilmenge ist genau dann kompakt wenn sie endlich ist. Das Integral einer Funktion $f : M \rightarrow \mathbb{R}$ über einer endlichen Teilmenge $B \subset M$ ist

$$\int_B f dS = \sum_{p \in B} f(p).$$

Eine Funktion $f : M \rightarrow \mathbb{R}$ ist also genau dann **absolut summierbar** (das heißt die Menge $\{\sum_{p \in B} |f(p)| \mid B \subset M, \#B < \infty\}$ ist beschränkt) wenn sie uneigentlich Riemann-integrierbar ist, und in dieser Situation ist Teil (ii) in Bemerkung 6.2.9 genau die Aussage von [4, Satz 4.3].

Beispiel 6.3.2 (Der Fall $d = n$). Eine n -dimensionale C^1 -Untermannigfaltigkeit des \mathbb{R}^n ist eine offene Teilmenge $M \subset \mathbb{R}^n$. Dann ist $V = U \cap M$ für jede offene Teilmenge $U \subset \mathbb{R}^n$ ein Kartengebiet mit $\psi = \text{id} : V \rightarrow U \cap M$ als C^1 -Parametrisierung. Eine kompakte Teilmenge $B \subset M$ ist in dieser Situation genau dann Jordan-messbar bezüglich M , wenn sie Jordan-messbar in \mathbb{R}^n ist, und das Integral einer stetigen Funktion $f : M \rightarrow \mathbb{R}$ über B in Definition 6.1.10 stimmt mit dem Integral in Definition 5.1.10 überein. Die Unabhängigkeit des Integrals von der Wahl der C^1 -Parametrisierung ist genau die Transformationsformel in Satz 5.7.11.

Beispiel 6.3.3 (Kurvenintegrale). Sei $M \subset \mathbb{R}^n$ eine 1-dimensionale C^1 -Untermannigfaltigkeit und sei $U \subset \mathbb{R}^n$ eine offene Menge, so dass $U \cap M$ ein zusammenhängendes Kartengebiet ist. Eine C^1 -Parametrisierung dieses Kartengebiets ist eine injektive C^1 -Abbildung $\gamma : I \rightarrow \mathbb{R}^n$ auf einem offenen Intervall $I \subset \mathbb{R}$ mit $\gamma(I) = U \cap M$ und $\dot{\gamma}(t) \neq 0$ für alle t . Sei $A = [a, b] \subset I$ ein kompaktes Teilintervall. Dann ist $B := \gamma(A)$ bezüglich M Jordan-messbar, die Funktion g^γ in (6.1.15) ist $g^\gamma(t) = \|\dot{\gamma}(t)\|$, und nach (6.1.16) ist das Integral einer stetigen Funktion $f : M \rightarrow \mathbb{R}$ über $B = \gamma([a, b])$ das Kurvenintegral

$$\int_B f dS = \int_a^b f(\gamma(s)) \|\dot{\gamma}(s)\| ds. \quad (6.3.1)$$

Für $f = 1$ ist also $\text{Vol}_1(B) = \int_B 1 dS = L(\gamma|_{[a,b]})$ die Länge der Kurve.

Beispiel 6.3.4 (Der Kreisumfang). Der Einheitskreis

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

ist eine 1-dimensionale glatte Untermannigfaltigkeit des \mathbb{R}^2 . Definiere die Kurve $\gamma : \mathbb{R} \rightarrow \mathbb{R}^2$ durch

$$\gamma(t) := (\cos(t), \sin(t)) \quad \text{für } t \in \mathbb{R}.$$

Diese Kurve ist glatt, erfüllt die Bedingung $\|\dot{\gamma}(t)\| = 1$ für alle $t \in \mathbb{R}$, und ihre Einschränkung auf jedes offene Intervall (a, b) mit $b - a \leq 2\pi$ ist eine glatte Parametrisierung eines Kartengebiets in S^1 . Für $\varepsilon > 0$ ist

$$B_\varepsilon := \gamma([\varepsilon, 2\pi - \varepsilon])$$

eine kompakte, bezüglich S^1 Jordan-messbare Teilmenge von S^1 , und es gilt

$$\bigcup_{\varepsilon > 0} B_\varepsilon = S^1 \setminus \{(1, 0)\}.$$

Ist $f : S^1 \rightarrow \mathbb{R}$ eine stetige Funktion, so gilt

$$\begin{aligned} \int_{S^1} f \, dS &= \int_{S^1 \setminus \{(1,0)\}} f \, dS \\ &= \lim_{\varepsilon \rightarrow 0} \int_{B_\varepsilon} f \, dS \\ &= \lim_{\varepsilon \rightarrow 0} \int_{\varepsilon}^{2\pi - \varepsilon} f(\gamma(t)) \|\dot{\gamma}(t)\| \, dt \\ &= \int_0^{2\pi} f(\gamma(t)) \, dt. \end{aligned}$$

Hier folgt die erste Gleichheit aus Bemerkung 6.2.10, die zweite aus Teil (i) von Bemerkung 6.2.9, die dritte aus Beispiel 6.3.3, und die vierte aus der Tatsache dass $\|\dot{\gamma}(t)\| = 1$ ist für alle t . Daraus folgt die Formel

$$\text{Vol}_1(S^1) = \int_{S^1} 1 \, dS = \int_0^{2\pi} dt = 2\pi$$

für den Umfang des Einheitskreises.

Beispiel 6.3.5 (Integration über Graphen). Sei $V \subset \mathbb{R}^{n-1}$ eine offene Teilmenge und $h : V \rightarrow \mathbb{R}$ eine glatte Funktion. Dann ist

$$M := \{(x, h(x)) \mid x = (x_1, \dots, x_{n-1}) \in V\}$$

eine glatte $(n-1)$ -dimensionale Untermannigfaltigkeit des \mathbb{R}^n . In diesem Fall ist ganz M ein Kartengebiet mit $U := V \times \mathbb{R}$ und $U \cap M = M$. Die Abbildung

$$V \rightarrow M : x \mapsto \psi(x) := (x, h(x))$$

ist eine glatte Paramaterisierung und die Jacobi-Matrix von ψ an der Stelle $x \in V$ ist

$$d\psi(x) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \\ \frac{\partial h}{\partial x_1} & \cdots & \frac{\partial h}{\partial x_{n-2}} & \frac{\partial h}{\partial x_{n-1}} \end{pmatrix} \in \mathbb{R}^{n \times (n-1)}.$$

Die Determinante der Matrix $d\psi(x)^T d\psi(x) \in \mathbb{R}^{(n-1) \times (n-1)}$ ist daher nach Beispiel D.3.14 durch

$$\det(d\psi(x)^T d\psi(x)) = 1 + \sum_{i=1}^{n-1} \left(\frac{\partial h}{\partial x_i}(x) \right)^2 = 1 + \|\nabla h(x)\|^2$$

gegeben. Damit erhalten wir die Gleichungen

$$g^\psi(x) = \sqrt{1 + \|\nabla h(x)\|^2}$$

und

$$\int_M f dS = \int_V f(x, h(x)) \sqrt{1 + \|\nabla h(x)\|^2} dx_1 \cdots dx_{n-1} \quad (6.3.2)$$

für jede Funktion $f : M \rightarrow \mathbb{R}$, die ausserhalb einer Jordanschen Nullmenge bezüglich M stetig und uneigentlich Riemann-integrierbar ist (Bemerkung 6.2.9). Die Mannigfaltigkeit M hat daher das Volumen

$$\text{Vol}_{n-1}(M) = \int_V \sqrt{1 + \|\nabla h(x)\|^2} dx_1 \cdots dx_{n-1}. \quad (6.3.3)$$

Dieses kann auch unendlich sein.

Beispiel 6.3.6 (Integration über Sphären). Für $r > 0$ und $n \in \mathbb{N}$ sei

$$M_r := S_r^{n-1} := \{x \in \mathbb{R}^n \mid \|x\| = r\}$$

die Sphäre vom radius r im \mathbb{R}^n . Dann ist M_r eine glatte $(n-1)$ -dimensionale Untermannigfaltigkeit des \mathbb{R}^n und die Menge

$$U_r := \{x = (x_1, \dots, x_n) \in M_r \mid x_2 = 0 \implies x_1 > 0\}$$

ist ein Kartengebiet, deren Komplement

$$N_r := M_r \setminus U_r = \{x \in M_r \mid x_2 = 0, x_1 \leq 0\}$$

eine Jordansche Nullmenge bezüglich M_r ist. Eine glatte Parametrisierung von U_r ist die Abbildung $\psi_r : V \rightarrow U_r$ mit

$$V := \{\theta \in \mathbb{R}^{n-1} \mid |\theta_1| < \pi, |\theta_i| < \pi/2 \text{ für } i = 2, \dots, n-1\}$$

und

$$\psi_r(\theta) := r \begin{pmatrix} \cos(\theta_1) \cdots \cos(\theta_{n-1}) \\ \sin(\theta_1) \cos(\theta_2) \cdots \cos(\theta_{n-1}) \\ \sin(\theta_2) \cos(\theta_3) \cdots \cos(\theta_{n-1}) \\ \vdots \\ \sin(\theta_{n-2}) \cos(\theta_{n-1}) \\ \sin(\theta_{n-1}) \end{pmatrix} \quad (6.3.4)$$

für $\theta = (\theta_1, \dots, \theta_{n-1}) \in V$. Man beachte, dass $\|\psi_r(\theta)\| = r$ und $\cos(\theta_i) > 0$ ist für $i = 2, \dots, n-1$ und alle $\theta \in V$, woraus sich ergibt, dass die Abbildung $\psi_r : V \rightarrow U_r$ bijektiv ist. Weiterhin sind die partiellen Ableitungen von ψ_r zueinander paarweise orthogonal und haben die Norm

$$\left\| \frac{\partial \psi_r}{\partial \theta_i}(\theta) \right\| = \begin{cases} r \cos(\theta_{i+1}) \cdots \cos(\theta_{n-1}), & \text{für } 1 \leq i \leq n-2, \\ r, & \text{für } i = n-1. \end{cases}$$

Daraus ergibt sich für die Funktion $g^{\psi_r}(\theta) = \sqrt{d\psi_r(\theta)^T d\psi_r(\theta)}$ die Formel

$$g^{\psi_r}(\theta) = r^{n-1} g(\theta), \quad g(\theta) := \cos(\theta_2) \cos^2(\theta_3) \cdots \cos^{n-2}(\theta_{n-1}). \quad (6.3.5)$$

Nach Definition 6.1.10 und Bemerkung 6.2.10 ist daher das Integral einer stetigen Funktion $f : S_r^{n-1} \rightarrow \mathbb{R}$ durch die Formel

$$\begin{aligned}
 \int_{S_r^{n-1}} f \, dS &= \int_{U_r} f \, dS \\
 &= \int_V f(\psi_r(\theta)) r^{n-1} g(\theta) \, d\theta \\
 &= r^{n-1} \int_V f(r\psi_1(\theta)) g(\theta) \, d\theta \\
 &= r^{n-1} \int_{S^{n-1}} f(rx) \, dS(x)
 \end{aligned} \tag{6.3.6}$$

gegeben. Insbesondere ergibt sich mit $f \equiv 1$ die Gleichung

$$\text{Vol}_{n-1}(S_r^{n-1}) = r^{n-1} \text{Vol}_{n-1}(S^{n-1}). \tag{6.3.7}$$

Beispiel 6.3.7 (Das Volumen der Einheitssphäre).

Das Volumen der Einheitssphäre im \mathbb{R}^n wird mit $\omega_n := \text{Vol}_{n-1}(S^{n-1})$ bezeichnet. Aus (6.3.5) und (6.3.6) ergibt sich dafür die explizite Formel

$$\begin{aligned}
 \omega_n &= \text{Vol}_{n-1}(S^{n-1}) \\
 &= \int_V g(\theta) \, d\theta \\
 &= \int_{-\pi/2}^{\pi/2} \cdots \int_{-\pi/2}^{\pi/2} \int_{-\pi}^{\pi} \cos(\theta_2) \cos^2(\theta_3) \cdots \cos^{n-2}(\theta_{n-1}) \, d\theta_1 \cdots d\theta_{n-1} \\
 &= 2\pi \prod_{k=1}^{n-2} \int_{-\pi/2}^{\pi/2} \cos^k(\theta) \, d\theta
 \end{aligned}$$

für $n \geq 3$. Es gilt also

$$\omega_{n+1} = \omega_n \int_{-\pi/2}^{\pi/2} \cos^{n-1}(\theta) \, d\theta \tag{6.3.8}$$

für jede natürliche Zahl $n \geq 2$. Für S^0 haben wir $\omega_1 = 2$ nach Beispiel 6.3.1 und für S^1 gilt $\omega_2 = 2\pi$ nach Beispiel 6.3.4. Daraus ergeben sich nach (6.3.8) die Formeln $\omega_3 = 2\pi \int_{-\pi/2}^{\pi/2} \cos(\theta) \, d\theta = 4\pi$ und $\omega_4 = 4\pi \int_{-\pi/2}^{\pi/2} \cos^2(\theta) \, d\theta = 2\pi^2$. Zusammenfassend haben wir für $1 \leq n \leq 4$ die Werte

$$\omega_1 = 2, \quad \omega_2 = 2\pi, \quad \omega_3 = 4\pi, \quad \omega_4 = 2\pi^2. \tag{6.3.9}$$

Eine allgemeine Formel für ω_n kann man mit Hilfe der Gleichung (6.3.8) durch Vollständige Induktion herleiten. Dazu definieren wir

$$c_n := \int_{-\pi/2}^{\pi/2} \cos^{n-1}(\theta) d\theta$$

für $n \in \mathbb{N}$. Dann gilt $\omega_{n+1} = c_n \omega_n$ für alle $n \in \mathbb{N}$. Für $n \geq 3$ ergibt sich durch partielle Integration mit $f(\theta) = \cos^{n-2}(\theta)$ und $g(\theta) = \sin(\theta)$ die Formel

$$\begin{aligned} c_n &= \int_{-\pi/2}^{\pi/2} \cos^{n-1}(\theta) d\theta \\ &= \int_{-\pi/2}^{\pi/2} f(\theta)g'(\theta) d\theta \\ &= - \int_{-\pi/2}^{\pi/2} f'(\theta)g(\theta) d\theta \\ &= (n-2) \int_{-\pi/2}^{\pi/2} \cos^{n-3}(\theta) \sin^2(\theta) d\theta \\ &= (n-2)(c_{n-2} - c_n). \end{aligned}$$

Daraus folgt

$$c_n = \frac{n-2}{n-1} c_{n-2}$$

für alle $n \geq 3$ und daher, mit $c_1 = \pi$ und $c_2 = 2$,

$$\begin{aligned} c_{2k} &= \frac{2k-2}{2k-1} \frac{2k-4}{2k-3} \cdots \frac{2}{3} 2 = \frac{2^k(k-1)!}{3 \cdot 5 \cdots (2k-1)}, \\ c_{2k+1} &= \frac{2k-1}{2k} \frac{2k-3}{2k-2} \cdots \frac{1}{2} \pi = \frac{3 \cdot 5 \cdots (2k-1)}{2^k k!} \pi. \end{aligned}$$

Daraus folgt $c_{2k+1}c_{2k} = \frac{\pi}{k}$ und $c_{2k}c_{2k-1} = \frac{2\pi}{2k-1}$ und daher gilt $\omega_{2k+2} = \frac{\pi}{k}\omega_{2k}$ und $\omega_{2k+1} = \frac{2\pi}{2k-1}\omega_{2k-1}$. Mit $\omega_1 = 2$ und $\omega_2 = 2\pi$ erhalten wir

$$\omega_{2k} = \frac{2\pi^k}{(k-1)!}, \quad \omega_{2k+1} = \frac{2^{k+1}\pi^k}{3 \cdot 5 \cdots (2k-1)}. \quad (6.3.10)$$

für alle $k \in \mathbb{N}$. Der Vergleich dieser Formel mit (5.6.2) in Beispiel 5.6.15 ergibt die Beziehung $\text{Vol}_n(B^n) = \text{Vol}_{n-1}(S^{n-1})/n$ zwischen dem Volumen des Einheitsballes und der Einheitskugel im \mathbb{R}^n . Diese Beziehung werden wir in Beispiel 6.3.8 noch einmal auf andere und erheblich einfachere Weise herleiten.

Beispiel 6.3.8 (Fubini für Polarkoordinaten).

Sei $\psi_r : V \rightarrow \mathbb{R}^n$ die Parametrisierung eines Kartengebiets auf der Sphäre vom Radius r in (6.3.4) und definiere die Abbildung $\psi : (0, \infty) \times V \rightarrow \mathbb{R}^n$ durch $\psi(r, x) := \psi_r(x)$ für $r > 0$ und $x \in V$. Diese Abbildung ist ein Diffeomorphismus auf die offene Teilmenge $U := \{x \in \mathbb{R}^n \mid x_2 = 0 \implies x_1 > 0\}$. Deren Komplement $N := \mathbb{R}^n \setminus U = \{x \in \mathbb{R}^n \mid x_2 = 0, x_1 \leq 0\}$ ist eine unbeschränkte abgeschlossene Teilmenge. Die Menge N kann zwar wegen der Unbeschränktheit keine Jordansche Nullmenge sein, worauf es jedoch ankommt ist, dass N jede kompakte Teilmenge des \mathbb{R}^n in einer Jordanschen Nullmenge schneidet. Nach Beispiel 6.3.6 hat die Jacobi-Matrix von ψ die Determinante

$$\det(d\psi(r, \theta)) = r^{n-1}g(\theta)$$

für alle $r > 0$ und alle $\theta \in V$, wobei die Funktion $g : V \rightarrow (0, \infty)$ durch (6.3.5) gegeben ist. Seien nun $0 < a < b$ reelle Zahlen und

$$I := [a, b], \quad A := \{x \in \mathbb{R}^n \mid a \leq \|x\| \leq b\}.$$

Dann ergibt sich aus der Gleichung (6.3.6) in Beispiel 6.3.6 mit Hilfe des Satzes von Fubini (Satz 5.4.1) und der Transformationsformel (Satz 5.7.11) für jede stetige Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ die Gleichung

$$\begin{aligned} \int_A f(x) dx &= \int_{I \times V} f(\psi(r, \theta)) \det(d\psi(r, \theta)) dr d\theta \\ &= \int_a^b \left(\int_V f(\psi(r, \theta)) r^{n-1} g(\theta) d\theta \right) dr \\ &= \int_a^b \left(r^{n-1} \int_{S^{n-1}} f(rx) dS(x) \right) dr \\ &= \int_a^b \left(\int_{S_r^{n-1}} f dS \right) dr. \end{aligned} \tag{6.3.11}$$

Mit $f \equiv 1$ und dem Grenzübergang $a \rightarrow 0$ und $b = 1$ ergibt sich aus der Gleichung (6.3.11) die Formel

$$\begin{aligned} \text{Vol}_n(B^n) &= \int_0^1 \text{Vol}_{n-1}(S_r^{n-1}) dr \\ &= \int_0^1 r^{n-1} dr \text{Vol}_{n-1}(S^{n-1}) \\ &= \frac{1}{n} \text{Vol}_{n-1}(S^{n-1}). \end{aligned} \tag{6.3.12}$$

Beispiel 6.3.9 (Sphärenvolumen und Gamma-Funktion).

Ist $f : \mathbb{R}^n \rightarrow \mathbb{R}$ eine stetige Funktion, die auf ganz \mathbb{R}^n uneigentlich Riemann-integrierbar ist, so ergibt sich aus der Gleichung (6.3.11) mit dem Grenzübergang $a \rightarrow 0$ und $b \rightarrow \infty$ die Formel

$$\int_{\mathbb{R}^n} f(x) dx = \int_0^\infty \left(\int_{S_r^{n-1}} f dS \right) dr. \quad (6.3.13)$$

Dieses Integral wird besonders einfach, wenn f durch $f(x) = h(\|x\|)$ für eine Funktion $h : (0, \infty) \rightarrow \mathbb{R}$ gegeben ist, und damit auf jeder Sphäre S_r^{n-1} konstant ist. Dann gilt

$$\int_{\mathbb{R}^n} h(\|x\|) dx = \omega_n \int_0^\infty r^{n-1} h(r) dr. \quad (6.3.14)$$

Für die Funktion $f(x) := e^{-\|x\|^2}$ ergibt die Gleichung (6.3.14) zusammen mit (5.8.11) in Beispiel 5.8.9 die Formel

$$\begin{aligned} \pi^{n/2} &= \int_{\mathbb{R}^n} e^{-\|x\|^2} dx \\ &= \omega_n \int_0^\infty r^{n-1} e^{-r^2} dr \\ &= \frac{\omega_n}{2} \int_0^\infty (r^2)^{\frac{n}{2}-1} e^{-r^2} 2r dr \\ &= \frac{\omega_n}{2} \int_0^\infty s^{\frac{n}{2}-1} e^{-s} ds \\ &= \frac{\omega_n}{2} \Gamma\left(\frac{n}{2}\right) \end{aligned}$$

Daraus folgt

$$\omega_n = \frac{2\pi^{n/2}}{\Gamma(n/2)}. \quad (6.3.15)$$

Verwendet man nun die Rekursionsformel

$$\Gamma(x+1) = x\Gamma(x)$$

für die Gamma-Funktion, sowie die Werte $\Gamma(1) = 1$ und $\Gamma(1/2) = \sqrt{\pi}$, so erhält man einen dritten Beweis für die Formel (6.3.10) für ω_n . (Der erste Beweis ist die Herleitung in Beispiel 6.3.7 und der zweite Beweis folgt aus der Kombination der Gleichung (5.6.2) in Beispiel 5.6.15 mit der Gleichung (6.3.12) in Beispiel 6.3.8.)

Beispiel 6.3.10 (Integration über Rotationsflächen).

Seien $z_0 < z_1$ reelle Zahlen und sei $f : (z_0, z_1) \rightarrow (0, \infty)$ eine glatte Funktion. Dann ist die Rotationsfläche

$$M := \{(x, y, z) \in \mathbb{R}^3 \mid z_0 < z < z_1, x^2 + y^2 = f(z)^2\} \quad (6.3.16)$$

eine 2-dimensionale glatte Untermannigfaltigkeit des \mathbb{R}^3 . Wir zeigen, dass ihr Flächeninhalt durch die Formel

$$\text{Vol}_2(M) = 2\pi \int_{z_0}^{z_1} f(z) \sqrt{1 + f'(z)^2} dz \quad (6.3.17)$$

gegeben ist. Diese Flächeninhalt kann auch unendlich sein.

Zum Beweis betrachten wir die Kartengebiete

$$M^\pm := \{(x, y, z) \in M \mid \pm x > 0\}$$

Dann ist M^+ der Graph der Funktion $h : V \rightarrow \mathbb{R}$ auf dem Gebiet

$$V := \{(y, z) \in \mathbb{R}^2 \mid z_0 < z < z_1, -f(z) < y < f(z)\},$$

die durch

$$h(y, z) := \sqrt{f(z)^2 - y^2}$$

für $(y, z) \in V$ definiert ist. Die partiellen Ableitungen von h sind

$$\frac{\partial h}{\partial y} = -\frac{y}{\sqrt{f(z)^2 - y^2}}, \quad \frac{\partial h}{\partial z} = \frac{f(z)f'(z)}{\sqrt{f(z)^2 - y^2}},$$

und daraus folgt

$$1 + \|\nabla h(y, z)\|^2 = 1 + \frac{y^2 + f(z)^2 f'(z)^2}{f(z)^2 - y^2} = \frac{1 + f'(z)^2}{1 - (y/f(z))^2}.$$

Damit ergibt sich aus der Gleichung (6.3.3) in Beispiel 6.3.5 die Formel

$$\begin{aligned} \text{Vol}_2(M^+) &= \int_V \sqrt{\frac{1 + f'(z)^2}{1 - (y/f(z))^2}} dy dz \\ &= \int_{z_0}^{z_1} \left(\int_{-f(z)}^{f(z)} \frac{dy/f(z)}{\sqrt{1 - (y/f(z))^2}} \right) f(z) \sqrt{1 + f'(z)^2} dz \\ &= \int_{-1}^1 \frac{dt}{\sqrt{1 - t^2}} \int_{z_0}^{z_1} f(z) \sqrt{1 + f'(z)^2} dz. \end{aligned}$$

Mit $t = \sin(\theta)$ für $-\pi/2 < \theta < \pi/2$ ergibt sich die Formel

$$\frac{dt}{d\theta} = \cos(\theta) = \sqrt{1 - \sin^2(\theta)} = \sqrt{1 - t^2}$$

und daraus folgt

$$\int_{-1}^1 \frac{dt}{\sqrt{1 - t^2}} = \int_{-\pi/2}^{\pi/2} d\theta = \pi.$$

Also gilt

$$\text{Vol}_2(M^+) = \pi \int_{z_0}^{z_1} f(z) \sqrt{1 + f'(z)^2} dz.$$

Da $\text{Vol}_2(M^+) = \text{Vol}_2(M^-)$ ist und $M \setminus (M^+ \cup M^-)$ eine 2-dimensionale Jordansche Nullmenge ist, folgt daraus die Gleichung (6.3.17).

Beispiel 6.3.11 (Kugeloberfläche). Betrachten wir in Beispiel 6.3.10 den Spezialfall $f(z) = \sqrt{1 - z^2}$ mit $-1 < z_0 < z_1 < 1$, so ist $M = S(z_0, z_1)$ der von den Meridianen auf den Höhen z_0 und z_1 begrenzte Teil der Oberfläche der Einheitskugel im \mathbb{R}^3 . Die Ableitung von f ist $f'(z) = -z/\sqrt{1 - z^2}$. Daraus folgt $1 + f'(z)^2 = 1/(1 - z^2)$ und daher $f(z)\sqrt{1 + f'(z)^2} = 1$. Damit ergibt sich aus der Gleichung (6.3.17) die Formel

$$\text{Vol}_2(S(z_0, z_1)) = 2\pi(z_1 - z_0). \quad (6.3.18)$$

Das heißt, der Flächeninhalt des Segments $S(z_0, z_1)$ der Kugeloberfläche ist proportional zur Höhendifferenz $z_1 - z_0$ (und dass der Proportionalitätsfaktor dann 2π sein muß, folgt aus der Betrachtung eines Segments in der Nähe des Äquators). Diese Beobachtung geht auf Archimedes zurück. Mit $z_0 = -1$ und $z_1 = 1$ ergibt sich der bekannte Flächeninhalt $\omega_3 = 4\pi$ der Oberfläche der Einheitskugel im \mathbb{R}^3 (Gleichung (6.3.9)).

Beispiel 6.3.12 (Kegeloberfläche). Ein weiterer Spezialfall ist die Oberfläche $M = K(r, h)$ eines Kegels der Höhe h über einer Kreisfläche vom Radius r . Dieser Fall entspricht $z_0 = 0$, $z_1 = h > 0$ und $f(z) = r(1 - z/h)$. Die Gleichung (6.3.17) liefert dann die Formel

$$\begin{aligned} \text{Vol}_2(K(r, h)) &= 2\pi \int_0^h r \left(1 - \frac{z}{h}\right) \sqrt{1 + \frac{r^2}{h^2}} dz \\ &= \pi r \sqrt{h^2 + r^2} \int_0^h 2 \left(1 - \frac{z}{h}\right) \frac{dz}{h} \\ &= \pi r \sqrt{h^2 + r^2} \end{aligned} \quad (6.3.19)$$

für die Kegeloberfläche.

6.4 Der Satz von Gauß

Der Divergenzsatz von Gauß ist der wichtigste Satz dieses Kapitels. Er ist eine Verallgemeinerung des Hauptsatzes der Infinitesimalrechnung auf Funktionen von mehreren Variablen. Er stellt einen Bezug her zwischen dem Integral der Divergenz eines Vektorfeldes über ein Gebiet im \mathbb{R}^n und einem geeigneten Integral über den Rand dieses Gebietes. Wir beweisen den Satz für Gebiete mit *glattem Rand*, beziehungsweise mit *C^ℓ -Rand* für ein $\ell \in \mathbb{N}$.

Definition 6.4.1 (Glatte Rand). Sei $\Omega \subset \mathbb{R}^n$ eine beschränkte offene Teilmenge, sei $G \subset \mathbb{R}^n$ eine kompakte Teilmenge, und sei $\ell \in \mathbb{N} \cup \{\infty\}$.

(i) Ω heißt **beschränkte offene Menge mit C^ℓ -Rand** (beziehungsweise mit **glattem Rand** im Fall $\ell = \infty$), wenn $\partial\Omega$ eine $(n-1)$ -dimensionale C^ℓ -Untermannigfaltigkeit von \mathbb{R}^n ist und mit $\partial\bar{\Omega}$ übereinstimmt (dass heißt, wenn jede offene Teilmenge $U \subset \mathbb{R}^n$ mit $U \cap \partial\Omega \neq \emptyset$ sowohl Elemente aus Ω als auch Elemente aus $\mathbb{R}^n \setminus \bar{\Omega}$ enthält).

(ii) G heißt **kompakte Menge mit C^ℓ -Rand** (beziehungsweise mit **glattem Rand** im Fall $\ell = \infty$), wenn ∂G eine $(n-1)$ -dimensionale C^ℓ -Untermannigfaltigkeit von \mathbb{R}^n ist und mit $\partial\overset{\circ}{G}$ übereinstimmt (dass heißt, wenn jede offene Teilmenge $U \subset \mathbb{R}^n$ mit $U \cap \partial G \neq \emptyset$ sowohl Elemente aus $\overset{\circ}{G}$ als auch Elemente aus $\mathbb{R}^n \setminus G$ enthält).

Bemerkung 6.4.2. Für das Verständnis von Definition 6.4.1 könnten die folgenden Anmerkungen hilfreich sein.

(i) Sei Ω eine offene Teilmenge eines metrischen Raumes M und sei $G := \bar{\Omega}$. Dann gilt $\Omega \subset \overset{\circ}{G}$ und $\partial G \subset \partial\Omega$, sowie $\Omega = \overset{\circ}{G} \iff \partial G = \partial\Omega$.

(ii) Sei G eine abgeschlossene Teilmenge eines metrischen Raumes M und sei $\Omega := \overset{\circ}{G}$. Dann gilt $\bar{\Omega} \subset G$ und $\partial\Omega \subset \partial G$, sowie $\bar{\Omega} = G \iff \partial\Omega = \partial G$.

Beispiel 6.4.3. (i) Sei $\Omega := \{x \in \mathbb{R}^n \mid 0 < \|x\| < 2, \|x\| \neq 1\}$. Dann ist Ω offen und ihr Rand ist eine glatte Untermannigfaltigkeit. Jedoch ist $\partial\bar{\Omega} \subsetneq \partial\Omega$ und daher ist Ω keine "offene Menge mit glattem Rand". Dennoch ist ihr Abschluss $G := \bar{\Omega}$ eine kompakte Menge mit glattem Rand.

(ii) Sei $B \subset \mathbb{C}$ die abgeschlossene Einheitskreisscheibe und $G := B \cup [1, 2]$. Dann ist $\partial G = S^1 \cup [1, 2] \subset \mathbb{C}$ keine Untermannigfaltigkeit. Dennoch ist ihr Inneres $\Omega := \overset{\circ}{G} = \overset{\circ}{B}$ eine offene Menge mit glattem Rand.

(iii) Sei $G \subset \mathbb{R}^n$ eine nichtleere kompakte glatte Untermannigfaltigkeit der Dimension $d < n$. Dann ist $\partial G = G$ eine glatte Untermannigfaltigkeit, jedoch ist $\partial\overset{\circ}{G} = \emptyset \subsetneq \partial G$ und daher ist G keine "kompakte Menge mit glattem Rand".

Wir werden in Lemma 6.4.5 offene Mengen mit C^ℓ -Rand charakterisieren. Zu dessen Beweis ist das folgende Lemma hilfreich.

Lemma 6.4.4. *Sei $U \subset M$ eine Teilmenge eines metrischen Raumes (M, d) , sei $V \subset \mathbb{R}^n$ eine konvexe offene Teilmenge von \mathbb{R}^n , und sei*

$$\phi = (\phi_1, \dots, \phi_n) : U \rightarrow V$$

ein Homöomorphismus. Wir definieren

$$\begin{aligned} U^+ &:= \{p \in U \mid \phi_n(p) > 0\}, \\ U^0 &:= \{p \in U \mid \phi_n(p) = 0\}, \\ U^- &:= \{p \in U \mid \phi_n(p) < 0\}. \end{aligned} \tag{6.4.1}$$

Sei $\Omega \subset M$ eine weitere offene Menge mit

$$U \cap \partial\Omega = U^0. \tag{6.4.2}$$

Dann gelten die Aussagen

$$U^+ \cap \Omega \neq \emptyset \quad \implies \quad U^+ \subset \Omega, \tag{6.4.3}$$

$$U^- \cap \Omega \neq \emptyset \quad \implies \quad U^- \subset \Omega. \tag{6.4.4}$$

Proof. Sei $I := [0, 1]$ und seien

$$p \in U^+ \cap \Omega, \quad x := \phi(p), \quad q \in U^+, \quad y := \phi(q).$$

Dann ist $(1-t)x + ty \in V$ für alle $t \in I$, da V eine konvexe Menge ist, und es gilt $x_n > 0$ und $y_n > 0$, da $p, q \in U^+$ sind. Sei $\gamma : I \rightarrow U$ die Kurve

$$\gamma(t) := \phi^{-1}((1-t)x + ty) \quad \text{für } t \in I.$$

Diese Kurve ist stetig, da ϕ ein Homöomorphismus ist. Außerdem ist

$$\phi_n(\gamma(t)) = (1-t)x_n + ty_n > 0 \quad \text{für alle } t \in I,$$

und daraus folgt $\gamma(t) \notin \partial\Omega$ für alle $t \in I$ nach (6.4.1) und (6.4.2). Daraus folgt $\gamma(t) \in \Omega \iff \gamma(t) \in \overline{\Omega}$ für alle $t \in I$, und daher ist die Menge

$$A := \{t \in I \mid \gamma(t) \in \Omega\} = \{t \in I \mid \gamma(t) \in \overline{\Omega}\}$$

sowohl I -offen als auch I -abgeschlossen. Außerdem ist diese Menge nicht-leer, da $\gamma(0) = p \in \Omega$ und daher $0 \in A$ ist. Daraus folgt $A = I$, da I zusammenhängend ist (Satz B.2.2). Daher ist $1 \in A$, und das heißt $q = \gamma(1) \in \Omega$. Damit ist (6.4.3) bewiesen. Die Aussage (6.4.4) folgt nun durch Vorzeichenumkehr von ϕ_n . Damit ist Lemma 6.4.4 bewiesen. \square

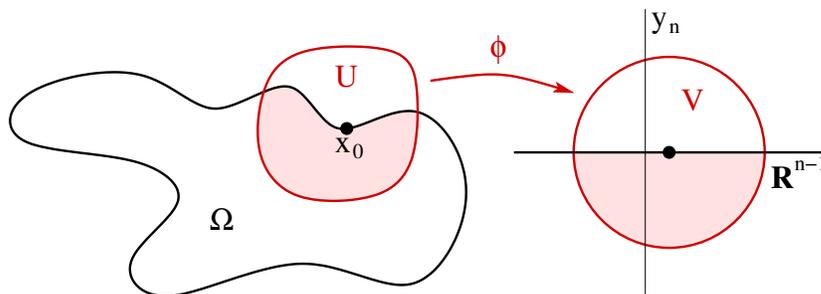


Abbildung 6.6: Eine offene Menge mit glattem Rand.

Lemma 6.4.5 (Glatte Rand). Sei $\Omega \subset \mathbb{R}^n$ eine beschränkte offene Menge und sei $\ell \in \mathbb{N} \cup \{\infty\}$. Dann sind folgende Aussagen äquivalent.

- (i) Ω ist eine beschränkte offene Menge mit C^ℓ -Rand.
- (ii) $\partial\Omega$ ist eine $(n - 1)$ -dimensionale C^ℓ -Untermannigfaltigkeit von \mathbb{R}^n und für jedes $x_0 \in \partial\Omega$ existiert ein $\xi_0 \in \mathbb{R}^n$ und ein $\delta > 0$ mit

$$\xi_0 \notin T_{x_0}\partial\Omega, \quad x_0 + t\xi_0 \notin \Omega \text{ für alle } t \in [0, \delta]. \quad (6.4.5)$$

- (iii) Für jedes $x_0 \in \partial\Omega$ existieren offene Mengen $U, V \subset \mathbb{R}^n$ mit $x_0 \in U$ und ein C^ℓ -Diffeomorphismus $\phi : U \rightarrow V$ (Abbildung 6.6) mit

$$\phi(U \cap \partial\Omega) = \{y \in V \mid y_n = 0\}, \quad \phi(U \cap \Omega) = \{y \in V \mid y_n < 0\}. \quad (6.4.6)$$

- (iv) Für jedes $x_0 \in \partial\Omega$ existiert eine offene Menge $U \subset \mathbb{R}^n$ mit $x_0 \in U$ und eine C^ℓ -Funktion $g : U \rightarrow \mathbb{R}$ mit $\nabla g(x) \neq 0$ für alle $x \in U \cap \partial\Omega$ und

$$U \cap \partial\Omega = \{x \in U \mid g(x) = 0\}, \quad U \cap \Omega = \{x \in U \mid g(x) < 0\}. \quad (6.4.7)$$

Beweis. Wir zeigen (i) \implies (ii). Sei also an $\partial\Omega = \partial\bar{\Omega}$ eine $(n-1)$ -dimensionale C^ℓ -Untermannigfaltigkeit von \mathbb{R}^n und sei $x_0 \in \partial\Omega$. Dann existiert ein C^ℓ -Diffeomorphismus $\phi : U \rightarrow V$ zwischen offenen Teilmengen des $U, V \subset \mathbb{R}^n$ mit $x_0 \in U$, so dass V convex ist und $U \cap \partial\Omega = U^0$ ist (in der Notation von Lemma 6.4.4). Daraus folgt, dass eine der Mengen $U^+ \cap \Omega$ und $U^- \cap \Omega$ leer sein muß, denn sonst wäre $U \subset \bar{\Omega}$ nach Lemma 6.4.4, und somit wäre x_0 kein Randpunkt von $\bar{\Omega}$. Ist $U^+ \cap \Omega = \emptyset$, so erfüllt jeder Vektor $\xi_0 \in \mathbb{R}^n$ mit $d\phi_n(x_0)\xi_0 > 0$ die Bedingung (6.4.5). Ist $U^- \cap \Omega = \emptyset$, so erfüllt jeder Vektor $\xi_0 \in \mathbb{R}^n$ mit $d\phi_n(x_0)\xi_0 < 0$ die Bedingung (6.4.5). Also erfüllt Ω die Bedingung (ii).

Wir zeigen (ii) \implies (iii). Sei $x_0 \in \partial\Omega$. Da $\partial\Omega$ eine $(n-1)$ -dimensionale C^ℓ -Untermannigfaltigkeit ist existieren offene Mengen $U, V \subset \mathbb{R}^n$ mit $x_0 \in U$ und ein C^1 -Diffeomorphismus $\phi : U \rightarrow V$ mit

$$\phi(U \cap \partial\Omega) = \{y \in V \mid y_n = 0\},$$

das heißt $U \cap \partial\Omega = U^0$ in der Notation von Lemma 6.4.4. Außerdem gilt

$$T_{x_0}\partial\Omega = d\phi(x_0)^{-1}(\mathbb{R}^{n-1} \times \{0\}).$$

Durch Verkleinern von U, V falls nötig, können wir annehmen, dass V konvex ist. Nun seien $\xi_0 \in \mathbb{R}^n$ und $\delta > 0$ so gewählt, dass (6.4.5) gilt. Da $\xi_0 \notin T_{x_0}\partial\Omega$ ist, gilt $d\phi(x_0)\xi_0 \notin \mathbb{R}^{n-1} \times \{0\}$. Daher ist $d\phi_n(x_0)\xi_0 \neq 0$ und, indem wir ϕ_n durch $-\phi_n$ ersetzen falls nötig, können wir annehmen, dass

$$d\phi_n(x_0)\xi_0 > 0$$

ist. Daraus folgt $\phi_n(x_0 + t\xi_0) > 0$ für kleine $t > 0$. Ist $0 < t_0 \leq \delta$ so klein, dass $x_0 + t_0\xi_0 \in U$ und $\phi_n(x_0 + t_0\xi_0) > 0$ ist, dann ist $x_0 + t_0\xi_0 \in U^+ \setminus \Omega$ in der Notation von Lemma 6.4.4. Also ist $U^+ \not\subset \Omega$ und daher $U^+ \cap \Omega = \emptyset$ nach Lemma 6.4.4. Da $(U^+ \cup U^0) \cap \Omega = \emptyset$ ist, muß $U^- \cap \Omega \neq \emptyset$ sein, und daraus folgt $U^- \subset \Omega$ nach Lemma 6.4.4. Also ist

$$U \cap \Omega = U^-,$$

und damit gilt (ii).

Die Implikation (iii) \implies (iv) folgt sofort mit $g := \phi_n$.

Wir zeigen (iv) \implies (i). Sei $x_0 \in \partial\Omega$ und sei $g : U \rightarrow \mathbb{R}$ wie in (iv). Dann ist der Durchschnitt von $\partial\Omega$ mit einer hinreichend kleinen Umgebung von x_0 das Urbild eines regulären Wertes unter einer reellwertigen C^ℓ -Funktion, und daher eine $(n-1)$ -dimensionale C^ℓ -Untermannigfaltigkeit. Außerdem gilt

$$x_0 + t\nabla g(x_0) \in U, \quad g(x_0 + t\nabla g(x_0)) > 0$$

für jedes hinreichen kleine $t > 0$. Also ist $x_0 + t\nabla g(x_0) \notin \overline{\Omega}$ für jedes solche t nach (6.4.7), und daraus folgt $x_0 = \lim_{t \searrow 0} (x_0 + t\nabla g(x_0)) \in \partial\overline{\Omega}$. Damit haben wir gezeigt, dass $\partial\Omega \subset \partial\overline{\Omega}$ ist. Die umgekehrte Inklusion gilt für jede offene Menge Ω nach Teil (i) von Bemerkung 6.4.2. Also ist $\partial\Omega = \partial\overline{\Omega}$ und damit erfüllt Ω die Bedingung (i). Damit ist Lemma 6.4.5 bewiesen. \square

Lemma 6.4.6. *Sei $\Omega \subset \mathbb{R}^n$ eine beschränkte offene Menge mit C^ℓ -Rand. Dann existiert eine offene Menge $U \subset \mathbb{R}^n$ mit $\partial\Omega \subset U$ und eine C^ℓ -Funktion $g : U \rightarrow \mathbb{R}$ mit den Eigenschaften*

$$\nabla g(x) \neq 0 \quad \text{für alle } x \in \partial\Omega, \quad (6.4.8)$$

$$U \cap \partial\Omega = \{x \in U \mid g(x) = 0\}, \quad (6.4.9)$$

$$U \cap \Omega = \{x \in U \mid g(x) < 0\}. \quad (6.4.10)$$

Beweis. Da Ω beschränkt, und $\partial\Omega$ daher kompakt ist, existieren nach Definition (Teil (iii) von Lemma 6.4.5) und Satz C.1.2, offene Mengen U_1, \dots, U_N mit $\partial\Omega \subset \bigcup_{i=1}^N U_i$ und C^ℓ -Funktionen $g_i : U_i \rightarrow \mathbb{R}$ für $i = 1, \dots, N$, die 0 als regulären Wert haben und (6.4.7) erfüllen. Nach Satz 6.2.4 existieren glatte Funktionen $\rho_i : \mathbb{R}^n \rightarrow [0, 1]$ mit kompaktem Träger, die die Bedingungen $\text{supp}(\rho_i) \subset U_i$ für $i = 1, \dots, N$ und

$$\sum_{i=1}^N \rho_i(x) = 1 \quad \text{für alle } x \in \partial\Omega$$

erfüllen. Dann ist

$$U := \left\{ x \in \mathbb{R}^n \mid \sum_{i=1}^N \rho_i(x) > 0 \right\} \subset \bigcup_{i=1}^N U_i$$

eine offene Menge, die $\partial\Omega$ enthält, und die Funktion

$$g := \sum_{i=1}^N \rho_i g_i : U \rightarrow \mathbb{R}$$

erfüllt die Behauptungen des Lemmas. Ist $x \in U \cap \partial\Omega$ so ist jeder Summand $\rho_i(x)g_i(x)$ mit $x \in U_i$ gleich Null. Ist $x \in U \cap \Omega$, so ist jeder Summand $\rho_i(x)g_i(x)$ mit $x \in U_i$ kleiner oder gleich Null und mindestens einer ist negativ. Ist $x \in U \setminus \bar{\Omega}$, so ist jeder Summand $\rho_i(x)g_i(x)$ mit $x \in U_i$ grösser oder gleich Null und mindestens einer ist positiv. Daraus folgen die Gleichungen (6.4.9) und (6.4.10). Die Aussage (6.4.8) gilt, weil $g_i(x) = 0$ ist für alle $x \in U_i \cap \partial\Omega$ und die Vektoren $\nabla g_i(x)$ und $\nabla g_j(x)$ für $x \in U_i \cap U_j \cap \partial\Omega$ positive Vielfache voneinander sind. Damit ist Lemma 6.4.6 bewiesen. \square

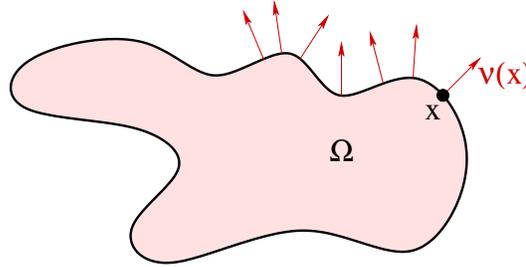


Abbildung 6.7: Das äußere Einheitsnormalenfeld.

Lemma 6.4.7 (Das äußere Einheitsnormalenfeld). Sei $\Omega \subset \mathbb{R}^n$ eine beschränkte offene Menge mit C^1 -Rand. Dann gelten folgende Aussagen.

(i) Für jedes $x \in \partial\Omega$ existiert genau ein $\nu(x) \in \mathbb{R}^n$ (Abbildung 6.7) mit

$$\begin{aligned} \|\nu(x)\| &= 1, & \nu(x) &\perp T_x\partial\Omega, \\ x + t\nu(x) &\notin \Omega \text{ für kleine } t > 0. \end{aligned} \quad (6.4.11)$$

(ii) Die Abbildung $\nu : \partial\Omega \rightarrow S^{n-1}$ in (i) ist stetig.

(iii) Wenn Ω einen C^ℓ -Rand hat, so existiert eine offene Menge $U \subset \mathbb{R}^n$ mit $\partial\Omega \subset U$, so dass sich die Abbildung $\nu : \partial\Omega \rightarrow S^{n-1}$ zu einer $C^{\ell-1}$ -Abbildung von U nach S^{n-1} fortsetzen läßt.

Beweis. Sei $g : U \rightarrow \mathbb{R}$ eine Funktion, welche die Bedingungen von Lemma 6.4.6 erfüllt. Dann ist 0 ein regulärer Wert von h mit $\partial\Omega = h^{-1}(0)$ nach (6.4.8) und (6.4.9). Daher gilt $T_x\partial\Omega = \nabla g(x)^\perp$ für alle $x \in \partial\Omega$. Daraus folgt die Existenz in Teil (i) folgt aus Lemma 6.4.6 mit

$$\nu(x) := \frac{\nabla g(x)}{\|\nabla g(x)\|}. \quad (6.4.12)$$

Dieser Vektor hat die Norm eins, und ist orthogonal zu $T_x\partial\Omega$. Außerdem ist $\langle \nabla g(x), \nu(x) \rangle = \|\nabla g(x)\| > 0$ und daraus folgt $g(x + t\nu(x)) > 0$, und daher $x + t\nu(x) \notin \Omega$ nach (6.4.10), für jedes hinreichend kleine $t > 0$. Es folgt ebenfalls, dass $g(x_0 - t\nu(x_0)) < 0$ und daher $x_0 - t\nu(x_0) \in \Omega$ ist für jedes hinreichend kleine $t > 0$. Da es nur zwei Einheitsvektoren in $(T_{x_0}\partial\Omega)^\perp$ gibt, ist damit Teil (i) bewiesen. Die Stetigkeit in Teil (ii) folgt ebenfalls direkt aus der Formel (6.4.12) und der Stetigkeit der partiellen Ableitungen von g . In Teil (iii) folgt die Existenz einer $C^{\ell-1}$ -Fortsetzung auch direkt aus der Formel (6.4.12), da die Funktion g nach Lemma 6.4.6 so gewählt werden kann, dass sie ℓ mal stetig differenzierbar ist. Damit ist Lemma 6.4.7 bewiesen. \square

Definition 6.4.8. Sei $\Omega \subset \mathbb{R}^n$ eine beschränkte offene Menge mit C^1 -Rand. Die Funktion $\nu : \partial\Omega \rightarrow S^{n-1}$ in Lemma 6.4.7 heißt **äußeres Einheitsnormalenfeld** von Ω . Sie wird auch die **Gauß-Abbildung von Ω** genannt.

Satz 6.4.9 (Divergenzsatz von Gauß). Sei $\Omega \subset \mathbb{R}^n$ eine beschränkte offene Menge mit C^1 -Rand, sei $\nu : \partial\Omega \rightarrow S^{n-1}$ das äußere Einheitsnormalenfeld, sei $W \subset \mathbb{R}^n$ eine offene Menge mit $\bar{\Omega} \subset W$, sei $f : W \rightarrow \mathbb{R}^n$ eine C^1 -Abbildung, und sei

$$\operatorname{div}(f) := \sum_{i=1}^n \frac{\partial f_i}{\partial x_i} : W \rightarrow \mathbb{R}, \quad \langle f, \nu \rangle := \sum_{i=1}^n f_i \nu_i : \partial\Omega \rightarrow \mathbb{R}.$$

Dann gilt

$$\int_{\Omega} \operatorname{div}(f) = \int_{\partial\Omega} \langle f, \nu \rangle dS, \quad (6.4.13)$$

Beweis. Zunächst sei erwähnt, dass es sich bei Ω um eine Jordan-messbare Teilmenge des \mathbb{R}^n handelt, da ihr Rand eine kompakte $(n-1)$ -dimensionale C^1 -Untermannigfaltigkeit von \mathbb{R}^n und daher nach Beispiel 5.6.13 eine Jordansche Nullmenge ist. Da f auf einer Umgebung der Menge $\bar{\Omega}$ stetig ist, ist also das Integral auf der linken Seite der Gleichung (6.4.13) wohldefiniert (Definition 5.5.2).

Wir beweisen den Divergenzsatz zunächst für Funktionen mit Träger in einer hinreichen kleinen Umgebung eines Randpunktes. Sei also $x_0 \in \partial\Omega$ gegeben. Dann existiert nach Voraussetzung (Teil (iii) von Lemma 6.4.5) eine offene Mengen $U \subset W$ mit $x_0 \in U$ und eine C^1 -Abbildung $g : U \rightarrow \mathbb{R}$ mit $\nabla g(x_0) \neq 0$ und

$$U \cap \partial\Omega = \{x \in U \mid g(x) = 0\}, \quad U \cap \Omega = \{x \in U \mid g(x) < 0\}.$$

Da $\nabla g(x_0)$ von Null ist, dürfen wir ohne Einschränkung der Allgemeinheit annehmen, dass $\partial g / \partial x_n(x_0) > 0$ ist. Genauer gesagt, es kann durchaus sein, dass $\partial g / \partial x_n(x_0) = 0$ ist; in diesem Fall existiert ein $i \in \{1, \dots, n-1\}$ mit $\partial g / \partial x_i(x_0) \neq 0$; durch Vertauschen der Koordinaten x_i und x_n erhalten wir dann $\partial g / \partial x_n(x_0) \neq 0$; anschließend können wir noch x_n durch $-x_n$ ersetzen, falls nötig, um die Ungleichung $\partial g / \partial x_n(x_0) > 0$ zu erhalten. Nach diesen Vorbereitungen erfüllt g die Bedingungen

$$g(x_0) = 0, \quad \frac{\partial g}{\partial x_n}(x_0) > 0. \quad (6.4.14)$$

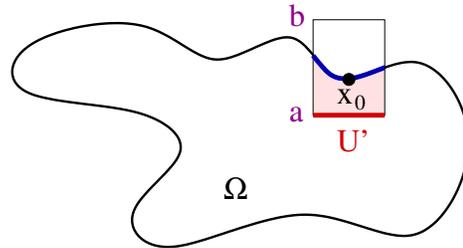


Abbildung 6.8: Eine Umgebung eines Randpunktes.

Es ist an dieser Stelle nützlich, die Bezeichnung

$$x' = (x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1}, \quad x = (x_1, \dots, x_n) = (x', x_n) \in \mathbb{R}^n$$

einzuführen. Insbesondere ist $x'_0 \in \mathbb{R}^{n-1}$ der Vektor, dessen Koordinaten mit den ersten $n - 1$ Koordinaten von x_0 übereinstimmen, $x_{0n} \in \mathbb{R}$ ist die letzte Koordinate von x_0 , und $x_0 = (x'_0, x_{0n})$. Dann folgt aus den Bedingungen in (6.4.14), dass reelle Zahlen $a < b$ existieren mit

$$a < x_{0n} < b, \quad \{x'_0\} \times [a, b] \subset U, \quad \frac{\partial g}{\partial x_n}(x'_0, x_n) > 0 \text{ für alle } x_n \in [a, b].$$

Da $g(x'_0, x_{0n}) = g(x_0) = 0$ ist, gilt dann nach dem Mittelwertsatz

$$g(x'_0, a) < 0 < g(x'_0, b).$$

Daher existiert eine offene Menge $U' \subset \mathbb{R}^{n-1}$ (Abbildung 6.8) mit

$$\begin{aligned} x'_0 \in U', \quad U' \times [a, b] \subset U, \\ \frac{\partial g}{\partial x_n}(x) > 0 \quad \text{für alle } x \in U' \times [a, b], \\ g(x', a) < 0 < g(x', b) \quad \text{für alle } x' \in U'. \end{aligned} \tag{6.4.15}$$

Aus (6.4.15) folgt nach dem Zwischenwertsatz und dem Mittelwertsatz, dass für jedes $x' \in U'$ genau eine reelle Zahl $h(x')$ existiert mit

$$a < h(x') < b, \quad g(x', h(x')) = 0. \tag{6.4.16}$$

Nach dem Satz über implizite Funktionen in Kapitel 3 ist die so definierte Funktion $h : U' \rightarrow (a, b)$ stetig differenzierbar (siehe Abbildung 6.9).

Der äußere Einheitsnormalenvektor an der Stelle $(x', h(x')) \in \partial\Omega$ ist in dieser Situation der eindeutige Einheitsvektor, der senkrecht auf dem Graphen der Ableitung von h an der Stelle $x' \in U'$ steht, und dessen n te Koordinate positiv ist. Dies ist der Vektor $\nu(x', h(x')) \in \mathbb{R}^n$ mit den Koordinaten

$$\begin{aligned} \nu_i(x', h(x')) &= -\frac{\partial_i h(x')}{\sqrt{1 + \|\nabla h(x')\|^2}} \quad \text{für } i = 1, \dots, n-1, \\ \nu_n(x', h(x')) &= \frac{1}{\sqrt{1 + \|\nabla h(x')\|^2}}. \end{aligned} \tag{6.4.17}$$

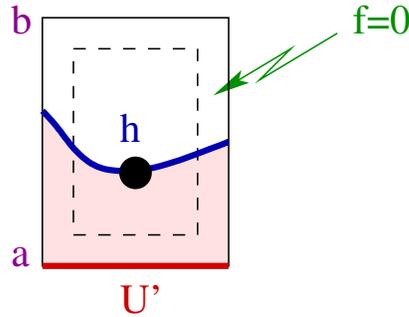


Abbildung 6.9: Der Rand als Graph einer Funktion $h : U' \rightarrow (a, b)$.

Sei nun $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld, dessen Träger in der offenen Menge $U' \times (a, b)$ enthalten ist. Dann ist das Integral der Funktion $\langle f, \nu \rangle : \partial\Omega \rightarrow \mathbb{R}$ nach (6.4.17) und Beispiel 6.3.5 durch die Formel

$$\begin{aligned} &\int_{\partial\Omega} \langle f, \nu \rangle dS \\ &= \int_{U'} \sum_{i=1}^n f_i(x', h(x')) \nu_i(x', h(x')) \sqrt{1 + \|\nabla h(x')\|^2} dx' \\ &= \int_{U'} \left(f_n(x', h(x')) - \sum_{i=1}^{n-1} f_i(x', h(x')) \frac{\partial h}{\partial x_i}(x') \right) dx' \end{aligned} \tag{6.4.18}$$

gegeben, und das Integral von $\text{div}(f)$ über Ω ist nach dem Satz von Fubini (Satz 5.4.1) das iterierte Integral

$$\int_{\Omega} \text{div}(f) = \sum_{i=1}^n \int_{U'} \int_a^{h(x')} \frac{\partial f_i}{\partial x_i}(x', x_n) dx_n dx'. \tag{6.4.19}$$

Zu zeigen ist nun, dass die beiden Integrale in (6.4.18) und (6.4.19) übereinstimmen. Für $i = n$ erhalten wir nach dem Hauptsatz der Infinitesimalrechnung die Formel

$$\begin{aligned} \int_a^{h(x')} \frac{\partial f_n}{\partial x_n}(x', x_n) dx_n &= f_n(x', h(x')) - f_n(x', a) \\ &= f_n(x', h(x')). \end{aligned}$$

Durch Integration über $x' \in U'$ ergibt sich daraus

$$\int_{\Omega} \frac{\partial f_n}{\partial x_n} = \int_{\partial\Omega} f_n \nu_n dS.$$

Für $i = 1, \dots, n-1$ und jede reelle Zahl t liefert der Hauptsatz der Infinitesimalrechnung die Formel

$$\int_{U'} \frac{\partial}{\partial x_i} f_i(x', t + h(x')) dx' = 0.$$

Durch Integration über $t \in (-\infty, 0]$ erhalten wir dann

$$\begin{aligned} 0 &= \int_{-\infty}^0 \int_{U'} \frac{\partial}{\partial x_i} f_i(x', t + h(x')) dx' dt \\ &= \int_{-\infty}^0 \int_{U'} \left(\frac{\partial f_i}{\partial x_i}(x', t + h(x')) + \frac{\partial f_i}{\partial x_n}(x', t + h(x')) \frac{\partial h}{\partial x_i}(x') \right) dx' dt \\ &= \int_{U'} \int_{-\infty}^0 \frac{\partial f_i}{\partial x_i}(x', t + h(x')) dt dx' \\ &\quad + \int_{U'} \left(\int_{-\infty}^0 \frac{\partial}{\partial t} f_i(x', t + h(x')) dt \right) \frac{\partial h}{\partial x_i}(x') dx' \\ &= \int_{U'} \int_a^{h(x')} \frac{\partial f_i}{\partial x_i}(x', x_n) dx_n dx' \\ &\quad + \int_{U'} f_i(x', h(x')) \frac{\partial h}{\partial x_i}(x') dx' \\ &= \int_{\Omega} \frac{\partial f_i}{\partial x_i} - \int_{\partial\Omega} f_i \nu_i dS. \end{aligned}$$

Damit ist der Divergenzsatz von Gauß für jedes stetig differenzierbaren Vektorfeld $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit $\text{supp}(f) \subset U' \times (a, b)$ bewiesen.

Zusammenfassend haben wir gezeigt, dass jeder Randpunkt von Ω eine offene Umgebung $U \subset W$ besitzt, so dass der Divergenzsatz von Gauß für alle Vektorfelder mit Träger in U gültig ist. Das gleiche gilt für Punkte in Ω , da beide Integrale in (6.4.13) für C^1 -Funktionen, deren Träger in Ω enthalten ist, verschwinden. Da $\bar{\Omega}$ eine kompakte Teilmenge des \mathbb{R}^n ist, folgt daraus nach Satz C.1.2, dass es endlich viele offene Teilmengen $U_1, \dots, U_N \subset W$ gibt, so dass

$$\bar{\Omega} \subset U_1 \cup U_2 \cup \dots \cup U_N$$

ist und die Gleichung (6.4.13) für jedes $i \in \{1, \dots, N\}$ und jedes stetig differenzierbare Vektorfeld $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit $\text{supp}(f) \subset U_i$ gilt. Nach Satz 6.2.4 existieren glatte Funktionen

$$\rho_1, \dots, \rho_N : \mathbb{R}^n \rightarrow [0, 1]$$

mit kompaktem Träger, welche die Bedingungen

$$\text{supp}(\rho_i) \subset U_i \quad \text{für } i = 1, \dots, N$$

und

$$\sum_{i=1}^N \rho_i(x) = 1 \quad \text{für alle } x \in \bar{\Omega}.$$

erfüllen. Ist nun $f : W \rightarrow \mathbb{R}^n$ ein beliebiges stetig differenzierbares Vektorfeld, so gilt (6.4.13) für jedes der Vektorfelder $\rho_i f$ und daraus folgt

$$\begin{aligned} \int_{\Omega} \text{div}(f) &= \int_{\Omega} \text{div} \left(\sum_{i=1}^N \rho_i f \right) \\ &= \sum_{i=1}^N \int_{\Omega} \text{div}(\rho_i f) \\ &= \sum_{i=1}^N \int_{\partial\Omega} \langle \rho_i f, \nu \rangle dS \\ &= \int_{\partial\Omega} \sum_{i=1}^N \rho_i \langle f, \nu \rangle dS \\ &= \int_{\partial\Omega} \langle f, \nu \rangle dS. \end{aligned}$$

Damit ist Satz 6.4.9 bewiesen. □

6.5 Weitere Beispiele

Beispiel 6.5.1 (Laplace-Operator). Der Laplace-Operator auf dem \mathbb{R}^n ist der Differentialoperator zweiter Ordnung

$$\Delta := \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}. \quad (6.5.1)$$

Sei nun $\Omega \subset \mathbb{R}^n$ eine beschränkte offene Menge mit glattem Rand. Eine reellwertige C^2 Funktion $u : \Omega \rightarrow \mathbb{R}$ auf einer offenen Teilmenge $\Omega \subset \mathbb{R}^n$ heißt **harmonisch** wenn $\Delta u = 0$ ist. Für jede C^2 -Funktionen $u : \Omega \rightarrow \mathbb{R}$ und jede C^1 -Funktion $v : \Omega \rightarrow \mathbb{R}$ gilt

$$\operatorname{div}(v\nabla u) = v\Delta u + \langle \nabla u, \nabla v \rangle. \quad (6.5.2)$$

Für $\ell \in \mathbb{N}$ nennen wir $u : \bar{\Omega} \rightarrow \mathbb{R}$ eine C^ℓ -Funktion, wenn sie sich zu einer C^ℓ -Funktion auf einer offenen Teilmenge des \mathbb{R}^n , die $\bar{\Omega}$ enthält, fortsetzen läßt. Ist $u : \bar{\Omega} \rightarrow \mathbb{R}$ eine C^1 -Funktion, so wird die Ableitung von u in Richtung des äußeren Normalenfeldes $\nu : \partial\Omega \rightarrow \mathbb{R}$ die **normale Ableitung** von u genannt. Wir bezeichnen sie mit

$$\frac{\partial u}{\partial \nu} := \langle \nabla u, \nu \rangle : \partial\Omega \rightarrow \mathbb{R}.$$

Damit erhalten wir aus Satz 6.4.9 und Gleichung (6.5.2) mit $v = 1$ und $v = u$ für jede C^2 -Funktion $u : \bar{\Omega} \rightarrow \mathbb{R}$ die Formeln

$$\int_{\Omega} \Delta u = \int_{\partial\Omega} \frac{\partial u}{\partial \nu} dS \quad (6.5.3)$$

und

$$\int_{\Omega} \|\nabla u\|^2 = \int_{\partial\Omega} u \frac{\partial u}{\partial \nu} dS - \int_{\Omega} u \Delta u. \quad (6.5.4)$$

Die Gleichung (6.5.4) zeigt, dass jede harmonische C^2 -Funktion $u : \bar{\Omega} \rightarrow \mathbb{R}$, die auf dem Rand verschwindet, auf ganz Ω verschwindet, denn sie ist auf jeder Zusammenhangskomponente von $\bar{\Omega}$ konstant und jede solche Zusammenhangskomponente enthält einen Randpunkt. Die Gleichung (6.5.3) zeigt, dass die normale Ableitung einer harmonischen C^2 -Funktion $u : \bar{\Omega} \rightarrow \mathbb{R}$ auf dem Rand den Mittelwert Null hat. Eine weitere wichtige Gleichung, die aus (6.5.2) und Satz 6.4.9 für zwei C^2 -Funktionen $u, v : \bar{\Omega} \rightarrow \mathbb{R}$ folgt, ist

$$\int_{\Omega} (u\Delta v - v\Delta u) = \int_{\partial\Omega} \left(u \frac{\partial v}{\partial \nu} - v \frac{\partial u}{\partial \nu} \right) dS. \quad (6.5.5)$$

Beispiel 6.5.2 (Volumen). Die Divergenz des Vektorfeld $f(x) = x$ ist die Konstante Funktion $\operatorname{div}(f) \equiv n$ und damit gilt $\int_{\Omega} \operatorname{div}(f) = n \operatorname{Vol}_n(\Omega)$. Nach Satz 6.4.9 ergibt sich daher für jede beschränkte offene Menge $\Omega \subset \mathbb{R}^n$ mit C^1 -Rand die Formel

$$\operatorname{Vol}_n(\Omega) = \frac{1}{n} \int_{\partial\Omega} \langle x, \nu(x) \rangle dS(x). \quad (6.5.6)$$

Ist $\Omega = B^n \subset \mathbb{R}^n$ der offene Einheitsball, so ist $\partial\Omega = S^{n-1}$ die Einheitssphäre und das äußere Einheitsnormalenfeld $\nu : \partial\Omega \rightarrow S^{n-1}$ ist in diesem Fall die Identität $\nu(x) = x$. Damit ergibt sich aus (6.5.6) die Gleichung

$$\operatorname{Vol}_n(B^n) = \frac{1}{n} \operatorname{Vol}_{n-1}(S^{n-1})$$

und wir erhalten damit einen weiteren Beweis der Gleichung (6.3.12).

Beispiel 6.5.3. Sei $\Omega \subset \mathbb{R}^n$ eine beschränkte offene Menge mit C^1 -Rand, sei $a \in \mathbb{R}^n \setminus \partial\Omega$, und sei $f : \mathbb{R}^n \setminus \{a\} \rightarrow \mathbb{R}^n$ das Vektorfeld

$$f(x) := \frac{x - a}{\|x - a\|^n}, \quad x \in \mathbb{R}^n \setminus \{a\}.$$

Dann gilt

$$\frac{\partial f_i}{\partial x_i}(x) = \frac{\partial}{\partial x_i} \frac{x_i - a_i}{\|x - a\|^n} = \frac{1}{\|x - a\|^n} - \frac{n(x_i - a_i)^2}{\|x - a\|^{n+2}}$$

für $i = 1, \dots, n$ und daher

$$(\operatorname{div}(f))(x) = \frac{n}{\|x - a\|^n} - \sum_{i=1}^n \frac{n(x_i - a_i)^2}{\|x - a\|^{n+2}} = 0.$$

Für $a \in \mathbb{R}^n \setminus \overline{\Omega}$ folgt daraus nach Satz 6.4.9 die Gleichung

$$\int_{\partial\Omega} \frac{\langle x - a, \nu(x) \rangle}{\|x - a\|^n} dS(x) = \int_{\partial\Omega} \langle f, \nu \rangle dS = \int_{\Omega} \operatorname{div}(f) = 0.$$

Ist $a \in \Omega$ und $r > 0$ mit $\overline{B}_r(a) = \{x \in \mathbb{R}^n \mid \|x - a\| \leq r\} \subset \Omega$ so ist $\Omega \setminus \overline{B}_r(a)$ eine beschränkte offene Teilmenge des \mathbb{R}^n mit C^1 -Rand

$$\partial(\Omega \setminus \overline{B}_r(a)) = \partial\Omega \cup \partial B_r(a).$$

Für $x \in \partial B_r(a)$ ist $\nu(x) = -\|x - a\|^{-1}(x - a)$ der äussere Einheitsnormalenvektor bezüglich der offenen Menge $\Omega \setminus \overline{B}_r(a)$ (siehe Abbildung 6.10).

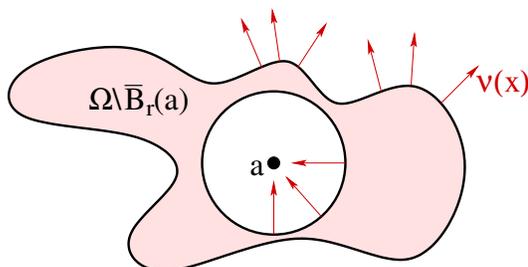


Abbildung 6.10: Der Satz von Gauß für das Komplement eines Balles.

Damit ergibt sich aus Satz 6.4.9 die Gleichung

$$\begin{aligned}
 0 &= \int_{\Omega \setminus \bar{B}_r(a)} \operatorname{div}(f) \\
 &= \int_{\partial\Omega} \langle f, \nu \rangle dS - \int_{\partial B_r(a)} \frac{\langle f(x), x - a \rangle}{\|x - a\|} dS(x) \\
 &= \int_{\partial\Omega} \langle f, \nu \rangle dS - \omega_n.
 \end{aligned}$$

Zusammenfassend erhalten wir die Formel

$$\frac{1}{\omega_n} \int_{\partial\Omega} \frac{\langle x - a, \nu(x) \rangle}{\|x - a\|^n} dS(x) = \begin{cases} 1, & \text{für } a \in \Omega, \\ 0, & \text{für } a \in \mathbb{R}^n \setminus \bar{\Omega}. \end{cases} \quad (6.5.7)$$

Im Fall $n = 2$ kann dieses Integral als *Umlaufzahl* von $\partial\Omega$ um einen Punkt a im Komplement des Randes interpretiert werden. (Vergleiche Beispiel 7.6.2.)

Beispiel 6.5.4 (Elektrodynamik). Seien m paarweise verschiedene Punkte $a_1, \dots, a_m \in \mathbb{R}^3$ gegeben. Wir nehmen an es befinde sich eine elektrische Ladung der Stärke $q_i \in \mathbb{R}$ an der Stelle a_i . Diese Ladungen erzeugen ein elektrisches Feld welches mit der Kraft

$$E(x) = \sum_{i=1}^m q_i \frac{x - a_i}{\|x - a_i\|^3} \quad (6.5.8)$$

auf eine positive Einheitsladung an der Stelle $x \in \mathbb{R}^3 \setminus \{a_1, \dots, a_m\}$ wirkt. Ist $\Omega \subset \mathbb{R}^3$ eine offene Menge mit C^1 -Rand $\partial\Omega \subset \mathbb{R}^3 \setminus \{a_1, \dots, a_m\}$, so ergibt sich aus (6.5.7) mit $\omega_3 = 4\pi$ die Formel

$$\int_{\partial\Omega} \langle E, \nu \rangle dS = \sum_{i=1}^m q_i \int_{\partial\Omega} \frac{\langle x - a_i, \nu(x) \rangle}{\|x - a_i\|^3} dS(x) = 4\pi \sum_{a_i \in \Omega} q_i. \quad (6.5.9)$$

Der Ladungsfluss über den Rand von Ω ist also das 4π -Fache der Gesamtladung im Gebiet Ω (Abbildung 6.11).

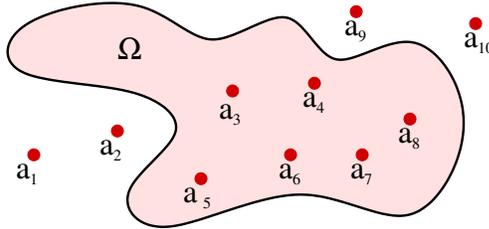


Abbildung 6.11: Elektrischer Ladungsfluss.

Beispiel 6.5.5 (Flächeninhalt). Sei $\Omega \subset \mathbb{R}^2$ eine beschränkte offene Teilmenge mit C^1 -Rand, so daß $\partial\Omega$ zusammenhängen ist. Dann existiert eine C^1 -Abbildung $\gamma : \mathbb{R} \rightarrow \partial\Omega$ mit

$$\dot{\gamma}(t) \neq 0 \quad \gamma(t + 1) = \gamma(t) \quad \text{für alle } t \in \mathbb{R},$$

so daß $\gamma|_{[0,1)} : [0, 1) \rightarrow \partial\Omega$ bijektiv ist. (Für einen Beweis siehe [3].) Diese Parametrisierung von $\partial\Omega$ kann so gewählt werden, daß γ den Rand von Ω gegen den Uhrzeigersinn durchläuft. Schreiben wir $\gamma(t) =: (x(t), y(t))$, so ist der äussere Einheitsnormalenvektor von Ω an der Stelle $\gamma(t) \in \partial\Omega$ unter diesen Voraussetzungen durch

$$\nu(\gamma(t)) = \frac{(\dot{y}(t), -\dot{x}(t))}{\sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}}$$

gegeben (siehe Abbildung 6.12).

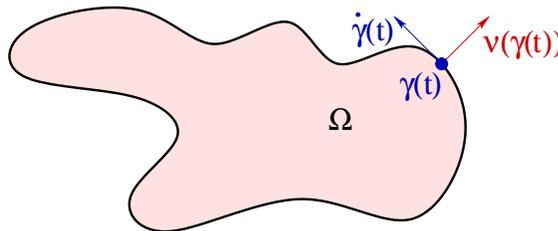


Abbildung 6.12: Der Flächeninhalt als Randintegral.

Wie in Beispiel 6.5.2 betrachten wir das Vektorfeld $f(x, y) := (x, y)$ auf dem \mathbb{R}^2 . Dann folgt aus Satz 6.4.9 mit Hilfe des Kurvenintegrals in Beispiel 6.3.3 die Gleichung

$$\begin{aligned}
 \text{Vol}_2(\Omega) &= \frac{1}{2} \int_{\Omega} \text{div}(f) \\
 &= \frac{1}{2} \int_{\partial\Omega} \langle f, \nu \rangle dS \\
 &= \frac{1}{2} \int_0^1 \langle f(\gamma(t)), \nu(\gamma(t)) \rangle \|\dot{\gamma}(t)\| dt \\
 &= \frac{1}{2} \int_0^1 (x(t)\dot{y}(t) - y(t)\dot{x}(t)) dt.
 \end{aligned} \tag{6.5.10}$$

Betrachten wir die offene Einheitskreisscheibe $\Omega = B^2 \subset \mathbb{R}^2$ mit der Parametrisierung $\gamma(t) := (\cos(2\pi t), \sin(2\pi t))$ des Randes $\partial B^2 = S^1$, so ergibt sich nochmals die Formel $\text{Vol}_2(B^2) = \pi$ für den Flächeninhalt.

Der letzte Term in (6.5.10) wird das “Integral der 1-Form $\frac{1}{2}(x dy - y dx)$ über der Kurve γ ” genannt. Diese Formulierung wird im Kapitel 7 genauer behandelt. Wie sich herausstellen wird, ist es durchaus nützlich, den Divergenzsatz von Gauß in die Sprache der Differentialformen zu übersetzen. Nicht nur führt diese Übersetzung zu einer natürlichen Verallgemeinerung in dem Satz von Stokes (Satz 7.5.4), sondern es lassen sich auch gewisse Anwendungen über den Formalismus der Differentialformen sehr viel leichter herleiten, als über die Formulierung mit der Divergenz von Vektorfeldern. Ein Beispiel dafür ist der Retraktionssatz 7.6.5 im nächsten Kapitel.

Kapitel 7

Integration von Differentialformen

Das Ziel dieses Kapitels ist es, das Integral einer Differentialform vom Grade d über einer orientierten d -dimensionalen Untermannigfaltigkeit $M \subset \mathbb{R}^n$ zu definieren. Hierzu bedarf es zunächst einer Einführung in die Begriffe der Alternierenden Formen (Abschnitt 7.1) und der Differentialformen (Abschnitt 7.2), sowie des Begriffs der Orientierung einer Untermannigfaltigkeit des \mathbb{R}^n (Abschnitt 7.3). Das Integral über einer Untermannigfaltigkeit wird im Abschnitt 7.4 eingeführt. Das wichtigste Resultat in diesem Kapitel ist der Satz von Stokes und dieser wird in Abschnitt 7.5 bewiesen. In diesem Kapitel betrachten wir stets nur glatte Untermannigfaltigkeiten, ohne dies in jedem Fall zu erwähnen.

7.1 Alternierende Formen

In diesem Abschnitt sind X, Y, Z reelle Vektorräume. Für $k \in \mathbb{N}$ bezeichnen wir mit S_k die Gruppe der Permutationen, das heißt der bijektiven Abbildungen

$$\sigma : \{1, \dots, k\} \rightarrow \{1, \dots, k\}.$$

Die Parität einer Permutation $\sigma \in S_k$ ist die Zahl

$$\varepsilon(\sigma) := (-1)^{\nu(\sigma)}, \quad \nu(\sigma) := \{(i, j) \mid 1 \leq i < j \leq k, \sigma(i) > \sigma(j)\}.$$

Die Parität definiert einen Gruppenhomomorphismus

$$\varepsilon : S_k \rightarrow \{\pm 1\}.$$

Definition 7.1.1. Sei $k \in \mathbb{N}$. Eine multi-lineare Abbildung $\omega : X^k \rightarrow \mathbb{R}$ heißt **alternierende k -Form** wenn sie für alle $i, j \in \{1, \dots, k\}$ mit $i < j$ und alle $\xi_1, \dots, \xi_k \in X$ die Bedingung

$$\omega(\xi_1, \dots, \xi_{i-1}, \xi_j, \xi_{i+1}, \dots, \xi_{j-1}, \xi_i, \xi_{j+1}, \dots, \xi_k) = -\omega(\xi_1, \dots, \xi_k)$$

erfüllt. Eine **alternierende 0-Form** auf X ist per Definition eine reelle Zahl. Für $k \in \mathbb{N}$ wird die Menge der alternierenden k -Formen auf X mit

$$\Lambda^k X^* := \{\omega : X^k \rightarrow \mathbb{R} \mid \omega \text{ ist eine alternierende } k\text{-form}\}$$

bezeichnet. Dies ist ein reeller Vektorraum. Für $k = 0$ ist

$$\Lambda^0 X^* := \mathbb{R}$$

und für $k = 1$ ist $\Lambda^1 X^* = X^* = \mathcal{L}(X, \mathbb{R})$ der Dualraum von X .

Übung 7.1.2. Für jedes $k \in \mathbb{N}$ ist $\Lambda^k X^*$ ein reeller Vektorraum.

Übung 7.1.3. Eine multi-lineare Abbildung $\omega : X^k \rightarrow \mathbb{R}$ ist genau dann eine alternierende k -Form ist wenn sie der Bedingung

$$\omega(\xi_{\sigma(1)}, \dots, \xi_{\sigma(k)}) = \varepsilon(\sigma)\omega(\xi_1, \dots, \xi_k)$$

für jede Permutation $\sigma \in S_k$ und alle Vektoren $\xi_1, \dots, \xi_k \in X$ genügt. **Hinweis 1:** Ist $\sigma \in S_k$ eine Transposition so gilt $\varepsilon(\sigma) = -1$. **Hinweis 2:** Jede Permutation läßt sich als Komposition von Transpositionen schreiben.

Übung 7.1.4. Sei $\omega \in \Lambda^k X^*$ und seien $\xi_1, \dots, \xi_k \in X$ linear abhängig. Dann gilt $\omega(\xi_1, \dots, \xi_k) = 0$. Insbesondere folgt daraus dass $\Lambda^k X^* = \{0\}$ ist für $k > \dim X$.

Beispiel 7.1.5. Sei $X = \mathbb{R}^n$ und $k \in \{1, \dots, n\}$. Für ein geordnetes k -Tupel

$$I = (i_1, \dots, i_k) \in \mathbb{N}^k, \quad 1 \leq i_1 < i_2 < \dots < i_k \leq n,$$

definieren wir die Abbildung $dx_I : (\mathbb{R}^n)^k \rightarrow \mathbb{R}$ durch die Formel

$$dx_I(\xi_1, \xi_2, \dots, \xi_k) := \det \left((\xi_{\mu i_\nu})_{\mu, \nu=1}^k \right)$$

für $\xi_\mu = (\xi_{\mu 1}, \dots, \xi_{\mu n}) \in \mathbb{R}^n$, $\mu = 1, \dots, k$. Es folgt aus den Eigenschaften der Determinante, dass dx_I in der Tat eine alternierende k -Form auf \mathbb{R}^n ist (siehe Anhang D).

Für $k = 1$ und $I = i \in \{1, \dots, n\}$ ist $dx_i : \mathbb{R}^n \rightarrow \mathbb{R}$ die Projektion auf die i te Koordinate, das heißt

$$dx_i(\xi) = \xi_i, \quad \xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n.$$

Die Menge der geordneten k -Tupel in $\{1, \dots, n\}$ bezeichnen wir mit

$$\mathcal{I}_k(n) := \{I = (i_1, \dots, i_k) \in \mathbb{N}^k \mid 1 \leq i_1 < i_2 < \dots < i_k \leq n\}.$$

Die Anzahl der geordneten k -Tupel in $\{1, \dots, n\}$ ist

$$\#\mathcal{I}_k(n) = \binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Für $k = 0$ ist es nützlich, die Konventionen $\mathcal{I}_0(n) := \{\emptyset\}$ und $dx_\emptyset := 1 \in \mathbb{R}$ zu verwenden.

Lemma 7.1.6. *Sei $n \in \mathbb{N}$ und $k \in \{0, 1, \dots, n\}$. Dann bilden die alternierenden k -Formen dx_I für $I \in \mathcal{I}_k(n)$ eine Basis von $\Lambda^k(\mathbb{R}^n)^*$. Insbesondere gilt*

$$\dim \Lambda^k(\mathbb{R}^n)^* = \binom{n}{k}.$$

Beweis. Für $k = 0$ folgt die Aussage direkt aus den Definitionen. Sei $k \geq 1$ und sei e_1, \dots, e_n die Standardbasis des \mathbb{R}^n . Dann gilt für alle $I, J \in \mathcal{I}_k(n)$ mit $J = (j_1, \dots, j_k)$ die Gleichung

$$dx_I(e_{j_1}, \dots, e_{j_k}) = \delta_{IJ} = \begin{cases} 1, & \text{falls } I = J, \\ 0, & \text{falls } I \neq J. \end{cases} \quad (7.1.1)$$

Sei nun $\omega \in \Lambda^k(\mathbb{R}^n)^*$ gegeben ebenso wie eine Abbildung $\mathcal{I}_k(n) \rightarrow \mathbb{R} : I \mapsto a_I$. Dann gilt

$$\omega = \sum_{I \in \mathcal{I}_k(n)} a_I dx_I \quad \iff \quad \begin{aligned} a_I &= \omega(e_{i_1}, \dots, e_{i_k}) \\ \forall I &= (i_1, \dots, i_k) \in \mathcal{I}_k(n). \end{aligned} \quad (7.1.2)$$

Um dies zu zeigen wählen wir ein k -Tupel $J = (j_1, \dots, j_k) \in \mathcal{I}_k(n)$ und werten die k -Formen ω und $\sum_I a_I dx_I$ auf den Vektoren $(e_{j_1}, \dots, e_{j_k})$ aus. Aufgrund von (7.1.1) erhalten wir für beide k -Formen genau dann die gleiche Antwort wenn $\omega(e_{j_1}, \dots, e_{j_k}) = a_J$ ist. Andererseits stimmen zwei k -Formen auf \mathbb{R}^n genau dann überein wenn sie auf jedem k -Tupel von Basisvektoren übereinstimmen. Damit haben wir (7.1.2) gezeigt.

Es folgt direkt aus (7.1.2) dass die k -Formen dx_I eine Basis on $\Lambda^k(\mathbb{R}^n)^*$ bilden. Damit ist Lemma 7.1.6 bewiesen. \square

Das äußere Produkt

Definition 7.1.7. Seien $k, \ell \in \mathbb{N}$. Das **äußere Produkt** von $\alpha \in \Lambda^k X^*$ und $\beta \in \Lambda^\ell X^*$ ist die alternierende $(k + \ell)$ -Form $\alpha \wedge \beta \in \Lambda^{k+\ell} X^*$, die durch

$$(\alpha \wedge \beta)(\xi_1, \dots, \xi_{k+\ell}) := \sum_{\sigma \in S_{k,\ell}} \varepsilon(\sigma) \alpha(\xi_{\sigma(1)}, \dots, \xi_{\sigma(k)}) \beta(\xi_{\sigma(k+1)}, \dots, \xi_{\sigma(k+\ell)})$$

für $\xi_1, \dots, \xi_{k+\ell} \in X$ definiert ist. Hier bezeichnen wir mit

$$S_{k,\ell} := \{\sigma \in S_{k+\ell} \mid \sigma(1) < \dots < \sigma(k), \sigma(k+1) < \dots < \sigma(k+\ell)\}$$

die Menge der (k, ℓ) -“shuffles”. Für $c \in \Lambda^0 X^* = \mathbb{R}$ und $\omega \in \Lambda^k X^*$ definieren wir

$$c \wedge \omega := c\omega.$$

Beispiel 7.1.8. Seien $\alpha, \beta \in \Lambda^1 X^*$ und $\omega \in \Lambda^2 X^*$. Dann gilt

$$\begin{aligned} (\alpha \wedge \beta)(\xi, \eta) &= \alpha(\xi)\beta(\eta) - \alpha(\eta)\beta(\xi), \\ (\alpha \wedge \omega)(\xi, \eta, \zeta) &= \alpha(\xi)\omega(\eta, \zeta) + \alpha(\eta)\omega(\zeta, \xi) + \alpha(\zeta)\omega(\xi, \eta) \end{aligned}$$

für alle $\xi, \eta, \zeta \in X$.

Lemma 7.1.9. Seien $k, \ell, m \in \mathbb{N} \cup \{0\}$.

(i) Die Abbildung

$$\Lambda^k X^* \times \Lambda^\ell X^* \rightarrow \Lambda^{k+\ell} X^* : (\alpha, \beta) \mapsto \alpha \wedge \beta$$

ist bi-linear.

(ii) Für $\alpha \in \Lambda^k X^*$ und $\beta \in \Lambda^\ell X^*$ gilt

$$\beta \wedge \alpha = (-1)^{k\ell} \alpha \wedge \beta.$$

(iii) Für $\alpha \in \Lambda^k X^*$, $\beta \in \Lambda^\ell X^*$, und $\gamma \in \Lambda^m X^*$ gilt

$$(\alpha \wedge \beta) \wedge \gamma = \alpha \wedge (\beta \wedge \gamma).$$

(iv) Für $\alpha \in \Lambda^k X^*$ und $\xi_1, \dots, \xi_k \in X$ gilt

$$(\alpha_1 \wedge \dots \wedge \alpha_k)(\xi_1, \dots, \xi_k) = \det \left((\alpha_\nu(\xi_\mu))_{\mu,\nu=1}^k \right). \quad (7.1.3)$$

(v) Für $I = (i_1, \dots, i_k) \in \mathcal{I}_k(n)$ gilt

$$dx_I = dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_k}.$$

Beweis. Die Aussagen (i) und (ii) folgen sofort aus der Definition des äusseren Produkts. Für den Beweis von (iii) definieren wir

$$S_{k,\ell,m} := \left\{ \sigma \in S_{k+\ell+m} \mid \begin{array}{l} \sigma(1) < \dots < \sigma(k), \\ \sigma(k+1) < \dots < \sigma(k+\ell), \\ \sigma(k+\ell+1) < \dots < \sigma(k+\ell+m) \end{array} \right\}.$$

Dann sind $((\alpha \wedge \beta) \wedge \gamma)(\xi_1, \dots, \xi_{k+\ell+m})$ und $(\alpha \wedge (\beta \wedge \gamma))(\xi_1, \dots, \xi_{k+\ell+m})$ beide gleich dem Ausdruck

$$\sum_{\sigma} \varepsilon(\sigma) \alpha(\xi_{\sigma(1)}, \dots, \xi_{\sigma(k)}) \beta(\xi_{\sigma(k+1)}, \dots, \xi_{\sigma(k+\ell)}) \gamma(\xi_{\sigma(k+\ell+1)}, \dots, \xi_{\sigma(k+\ell+m)}),$$

wobei die Summe über alle $\sigma \in S_{k,\ell,m}$ läuft. Hieraus folgt die Assoziativität des äusseren Produkts.

Iterieren wir die Formel im Beweis von (iii) so ergibt sich mit vollständiger Induktion die Gleichung

$$\begin{aligned} (\alpha_1 \wedge \dots \wedge \alpha_k)(\xi_1, \dots, \xi_k) &= \sum_{\sigma \in S_k} \varepsilon(\sigma) \alpha_1(\xi_{\sigma(1)}) \dots \alpha_k(\xi_{\sigma(k)}) \\ &= \det \left((\alpha_\nu(\xi_\mu))_{\mu,\nu=1}^k \right) \end{aligned}$$

für $\alpha_1, \dots, \alpha_k \in X^*$ und $\xi_1, \dots, \xi_k \in X$. Dies beweist (iv), und mit $X = \mathbb{R}^n$ und $\alpha_\nu = dx_{i_\nu}$ folgt daraus Teil (v). Damit ist Lemma 7.1.9 bewiesen. \square

Übung 7.1.10. Sei $X = \mathbb{R}^{2n}$ mit den Koordinaten $x_1, \dots, x_n, y_1, \dots, y_n$ und

$$\omega := dx_1 \wedge dy_1 + \dots + dx_n \wedge dy_n \in \Lambda^2(\mathbb{R}^{2n})^*.$$

Dann ist das n -fache äussere Produkt von ω mit sich selbst durch die Formel

$$\omega^{\wedge n} = (-1)^{n(n-1)/2} n! dx_1 \wedge \dots \wedge dx_n \wedge dy_1 \wedge \dots \wedge dy_n$$

gegeben.

Zurückholen

Definition 7.1.11. Sei $A : X \rightarrow Y$ eine lineare Abbildung und $\omega \in \Lambda^k Y^*$. Die durch A zurückgeholte alternierende k -Form $A^* \omega \in \Lambda^k X^*$ ist durch

$$(A^* \omega)(\xi_1, \dots, \xi_k) := \omega(A\xi_1, \dots, A\xi_k)$$

für $\xi_1, \dots, \xi_k \in X$ definiert.

Lemma 7.1.12. Seien $A : X \rightarrow Y$ und $B : Y \rightarrow Z$ lineare Abbildungen.

(i) Die Abbildung $\Lambda^k Y^* \rightarrow \Lambda^k X^* : \omega \mapsto A^* \omega$ ist linear.

(ii) Für $\omega \in \Lambda^k Z^*$ gilt

$$(B \circ A)^* \omega = A^* B^* \omega.$$

(iii) Für $\alpha \in \Lambda^k Y^*$ und $\beta \in \Lambda^\ell Y^*$ gilt

$$A^*(\alpha \wedge \beta) = (A^* \alpha) \wedge (A^* \beta).$$

(iv) Seien $X = \mathbb{R}^n$, $Y = \mathbb{R}^p$, und

$$A = (a_{ji})_{i=1, \dots, n}^{j=1, \dots, p} \in \mathbb{R}^{p \times n}$$

(verstanden als lineare Abbildung von \mathbb{R}^n nach \mathbb{R}^p). Wir bezeichnen die Koordinaten auf \mathbb{R}^n mit $x = (x_1, \dots, x_n)$ und die auf \mathbb{R}^p mit $y = (y_1, \dots, y_p)$. Dann gilt für $J = (j_1, \dots, j_k) \in \mathcal{I}_k(p)$ dass

$$A^* dy_J = \sum_{I \in \mathcal{I}_k(n)} \det(A_{JI}) dx_I$$

wobei

$$A_{JI} := (a_{j_\nu i_\mu})_{\mu, \nu=1}^k \in \mathbb{R}^{k \times k}$$

für $I = (i_1, \dots, i_k) \in \mathcal{I}_k(n)$.

Beweis. Die Aussagen (i), (ii), (iii) folgen sofort aus den Definitionen. Zum Beweis von (iv) bezeichnen wir mit e_1, \dots, e_n die Standardbasis des \mathbb{R}^n und wählen ein k -Tupel $I = (i_1, \dots, i_k) \in \mathcal{I}_k(n)$. Dann gilt

$$\begin{aligned} (A^* dy_J)(e_{i_1}, \dots, e_{i_k}) &= dy_J(Ae_{i_1}, \dots, Ae_{i_k}) \\ &= \det \left(((Ae_{i_\mu})_{j_\nu})_{\mu, \nu=1}^k \right) \\ &= \det(A_{JI}). \end{aligned}$$

Hier folgt die letzte Gleichung aus der Tatsache, dass die j te Koordinate des Vektors Ae_i der Eintrag $a_{ji} = (Ae_i)_j$ der Matrix A ist (i te Spalte und j te Zeile). Damit folgt die Behauptung aus (7.1.2). \square

7.2 Differentialformen

In diesem Abschnitt sind $n, p, q \in \mathbb{N} \cup \{0\}$ und

$$U \subset \mathbb{R}^n, \quad V \subset \mathbb{R}^p, \quad W \subset \mathbb{R}^q$$

nichtleere offene Mengen. Die Elemente von U werden mit $x = (x_1, \dots, x_n)$ und die von V mit $y = (y_1, \dots, y_p)$ bezeichnet.

Definition 7.2.1. Sei $k \in \mathbb{N} \cup \{0\}$. Eine **Differentialform auf U vom Grade k** ist eine glatte (d.h. C^∞) Abbildung $\omega : U \times (\mathbb{R}^n)^k \rightarrow \mathbb{R}$ so dass die Abbildung

$$(\mathbb{R}^n)^k \longrightarrow \mathbb{R} : (\xi_1, \dots, \xi_k) \mapsto \omega(x; \xi_1, \dots, \xi_k)$$

für jedes $x \in U$ eine alternierende k -Form ist.

Ist $\omega : U \times (\mathbb{R}^n)^k \rightarrow \mathbb{R}$ eine Differentialform so bezeichnen wir die durch $x \in U$ bestimmte alternierende k -Form mit $\omega_x \in \Lambda^k(\mathbb{R}^n)^*$, das heißt

$$\omega_x(\xi_1, \dots, \xi_k) := \omega(x; \xi_1, \dots, \xi_k)$$

für $\xi_1, \dots, \xi_k \in \mathbb{R}^n$. In dieser Schreibweise können wir eine Differentialform vom Grade k auch als glatte Abbildung $U \rightarrow \Lambda^k(\mathbb{R}^n)^* : x \mapsto \omega_x$ verstehen. Nach Lemma 7.1.6 können wir nun die alternierende k -Form ω_x in der Basis $\{dx_I\}_{I \in \mathcal{I}_k(n)}$ darstellen. Das heißt, es gibt glatte Funktionen $a_I : U \rightarrow \mathbb{R}$, eine für jedes k -Tupel $I = (i_1, \dots, i_k) \in \mathcal{I}_k(n)$, so dass für alle $x \in U$ gilt:

$$\omega_x = \sum_{I \in \mathcal{I}_k(n)} a_I(x) dx_I. \quad (7.2.1)$$

Der Beweis von Lemma 7.1.6 zeigt auch, dass die Koeffizienten sich in der Form $a_I(x) = \omega_x(e_{i_1}, \dots, e_{i_k})$ schreiben lassen. Die Menge der Differentialformen auf U vom Grade k bezeichnen wir mit

$$\Omega^k(U) := C^\infty(U, \Lambda^k(\mathbb{R}^n)^*).$$

Dies ist ein (unendlich dimensionaler) reeller Vektorraum. Manchmal bezeichnen wir die Elemente von $\Omega^k(U)$ kurz als k -Formen. Hierbei ist jedoch auf den Unterschied zwischen alternierenden k -Formen (Elementen von $\Lambda^k(\mathbb{R}^n)^*$) und k -Formen als Differentialformen (Element von $\Omega^k(U)$) zu achten.

Zurückholen

Definition 7.2.2. Sei $f : U \rightarrow V$ eine glatte Abbildung und $\omega \in \Omega^k(V)$. Die durch f zurückgeholte Differentialform

$$f^*\omega \in \Omega^k(U)$$

ist definiert durch

$$(f^*\omega)(x; \xi_1, \dots, \xi_k) := \omega(f(x); df(x)\xi_1, \dots, df(x)\xi_k)$$

für $x \in U$ und $\xi_1, \dots, \xi_k \in \mathbb{R}^n$.

Lemma 7.2.3. Seien $f : U \rightarrow V$ und $g : V \rightarrow W$ glatte Abbildungen.

- (i) Die Abbildung $\Omega^k(V) \rightarrow \Omega^k(U) : \omega \mapsto f^*\omega$ ist linear.
(ii) Für $\omega \in \Omega^k(W)$ gilt

$$(g \circ f)^*\omega = f^*g^*\omega.$$

- (iii) Für $\alpha \in \Omega^k(V)$ und $\beta \in \Omega^\ell(V)$ gilt

$$f^*(\alpha \wedge \beta) = (f^*\alpha) \wedge (f^*\beta).$$

- (iv) Ist

$$\omega = \sum_{J \in \mathcal{I}_k(p)} b_J dy_J \in \Omega^k(V)$$

mit $b_J \in C^\infty(V)$ so gilt

$$f^*\omega = \sum_{J \in \mathcal{I}_k(p)} \sum_{I \in \mathcal{I}_k(n)} (b_J \circ f) \det \left(\frac{\partial f_J}{\partial x_I} \right) dx_I, \quad (7.2.2)$$

wobei

$$\frac{\partial f_J}{\partial x_I} := \left(\frac{\partial f_{j_\nu}}{\partial x_{i_\mu}} \right)_{\nu, \mu=1}^k \in C^\infty(U, \mathbb{R}^{k \times k})$$

für $I = (i_1, \dots, i_k) \in \mathcal{I}_k(n)$ und $J = (j_1, \dots, j_k) \in \mathcal{I}_k(p)$.

Beweis. Die Aussagen (i), (ii), (iii) folgen sofort aus den Definitionen und (iv) folgt aus Lemma 7.1.12 (iv). \square

Beispiel 7.2.4. Sei $V = \mathbb{R}^2$ mit den Koordinaten (x, y) und $U = \mathbb{R}^2$ mit den Koordinaten (r, θ) . Sei $f : U \rightarrow V$ die Abbildung

$$f(r, \theta) := (r \cos(\theta), r \sin(\theta))$$

Also ist

$$x = f_1(r, \theta) = r \cos(\theta), \quad y = f_2(r, \theta) = r \sin(\theta).$$

Die Formel (7.2.2) für das Zurückholen durch f ergibt

$$f^* dx = \frac{\partial f_1}{\partial r} dr + \frac{\partial f_1}{\partial \theta} d\theta = \cos(\theta) dr - r \sin(\theta) d\theta$$

und

$$f^* dy = \frac{\partial f_2}{\partial r} dr + \frac{\partial f_2}{\partial \theta} d\theta = \sin(\theta) dr + r \cos(\theta) d\theta.$$

Hieraus folgt nach Lemma 7.2.3 (iii) dass

$$f^*(dx \wedge dy) = r dr \wedge d\theta.$$

Wir haben dabei verwendet dass

$$d\theta \wedge dr = -dr \wedge d\theta$$

ist. Dies stimmt überein mit der Gleichung

$$\det(df(r, \theta)) = r.$$

Das Differential

Definition 7.2.5. Sei $k \in \mathbb{N} \cup \{0\}$ und $\omega \in \Omega^k(U)$. Das **Differential** von ω ist die Differentialform $d\omega \in \Omega^{k+1}(U)$, die durch

$$d\omega(x; \xi_0, \dots, \xi_k) := \sum_{i=0}^k \sum_{\nu=1}^n (-1)^i \xi_{i\nu} \frac{\partial \omega}{\partial x_\nu}(x; \xi_0, \dots, \xi_{i-1}, \xi_{i+1}, \dots, \xi_k) \quad (7.2.3)$$

für $x \in U$ und $\xi_0, \dots, \xi_k \in \mathbb{R}^n$ definiert ist. Die resultierende Abbildung

$$d : \Omega^k(U) \rightarrow \Omega^{k+1}(U)$$

ist ein linearer Differentialoperator erster Ordnung.

Bemerkung 7.2.6. Eine 0-Form auf U ist eine glatte Funktion $f : U \rightarrow \mathbb{R}$. Das Differential $df \in \Omega^1(U)$ einer solchen 0-Form $f \in \Omega^0(U) = C^\infty(U)$ ist nach Definition 7.2.5 die durch

$$df(x; \xi) := \sum_{\nu=1}^n \xi_{i\nu} \frac{\partial f}{\partial x_\nu}(x)$$

für $x \in U$ und $\xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$ definierte 1-Form. Also ist

$$(df)_x = df(x) \in (\mathbb{R}^n)^*$$

in der Tat die Ableitung von f an der Stelle $x \in U$ wie wir sie kennen. In etwas anderer Schreibweise haben wir

$$df = \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i.$$

Angewendet auf die Funktion $U \rightarrow \mathbb{R} : x = (x_1, \dots, x_n) \mapsto x_i$ rechtfertigt dies nachträglich die Bezeichnung dx_i für die Projektion auf die i te Koordinate.

Satz 7.2.7. (i) Ist $\omega = \sum_{I \in \mathcal{I}_k(n)} a_I dx_I$ mit $a_I \in C^\infty(U)$ so gilt

$$d\omega = \sum_{\nu=1}^n \sum_{I \in \mathcal{I}_k(n)} \frac{\partial a_I}{\partial x_\nu} dx_\nu \wedge dx_I = \sum_{I \in \mathcal{I}_k(n)} da_I \wedge dx_I. \quad (7.2.4)$$

(ii) Für $\alpha \in \Omega^k(U)$ und $\beta \in \Omega^\ell(U)$ gilt

$$d(\alpha \wedge \beta) = (d\alpha) \wedge \beta + (-1)^k \alpha \wedge (d\beta)$$

(iii) Für jedes $\omega \in \Omega^k(U)$ gilt

$$d(d\omega) = 0.$$

(iv) Ist $f : U \rightarrow V$ eine glatte Abbildung so gilt für alle $\omega \in \Omega^k(V)$ dass

$$d(f^*\omega) = f^*(d\omega).$$

Beweis. Wir zeigen, dass die Formeln (7.2.3) und (7.2.4) für $d\omega$ übereinstimmen. Ist $\omega = \sum_I a_I dx_I$ so erhalten wir für $x \in U$ und $\xi_0, \dots, \xi_k \in \mathbb{R}^n$:

$$\begin{aligned}
d\omega(x; \xi_0, \dots, \xi_k) &= \sum_{i=0}^k \sum_{\nu=1}^n (-1)^i \xi_{i\nu} \frac{\partial \omega}{\partial x_\nu}(x; \xi_0, \dots, \xi_{i-1}, \xi_{i+1}, \dots, \xi_k) \\
&= \sum_{i=0}^k \sum_{\nu=1}^n \sum_I (-1)^i \xi_{i\nu} \frac{\partial a_I}{\partial x_\nu} dx_I(\xi_0, \dots, \xi_{i-1}, \xi_{i+1}, \dots, \xi_k) \\
&= \sum_{\nu=1}^n \sum_I \frac{\partial a_I}{\partial x_\nu} \left(\sum_{i=0}^k (-1)^i \xi_{i\nu} dx_I(\xi_0, \dots, \xi_{i-1}, \xi_{i+1}, \dots, \xi_k) \right) \\
&= \sum_{\nu=1}^n \sum_I \frac{\partial a_I}{\partial x_\nu} (dx_\nu \wedge dx_I)(\xi_0, \dots, \xi_k).
\end{aligned}$$

Damit ist (i) bewiesen.

Wir beweisen (ii). Da $d : \Omega^k(U) \rightarrow \Omega^{k+1}(U)$ ein linearer Operator ist, genügt es, die Leibnitz-Regel für $\alpha = a dx_I$ und $\beta = b dx_J$ zu zeigen, wobei $a, b : U \rightarrow \mathbb{R}$ glatte Funktionen sind und $I \in \mathcal{I}_k(n)$, $J \in \mathcal{I}_\ell(n)$. In diesem Fall erhalten wir

$$\begin{aligned}
d(\alpha \wedge \beta) &= d(ab dx_I \wedge dx_J) = \sum_{\nu=1}^n \frac{\partial(ab)}{\partial x_\nu} dx_\nu \wedge dx_I \wedge dx_J \\
&= \sum_{\nu=1}^n \frac{\partial a}{\partial x_\nu} b dx_\nu \wedge dx_I \wedge dx_J + \sum_{\nu=1}^n a \frac{\partial b}{\partial x_\nu} dx_\nu \wedge dx_I \wedge dx_J \\
&= \left(\sum_{\nu=1}^n \frac{\partial a}{\partial x_\nu} dx_\nu \wedge dx_I \right) \wedge \beta + (-1)^k \alpha \wedge \left(\sum_{\nu=1}^n \frac{\partial b}{\partial x_\nu} dx_\nu \wedge dx_J \right) \\
&= d\alpha \wedge \beta + (-1)^k \alpha \wedge d\beta.
\end{aligned}$$

Hier haben wir in der zweiten und letzten Gleichung die Formel (7.2.4) verwendet. Damit ist (ii) bewiesen.

Wir beweisen (iii). Für eine 0-Form $f \in \Omega^0(U) = C^\infty(U)$ gilt

$$\begin{aligned}
d(df) &= d\left(\sum_{\nu=1}^n \frac{\partial f}{\partial x_\nu} dx_\nu \right) = \sum_{\mu, \nu=1}^n \frac{\partial}{\partial x_\mu} \frac{\partial f}{\partial x_\nu} dx_\mu \wedge dx_\nu \\
&= \sum_{\mu < \nu} \left(\frac{\partial}{\partial x_\mu} \frac{\partial f}{\partial x_\nu} - \frac{\partial}{\partial x_\nu} \frac{\partial f}{\partial x_\mu} \right) dx_\mu \wedge dx_\nu = 0.
\end{aligned}$$

Außerdem gilt, nach Definition, $d(dx_I) = 0$ für jedes k -Tupel $I \in \mathcal{I}_k(n)$. Für $\omega = \sum_I a_I dx_I$ ergibt sich daher aus (i) und (ii) die Gleichung

$$d(d\omega) = d\left(\sum_I da_I \wedge dx_I\right) = \sum_I d(da_I) \wedge dx_I = 0.$$

Damit ist (iii) bewiesen.

Wir beweisen (iv). Ist $g \in \Omega^0(V) = C^\infty(V)$ so ist $f^*g = g \circ f$ und daher ist die Formel

$$f^*(dg) = d(f^*g) \tag{7.2.5}$$

eine Umformulierung der Kettenregel. Mit $g(y) = y_j$ gilt insbesondere

$$f^*dy_j = df_j$$

für $j = 1, \dots, p$. Daraus folgt für $J = (j_1, \dots, j_k) \in \mathcal{I}_k(p)$

$$f^*dy_J = df_{j_1} \wedge \dots \wedge df_{j_k}.$$

Wenden wir nun die Leibnitz-Regel in (ii) auf diese Formel an so ergibt sich (mit vollständiger Induktion)

$$d(f^*dy_J) = 0. \tag{7.2.6}$$

Für eine k -Form $\omega = \sum_{J \in \mathcal{I}_k(p)} b_J dy_J$ mit $b_J \in C^\infty(V)$ gilt daher

$$\begin{aligned} d(f^*\omega) &= d\left(\sum_{J \in \mathcal{I}_k(p)} (f^*b_J)(f^*dy_J)\right) \\ &= \sum_{J \in \mathcal{I}_k(p)} (d(f^*b_J)) \wedge (f^*dy_J) \\ &= \sum_{J \in \mathcal{I}_k(p)} (f^*(db_J)) \wedge (f^*dy_J) \\ &= f^*\left(\sum_{J \in \mathcal{I}_k(p)} (db_J) \wedge dy_J\right) \\ &= f^*(d\omega) \end{aligned}$$

Hier folgt die zweite Gleichung aus (7.2.6) und der Leibnitz-Regel in (ii), die dritte aus (7.2.5), die vierte aus Lemma 7.2.3 (ii), und die letzte aus (7.2.4). Damit ist Satz 7.2.7 bewiesen. \square

Definition 7.2.8. Eine Differentialform $\omega \in \Omega^k(U)$ heißt **geschlossen** wenn $d\omega = 0$ ist. Sie heißt **exakt** wenn eine $(k-1)$ -Form $\alpha \in \Omega^{k-1}(U)$ existiert, so dass $d\alpha = \omega$ ist.

Nach Teil (iii) von Satz 7.2.7 ist jede exakte k -Form geschlossen. Wir werden aber sehen dass eine geschlossene k -Form nicht unbedingt exakt sein muss. Die Übung 7.2.10 wird jedoch zeigen, dass für sternförmige Gebiete $U \subset \mathbb{R}^n$ jede geschlossene k -Form auf U mit $k \geq 1$ notwendigerweise auch exakt ist. Mit anderen Worten, für jede offene Teilmenge $U \subset \mathbb{R}^n$ haben wir eine endliche Folge von linearen Operatoren

$$\Omega^0(U) \xrightarrow{d} \Omega^1(U) \xrightarrow{d} \Omega^2(U) \xrightarrow{d} \dots \xrightarrow{d} \Omega^{n-1}(U) \xrightarrow{d} \Omega^n(U)$$

mit der Eigenschaft, dass das Bild eines jeden Operators enthalten ist im Kern des darauf folgenden. Eine solche Folge von Operatoren wird auch **Kettenkomplex** genannt. Wenn das Bild des Operators $d : \Omega^{k-1}(U) \rightarrow \Omega^k(U)$ ein echter Unterraum vom Kern des Operators $d : \Omega^k(U) \rightarrow \Omega^{k+1}(U)$ ist, so ist der Quotient dieser Räume ein reeller Vektorraum positiver Dimension. Diese Quotientenräume werden **deRham Kohomologiegruppen** von U genannt und mit $H^k(U; \mathbb{R})$ bezeichnet. Sie enthalten interessante Informationen über die Eigenschaften der Menge U . Dieses Thema geht jedoch weit über die Inhalte dieser Vorlesung hinaus.

Differentialformen in der Dimension drei

Beispiel 7.2.9. Wir betrachten den 3-dimensionalen Euklidischen Raum \mathbb{R}^3 mit den Koordinaten x, y, z . Dann ist

$$\Lambda^0(\mathbb{R}^3)^* = \mathbb{R}, \quad \Lambda^1(\mathbb{R}^3)^* \cong \mathbb{R}^3, \quad \Lambda^2(\mathbb{R}^3)^* \cong \mathbb{R}^3, \quad \Lambda^3(\mathbb{R}^3)^* \cong \mathbb{R}.$$

Der Isomorphismus $\Lambda^1(\mathbb{R}^3)^* \cong \mathbb{R}^3$ ist gegeben durch die Basis dx, dy, dz und der Isomorphismus $\Lambda^2(\mathbb{R}^3)^* \cong \mathbb{R}^3$ durch die Basis $dy \wedge dz, dz \wedge dx, dx \wedge dy$.

Sei $U \subset \mathbb{R}^3$ eine offene Menge und $\text{Vekt}(U) := C^\infty(U, \mathbb{R}^3)$ der Raum der glatten Vektorfelder auf U . Die einem Vektorfeld $f = (a, b, c) : U \rightarrow \mathbb{R}^3$ zugeordnete 1-Form und 2-Form sind

$$\alpha_f := a dx + b dy + c dz, \quad \omega_f := a dy \wedge dz + b dz \wedge dx + c dx \wedge dy. \quad (7.2.7)$$

Dies ergibt Isomorphismen $\text{Vekt}(U) \rightarrow \Omega^1(U)$ und $\text{Vekt}(U) \rightarrow \Omega^2(U)$. Das Differential der 1-Form α_f ist

$$d\alpha_f = \left(\frac{\partial c}{\partial y} - \frac{\partial b}{\partial z} \right) dy \wedge dz + \left(\frac{\partial a}{\partial z} - \frac{\partial c}{\partial x} \right) dz \wedge dx + \left(\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y} \right) dx \wedge dy.$$

Diese 2-Form wird wiederum mit dem Vektorfeld

$$\operatorname{rot}(f) := \begin{pmatrix} \partial c/\partial y - \partial b/\partial z \\ \partial a/\partial z - \partial c/\partial x \\ \partial b/\partial x - \partial a/\partial y \end{pmatrix} = \begin{pmatrix} \partial/\partial x \\ \partial/\partial y \\ \partial/\partial z \end{pmatrix} \times \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \nabla \times f$$

identifiziert. Hier verwenden wir das Kreuzprodukt und rechnen formal mit dem Ausdruck $\nabla := (\partial/\partial x, \partial/\partial y, \partial/\partial z)$ wie mit einem Vektor. Ebenso erhalten wir

$$d\omega_f = \left(\frac{\partial a}{\partial x} + \frac{\partial b}{\partial y} + \frac{\partial c}{\partial z} \right) dx \wedge dy \wedge dz = (\nabla \cdot f) dx \wedge dy \wedge dz,$$

wobei mit $\nabla \cdot f = \langle \nabla, f \rangle$ formal das standard innere Produkt gemeint ist. Zusammenfassend ergibt sich

$$du = \alpha_{\operatorname{grad}(u)}, \quad d\alpha_f = \omega_{\operatorname{rot}(f)}, \quad d\omega_f = \operatorname{div}(f) dx \wedge dy \wedge dz \quad (7.2.8)$$

für $u \in C^\infty(U)$ und $f \in \operatorname{Vekt}(U)$ mit

$$\operatorname{grad}(u) = \nabla u, \quad \operatorname{rot}(f) = \nabla \times f, \quad \operatorname{div}(f) = \nabla \cdot f$$

Es folgt nun aus Satz 7.2.7 (oder auch durch direktes Nachrechnen) dass

$$\operatorname{rot}(\operatorname{grad}(u)) = 0, \quad \operatorname{div}(\operatorname{rot}(f)) = 0 \quad (7.2.9)$$

für $u \in C^\infty(U)$ und $f \in \operatorname{Vekt}(U)$. Außerdem gilt

$$\operatorname{div}(\operatorname{grad}(u)) = \Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}. \quad (7.2.10)$$

Dies ist der **Laplace-Operator**.

Übungen

Übung 7.2.10 (Lemma von Poincaré). Sei $U \subset \mathbb{R}^n$ offen und **sternförmig** (das heißt wenn $x \in U$ ist dann ist auch $tx \in U$ für jedes $t \in [0, 1]$). Sei $k \in \mathbb{N}$ und $\omega \in \Omega^k(U)$ eine geschlossene k -Form. Dann ist ω exakt. **Hinweis:** Definiere $\alpha \in \Omega^{k-1}(U)$ durch

$$\alpha(x; \xi_1, \dots, \xi_{k-1}) := \int_0^1 \omega(tx; x, \xi_1, \dots, \xi_{k-1}) t^{k-1} dt$$

für $x \in U$ und $\xi_1, \dots, \xi_{k-1} \in \mathbb{R}^n$. Dann gilt $d\alpha = \omega$.

Übung 7.2.11. Sei $U \subset \mathbb{R}^n$ eine offene Menge, $f = (f_1, \dots, f_n) : U \rightarrow \mathbb{R}^n$ ein Vektorfeld, und $\omega_f \in \Omega^{n-1}(U)$ gegeben durch

$$\omega_f(x; \xi_1, \dots, \xi_{n-1}) := \det(f(x), \xi_1, \dots, \xi_{n-1})$$

für $x \in U$ und $\xi_1, \dots, \xi_{n-1} \in \mathbb{R}^n$. Dann gilt

$$\omega_f = \sum_{i=1}^n (-1)^{i-1} f_i(x) dx_1 \wedge \dots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \dots \wedge dx_n, \quad (7.2.11)$$

$$d\omega_f = \operatorname{div}(f) dx_1 \wedge \dots \wedge dx_n. \quad (7.2.12)$$

Übung 7.2.12. Der *-Operator $* : \Omega^k(\mathbb{R}^n) \rightarrow \Omega^{n-k}(\mathbb{R}^n)$ ist durch

$$*dx_I := \varepsilon(\sigma) dx_{*I}$$

für $I = (i_1, \dots, i_k) \in \mathcal{I}_k(n)$ definiert, wobei $*I = (j_1, \dots, j_{n-k}) \in \mathcal{I}_{n-k}(n)$ so gewählt ist, dass $\{1, \dots, n\} = \{i_1, \dots, i_k, j_1, \dots, j_{n-k}\}$, und die Permutation $\sigma \in S_n$ durch $\sigma(\nu) = i_\nu$ für $\nu = 1, \dots, k$ und $\sigma(\nu) = j_{\nu-k}$ für $\nu = k+1, \dots, n$ definiert ist. Da der *-Operator linear ist, ist er durch die Bilder der Basisvektoren dx_I eindeutig bestimmt. Er hat folgende Eigenschaften.

(i) Für jedes $I \in \mathcal{I}_k(n)$ gilt

$$**dx_I = (-1)^{k(n-k)} dx_I, \quad dx_I \wedge *dx_I = dx_1 \wedge \dots \wedge dx_n.$$

Außerdem ist $dx_I \wedge *dx_J = 0$ für $I, J \in \mathcal{I}_k(n)$ mit $I \neq J$.

(ii) Sei $U \subset \mathbb{R}^n$ eine offene Menge, $f = (f_1, \dots, f_n) : U \rightarrow \mathbb{R}^n$ ein glattes Vektorfeld, und

$$\alpha_f := \sum_{i=1}^n f_i(x) dx_i \in \Omega^1(U).$$

Dann gilt $*\alpha_f = \omega_f$ und $d(*\alpha_f) = \operatorname{div}(f) dx_1 \wedge \dots \wedge dx_n$ (siehe Übung 7.2.11). Für jede glatte Funktion $u : \mathbb{R}^n \rightarrow \mathbb{R}$ gilt

$$*d*du = \Delta u := \frac{\partial^2 u}{\partial x_1^2} + \dots + \frac{\partial^2 u}{\partial x_n^2}.$$

(iii) Für $n = 3$ gilt

$$*dx = dy \wedge dz, \quad *dy = dz \wedge dx, \quad *dz = dx \wedge dy,$$

sowie $*\omega_f = \alpha_f$ und $*d\alpha_f = \alpha_{\operatorname{rot}(f)}$ für jedes Vektorfeld $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$.

7.3 Orientierung

Unser Ziel ist es, Differentialformen vom Grade d über Untermannigfaltigkeiten der Dimension d zu integrieren. Dabei spielt der Begriff der Orientierung eine zentrale Rolle. Den Ausgangspunkt bildet folgende Definition.

Orientierte Vektorräume

Sei X ein d -dimensionaler reeller Vektorraum. Zwei geordnete Basen

$$e_1, \dots, e_d, \quad \tilde{e}_1, \dots, \tilde{e}_d$$

haben **die gleiche Orientierung** wenn die $(d \times d)$ -Matrix $A = (a_{ij})_{i,j=1}^d$, die durch die Formel

$$\tilde{e}_i = \sum_{j=1}^d a_{ij} e_j, \quad i = 1, \dots, d,$$

definiert ist, positive Determinante hat. Der Begriff “gleiche Orientierung” definiert eine Äquivalenzrelation auf der Menge der Basen von X mit genau zwei Äquivalenzklassen. Eine **Orientierung** von X ist die Wahl einer dieser Äquivalenzklassen von Basen. Ist eine Orientierung von X gegeben so nennen wir die Basen in dieser ausgewählten Äquivalenzklasse **positiv** und die anderen Basen **negativ**. Ein **orientierter Vektorraum** ist ein endlichdimensionaler reeller Vektorraum zusammen mit einer Orientierung.

Eine Orientierung von X wird also durch eine Basis bestimmt. Die positiven (bzw. negativen) Basen sind dann jene die aus dieser Basis durch eine Transformation mit positiver (bzw. negativer) Determinante hervorgehen. Vertauschen wir zum Beispiel zwei Elemente einer positiven Basis so erhalten wir eine negative Basis. Die **Standardorientierung** des \mathbb{R}^d wird durch die Standardbasis

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad e_d = \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

bestimmt. Ist $\Phi : X \rightarrow Y$ ein Vektorraumisomorphismus zwischen zwei orientierten Vektorräumen, so nennen wir ihn **orientierungserhaltend** wenn er

positive Basen von X in positive Basen von Y überführt und **orientierungsumkehrend** wenn er positive Basen von X in negative Basen von Y überführt. Ein orientierungserhaltender Isomorphismus überführt natürlich auch negative Basen von X in negative Basen von Y und ein orientierungsumkehrender Isomorphismus überführt negative Basen von X in positive Basen von Y . Mit anderen Worten, stellen wir die lineare Abbildung Φ als Matrix bezüglich zweier positiver Basen von X und Y dar, so ist Φ genau dann orientierungserhaltend (beziehungsweise orientierungsumkehrend), wenn die Determinante dieser Matrix positiv (beziehungsweise negativ) ist.

Untermannigfaltigkeiten

Wir wollen nun diesen Begriff der Orientierung auf d -dimensionale glatte (C^∞) Untermannigfaltigkeiten $M \subset \mathbb{R}^n$ übertragen. Zur Erinnerung diskutieren wir nochmals die zugrundeliegenden Begriffe (Definition 6.1.9). Es sei hier noch einmal betont, dass in diesem Kapitel alle Diffeomorphismen C^∞ -Diffeomorphismen sind.

Gegeben seien ganze Zahlen d, n mit $0 \leq d \leq n$. Es sei daran erinnert, dass eine Teilmenge $M \subset \mathbb{R}^n$ eine *d -dimensionale glatte Untermannigfaltigkeit* von \mathbb{R}^n genannt wird, wenn es für jeden Punkt $p_0 \in M$ eine M -offene Teilmenge $U_0 \subset M$ gibt, die p_0 enthält und zu einer offenen Teilmenge $V_0 \subset \mathbb{R}^d$ diffeomorph ist. Einen solchen Diffeomorphismus $\phi_0 : U_0 \rightarrow V_0$ nennen wir eine *Karte* von M , die M -offene Menge U_0 nennen wir ein *Kartengebiet* von M , und die Umkehrabbildung $\psi_0 := \phi_0^{-1} : V_0 \rightarrow U_0$ nennen wir eine *glatte Parametrisierung* des Kartengebiets.

Bei dieser Formulierung sind verschiedene Dinge zu beachten. Zunächst einmal ist die Teilmenge $U_0 \subset M$, von der hier die Rede ist, **keine** offene Teilmenge von \mathbb{R}^n (ausser im Fall $d = n$) sondern sie ist lediglich offen bezüglich der Relativtopologie von M . Nach Lemma B.1.2 heißt das, dass eine offene Teilmenge $U \subset \mathbb{R}^n$ mit $U \cap M = U_0$ existiert. Wenn wir also von einer glatten Abbildung $\phi_0 : U_0 \rightarrow \mathbb{R}^d$ sprechen, so bedeutet dies, dass eine glatte Abbildung $\phi : U \rightarrow \mathbb{R}^d$ auf einer offenen Menge $U \subset \mathbb{R}^n$ existiert, so dass $U \cap M = U_0$ ist und die Einschränkung von ϕ auf $U \cap M = U_0$ mit ϕ_0 übereinstimmt. Eine Karte $\phi_0 : U_0 \rightarrow V_0$ ist also in diesem Sinne eine glatte Abbildung, die zusätzlich noch bijektiv ist und eine glatte Umkehrabbildung besitzt, eben die genannte Parametrisierung $\psi_0 : V_0 \rightarrow U_0$ des Kartengebiets. Diese Umkehrabbildung ist natürlich glatt im üblichen Sinne, da V_0 eine offene Teilmenge von \mathbb{R}^d ist und wir ψ_0 als Abbildung von V_0 nach \mathbb{R}^n betrach-

ten können (und dabei vorübergehend ignorieren, dass die Werte von ψ_0 tatsächlich in $U_0 \subset M$ enthalten sind).

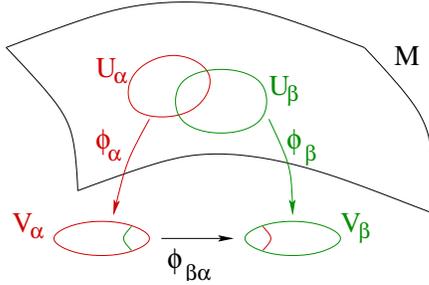


Abbildung 7.1: Übergangsabbildungen.

Kehren wir nun zurück zu unserer Untermannigfaltigkeit M , so können wir also M (nach Definition) durch Kartengebiete U_α (mit Indizes α in einer Menge A) überdecken und können dazu noch glatte Karten $\phi_\alpha : U_\alpha \rightarrow V_\alpha$ wählen. Ein solches System von Karten $\{U_\alpha, \phi_\alpha\}_{\alpha \in A}$ heißt **Atlas** von M . Ist M kompakt so folgt aus der Charakterisierung der Kompaktheit durch die Überdeckungseigenschaft in Satz C.1.2, dass M einen endlichen Atlas besitzt. Für je zwei Karten

$$\phi_\alpha : U_\alpha \rightarrow V_\alpha, \quad \phi_\beta : U_\beta \rightarrow V_\beta$$

spielen die **Übergangsabbildungen**

$$\phi_{\beta\alpha} := \phi_\beta \circ \phi_\alpha^{-1} : \phi_\alpha(U_\alpha \cap U_\beta) \rightarrow \phi_\beta(U_\alpha \cap U_\beta)$$

eine wichtige Rolle (siehe Abbildung 7.1). Da ϕ_α und ϕ_β insbesondere Homöomorphismen sind (also bijektiv und in beide Richtungen stetig), und $U_\alpha \cap U_\beta$ eine M -offene Teilmenge von M ist, sind auch die Teilmengen

$$\phi_\alpha(U_\alpha \cap U_\beta) \subset V_\alpha, \quad \phi_\beta(U_\alpha \cap U_\beta) \subset V_\beta$$

offene Mengen in \mathbb{R}^d und $\phi_{\beta\alpha}$ ist ein glatter Diffeomorphismus zwischen diesen Mengen.

Sei $M \subset \mathbb{R}^n$ eine d -dimensionale glatte Untermannigfaltigkeit. Der **Tangententialraum** $T_p M \subset \mathbb{R}^n$ ist die Menge

$$T_p M := \{ \dot{\gamma}(0) \mid \gamma \in C^1(\mathbb{R}, \mathbb{R}^n), \gamma(t) \in M \forall t \in \mathbb{R}, \gamma(0) = p \}.$$

Wir hatten in Kapitel 3 gezeigt, dass T_pM ein d -dimensionaler linearer Unterraum von \mathbb{R}^n ist und dass

$$T_pM = \text{im } d\psi_\alpha(x)$$

ist für jede glatte Parametrisierung $\psi_\alpha : V_\alpha \rightarrow U_\alpha$ eines Kartengebiets U_α mit $p \in U_\alpha$ und $x = \psi_\alpha^{-1}(p) \in V_\alpha$. Insbesondere ist die Ableitung

$$d\psi_\alpha(x) : \mathbb{R}^d \rightarrow T_pM, \quad p := \psi_\alpha(x),$$

ein Vektorraumisomorphismus für alle $\alpha \in A$ und $x \in V_\alpha$.

Übung 7.3.1. Sei $M \subset \mathbb{R}^n$ eine glatte d -dimensionale Untermannigfaltigkeit. Dann ist das **Tangentialbündel**

$$TM := \{(p, v) \in \mathbb{R}^n \times \mathbb{R}^n \mid p \in M, v \in T_pM\}$$

eine glatte $2d$ -dimensionale Untermannigfaltigkeit von \mathbb{R}^{2n} .

Beispiel 7.3.2. Zur Erinnerung nochmals ein Beispiel. Die Menge

$$M := \{(x, y) \in \mathbb{R}^2 \mid xy = 0\}$$

ist keine Untermannigfaltigkeit von \mathbb{R}^2 . Zwar ist die Teilmenge

$$U_0 := \{(x, y) \in \mathbb{R}^2 \mid -1 < x < 1, y = 0\} \subset M$$

diffeomorph zu einer offenen Teilmenge von \mathbb{R} , jedoch ist U_0 nicht M -offen. In Kapitel 3 hatten wir gezeigt dass eine Teilmenge $M \subset \mathbb{R}^n$ genau dann eine glatte d -dimensionale Untermannigfaltigkeit ist, wenn es für jedes $p_0 \in M$ eine offene Teilmenge $U \subset \mathbb{R}^n$ und eine glatte Abbildung

$$f : U \rightarrow \mathbb{R}^{n-d}$$

gibt, so dass $p_0 \in U$ ist, 0 ein regulärer Wert von f ist, und

$$U \cap M = \{x \in U \mid f(x) = 0\}.$$

In unserem Beispiel wäre $n = 2$ und $d = 1$, jedoch besitzt der Punkt

$$p_0 = (0, 0)$$

keine solche Umgebung. Denn wenn $f : U \rightarrow \mathbb{R}$ auf $U \cap M$ verschwindet so ist auch $df(0, 0) = 0$ und somit ist 0 kein regulärer Wert von f .

Orientierte Untermannigfaltigkeiten

Definition 7.3.3. Sei $M \subset \mathbb{R}^n$ eine glatte d -dimensionale Untermannigfaltigkeit. Eine **Orientierung von M** ist die Wahl eine Orientierung für jeden der Tangentialräume $T_p M$ so dass diese Orientierungen wie folgt verknüpft sind. Für jeden Punkt $p_0 \in M$ existiert ein Kartengebiet $U_0 \subset M$ mit $p_0 \in U_0$ und eine glatte Parametrisierung

$$\psi_0 : V_0 \rightarrow U_0$$

(auf einer offenen Teilmenge $V_0 \subset \mathbb{R}^d$) so dass die Ableitung

$$d\psi_0(x) : \mathbb{R}^d \rightarrow T_{\psi_0(x)}M$$

für jedes $x \in V_0$ ein orientierungserhaltender Isomorphismus ist (das heißt, der Isomorphismus $d\psi_0(x)$ überführt die Standardbasis des \mathbb{R}^d in eine positive Basis von $T_{\psi_0(x)}M$). Eine solches ψ_0 nennen wir eine **orientierte Parametrisierung** von U_0 und die Umkehrabbildung

$$\phi_0 := \psi_0^{-1} : U_0 \rightarrow V_0$$

eine **orientierte Karte**. M heißt **orientierbar** wenn eine solche Orientierung existiert. Eine **orientierte Untermannigfaltigkeit** von \mathbb{R}^n ist eine Untermannigfaltigkeit zusammen mit einer Orientierung.

Ist $M \subset \mathbb{R}^n$ eine d -dimensionale orientierte Untermannigfaltigkeit so existiert offensichtlich ein Atlas $\{U_\alpha, \phi_\alpha\}_{\alpha \in A}$, der ausschliesslich aus orientierten Karten besteht. In diesem Fall sind alle Übergangsabbildungen $\phi_{\beta\alpha}$ **orientierungserhaltend**, das heißt für $\alpha, \beta \in A$ und $x \in \phi_\alpha(U_\alpha \cap U_\beta)$ gilt

$$\det(d\phi_{\beta\alpha}(x)) > 0. \quad (7.3.1)$$

Wir sprechen dann von einem **orientierten Atlas**. Umgekehrt induziert jeder orientierte Atlas eine Orientierung von M ; wir definieren die Orientierung von $T_p M$ als die durch den Isomorphism

$$d\psi_\alpha(x) : \mathbb{R}^d \rightarrow T_p M$$

induzierte Orientierung, wobei $\alpha \in A$ so gewählt ist, dass $p \in U_\alpha$ ist, und

$$x := \phi_\alpha(p), \quad \psi_\alpha := \phi_\alpha^{-1}.$$

Nach (7.3.1) ist diese Definition der Orientierung von $T_p M$ unabhängig von α .

Beispiel 7.3.4. Die Einheitssphäre

$$S^{n-1} := \{x \in \mathbb{R}^n \mid \|x\| = 1\}$$

ist eine orientierte Untermannigfaltigkeit von \mathbb{R}^n . Eine Basis v_1, \dots, v_{n-1} von $T_x S^{n-1} = x^\perp$ ist genau dann positiv wenn x, v_1, \dots, v_{n-1} eine positive Basis des \mathbb{R}^n ist.

Beispiel 7.3.5. Der standard n -Torus

$$\mathbb{T}^n := \{z = (z_1, \dots, z_n) \in \mathbb{C}^n \mid |z_i| = 1 \text{ für } i = 1 \dots, n\}.$$

ist orientiert. Der Tangentialraum $T_z \mathbb{T}^n = \{(\mathbf{i}t_1 z_1, \dots, \mathbf{i}t_n z_n) \mid t_\nu \in \mathbb{R}\}$ besitzt einen kanonischen Isomorphismus zu \mathbb{R}^n (nämlich $t \mapsto (\mathbf{i}t_1 z_1, \dots, \mathbf{i}t_n z_n)$) und dieser induziert die Standardorientierung des Tangentialraumes.

Beispiel 7.3.6. Sei $M \subset \mathbb{R}^n$ eine glatte $(n-1)$ -dimensionale Untermannigfaltigkeit und $\nu : M \rightarrow S^{n-1}$ eine glatte Abbildung mit $T_x M = \nu(x)^\perp$ für alle $x \in M$. Dann bestimmt ν eine Orientierung von M : Eine Basis v_1, \dots, v_{n-1} von $T_x M$ genau dann positiv ist, wenn $\det(\nu(x), v_1, \dots, v_{n-1}) > 0$ ist. Beispiel 7.3.4 ist der Spezialfall $M = S^{n-1}$ mit $\nu(x) = x$.

Beispiel 7.3.7. Sei $f = (f_1, \dots, f_m) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine glatte Abbildung und sei 0 ein regulärer wert von f . Dann ist $M := f^{-1}(0)$ eine glatte orientierbare Untermannigfaltigkeit von \mathbb{R}^n der Dimension $d := n - m$. Für $x \in M$ kann eine Orientierung von $T_x M$ wie folgt definiert werden. Eine Basis v_1, \dots, v_d von $T_x M = \ker df(x)$ ist genau dann positiv, wenn

$$\det(\nabla f_1(x), \dots, \nabla f_m(x), v_1, \dots, v_d) > 0$$

ist. Beispiel 7.3.6 ist der Spezialfall $m = 1$ mit $\nu(x) = \|\nabla f(x)\|^{-1} \nabla f(x)$.

Beispiel 7.3.8. Das Moebiusband $M \subset \mathbb{R}^3$ in Abbildung 7.2 ist nicht orientierbar.

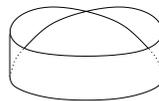


Abbildung 7.2: Das Möbiusband.

7.4 Integration

Sei $V \subset \mathbb{R}^d$ eine offene Teilmenge und $\tau \in \Omega^d(V)$ eine Differentialform vom Grade d . Die Koordinaten auf V bezeichnen wir mit $y = (y_1, \dots, y_d)$. Dann gibt es eine glatte (d.h. C^∞) Funktion $h : V \rightarrow \mathbb{R}$ so dass

$$\tau = h(y) dy_1 \wedge \dots \wedge dy_d.$$

Nehmen wir nun an τ habe **kompakten Träger** in V . Das heißt, die Menge

$$\text{supp}(\tau) := \overline{\{y \in V \mid h(y) \neq 0\}}$$

(der Abschluss ist hier in \mathbb{R}^d zu verstehen) ist kompakt und in V enthalten. Dazu äquivalent ist die Bedingung, dass der Abschluss der Menge $\{y \in V \mid h(y) \neq 0\}$ bezüglich der Relativtopologie von V kompakt ist. Wieder anders formuliert, es gibt eine kompakte Menge $K \subset V$ so dass h auf dem Komplement $V \setminus K$ verschwindet. Daraus folgt dass das Riemann-Integral von h auf V definiert ist: wir können ein Quadergebäude B konstruieren mit $K \subset B \subset V$ (Abbildung 7.3 und Übung 5.1.5) und h über B integrieren. Wir schreiben dies als Integral über V und definieren

$$\int_V \tau := \int_V h(y) dy_1 \cdots dy_d.$$

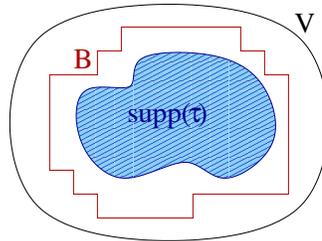


Abbildung 7.3: Differentialform mit kompaktem Träger.

Sei nun $V' \subset \mathbb{R}^d$ eine weitere offene Teilmenge und $\phi : V' \rightarrow V$ ein glatter Diffeomorphismus. Bezeichnen wir die Koordinaten auf V' mit $x = (x_1, \dots, x_d)$, so gilt

$$\phi^* \tau = h(\phi(x)) \det(d\phi(x)) dx_1 \wedge \dots \wedge dx_d$$

(siehe Gleichung (7.2.2) in Lemma 7.2.3). Hat ω kompakten Träger in V so hat auch die zurückgezogene d -Form $\phi^*\tau \in \Omega^d(V')$ kompakten Träger in V' . Ist die Determinante von $d\phi(x)$ überall positiv so folgt aus der Transformationsformel in Satz 5.7.11, dass

$$\int_{V'} \phi^*\tau = \int_V \tau. \quad (7.4.1)$$

Diese Gleichung gilt nur für orientierungserhaltende Diffeomorphismen ϕ . Bei einem orientierungsumkehrenden Diffeomorphismus ändert sich das Vorzeichen des Integrals. Diese Beobachtungen erlauben es uns, das Integral auf Kartengebiete in orientierten Untermannigfaltigkeiten zu übertragen.

Integration über Kartengebiete

Sei $M \subset \mathbb{R}^n$ eine d -dimensionale orientierte Untermannigfaltigkeit von \mathbb{R}^n , sei $W \subset \mathbb{R}^n$ eine offene Teilmenge die M enthält, und sei $\omega \in \Omega^d(W)$. Sei

$$U \subset M$$

ein Kartengebiet und sei $\psi : V \rightarrow U$ eine glatte orientierte Parametrisierung von U . Wir nehmen an dass der Träger

$$\text{supp}(\omega) := \overline{\{x \in \mathbb{R}^n \mid \omega_x \neq 0\}}$$

das Kartengebiet U in einer kompakten Menge schneidet. Dann hat $\psi^*\omega$ kompakten Träger in V und wir definieren das **Integral von ω über U** durch

$$\int_U \omega := \int_V \psi^*\omega.$$

Das Integral hängt nicht von der Wahl der Parametrisierung ab. Ist nämlich $\psi' : V' \rightarrow U$ eine weitere orientierte Parametrisierung von U , so ist der Diffeomorphismus $\phi := \psi^{-1} \circ \psi' : V' \rightarrow V$ orientierungserhaltend und es gilt

$$\int_{V'} (\psi')^*\omega = \int_{V'} (\psi \circ \phi)^*\omega = \int_{V'} \phi^*\psi^*\omega = \int_V \psi^*\omega.$$

Die letzte Gleichung folgt aus (7.4.1) mit $\tau = \psi^*\omega$.

Integral über kompakte Untermannigfaltigkeiten

Sei $M \subset \mathbb{R}^n$ eine kompakte orientierte d -dimensionale Untermannigfaltigkeit, enthalten in einer offenen Menge $W \subset \mathbb{R}^n$, und $\omega \in \Omega^d(W)$. Dann gibt es einen endlichen orientierten Atlas $\{U_\alpha, \phi_\alpha\}_{\alpha \in A}$. Wir wählen eine glatte Partition der Eins $\rho_\alpha : \mathbb{R}^n \rightarrow [0, 1]$ so dass

$$\text{supp}(\rho_\alpha) \cap M \subset U_\alpha, \quad \sum_{\alpha} \rho_\alpha(x) = 1 \quad \forall x \in M. \quad (7.4.2)$$

Daß eine solche Partition der Eins existiert folgt aus Satz 6.2.4. (Zunächst wählen wir offene Teilmengen $U'_\alpha \subset \mathbb{R}^n$ mit $U'_\alpha \cap M = U_\alpha$. Dann wenden wir Satz 6.2.4 auf die Überdeckung von M durch die Mengen U'_α an.)

Definition 7.4.1. Seien $M, \omega, U_\alpha, \rho_\alpha$ wie oben. Das **Integral von ω über M** ist definiert durch

$$\int_M \omega = \sum_{\alpha} \int_{U_\alpha} \rho_\alpha \omega. \quad (7.4.3)$$

Lemma 7.4.2. Das Integral von ω über M ist unabhängig von der Wahl des orientierten Atlas und der Partition der Eins.

Beweis. Ist $\{\tilde{U}_\beta, \tilde{\phi}_\beta\}_{\beta \in \tilde{A}}$ ein weiterer orientierter Atlas mit dazugehöriger Partition der Eins $\tilde{\rho}_\beta : \mathbb{R}^n \rightarrow [0, 1]$, so gilt

$$\sum_{\alpha} \int_{U_\alpha} \rho_\alpha \omega = \sum_{\alpha} \sum_{\beta} \int_{U_\alpha \cap \tilde{U}_\beta} \rho_\alpha \tilde{\rho}_\beta \omega = \sum_{\beta} \int_{\tilde{U}_\beta} \tilde{\rho}_\beta \omega.$$

Damit ist das Lemma bewiesen. \square

Bemerkung 7.4.3. Bezeichnen wir die Koordinaten auf \mathbb{R}^n mit x_1, \dots, x_n , und die Koordinaten auf \mathbb{R}^d mit y_1, \dots, y_d , und nehmen wir an, dass ω in der Form $\omega = \sum_{I \in \mathcal{I}_d(n)} a_I dx_I$ mit $a_I \in C^\infty(W)$ gegeben ist, so hat das Integral in Definition 7.4.1 die explizite Form

$$\int_M \omega = \sum_{\alpha} \sum_{I \in \mathcal{I}_d(n)} \int_{V_\alpha} \rho_\alpha(\psi_\alpha(y)) a_I(\psi_\alpha(y)) \det(d\psi_{\alpha I}(y)) dy_1 \cdots dy_d. \quad (7.4.4)$$

Hier bezeichnen wir mit $\psi_\alpha := \phi_\alpha^{-1} : V_\alpha \rightarrow U_\alpha$ die orientierte Parametrisierung von U_α und mit $\psi_{\alpha I}$ die Komposition von ψ_α mit der Projektion auf die durch $I \in \mathcal{I}_d(n)$ bezeichneten Koordinaten.

Integration über kompakte Teilmengen

Wie bisher sei $M \subset \mathbb{R}^n$ eine orientierte d -dimensionale Untermannigfaltigkeit, enthalten in einer offenen Menge $W \subset \mathbb{R}^n$, und $\omega \in \Omega^d(W)$. Wir verlangen jedoch nicht mehr dass M kompakt ist und integrieren ω statt dessen über eine geeignete kompakte Teilmenge $G \subset M$.

Für $G \subset M$ bezeichnen wir mit $\partial G := \partial_M G$ den Rand von G bezüglich der Relativtopologie auf M . Das heißt, ein Punkt $p \in M$ gehört zu ∂G wenn es sowohl eine Folge in G also auch eine Folge in $M \setminus G$ gibt, die gegen p konvergiert. Wir setzen voraus, dass ∂G eine d -dimensionale Jordan-Nullmenge ist. Das heißt, dass die Menge $K \cap \psi^{-1}(\partial G) = \{y \in K \mid \psi(y) \in \partial G\}$ für jede Parametrisierung $\psi : V \rightarrow U$ eines Kartengebiets und jede kompakte Teilmenge $K \subset V$ eine Jordansche Nullmenge ist (Definition 6.2.1). Das folgende Lemma erinnert noch einmal die Inklusion (6.2.2) im Beweis von Lemma 6.2.3.

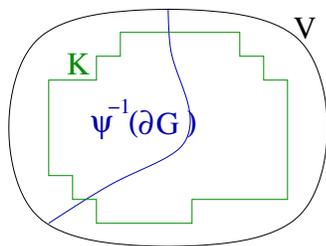


Abbildung 7.4: Die Menge $\psi^{-1}(\partial G)$.

Lemma 7.4.4. *Sei $G \subset M$ eine kompakte Teilmenge deren M -Rand ∂G eine d -dimensionale Nullmenge ist, $\psi : V \rightarrow U$ eine orientierte Parametrisierung eines Kartengebiets, und $K \subset V$ eine kompakte Menge (siehe Abbildung 7.4). Dann gilt*

$$\partial(K \cap \psi^{-1}(G)) \subset \partial K \cup (K \cap \psi^{-1}(\partial G)). \quad (7.4.5)$$

Beweis. Sei $y \in \partial(K \cap \psi^{-1}(G))$ und $y \notin \partial K$. Dann gibt es zwei Folgen $y_i \in K \cap \psi^{-1}(G)$ und $y'_i \in V \setminus (K \cap \psi^{-1}(G))$ die gegen y konvergieren. Da K kompakt und damit abgeschlossen ist, gilt $y \in K \subset V$, und da $\psi(y_i) \in G$ ist, gilt auch $\psi(y) = \lim_{i \rightarrow \infty} \psi(y_i) \in G$. Insbesondere ist y ein Element der offenen Menge $K \setminus \partial K$ und daher gilt $y'_i \in K$ für hinreichend grosse i . Daraus folgt $\psi(y'_i) \notin G$ für grosse i und deshalb ist $\psi(y)$ sowohl der Grenzwert einer Folge $\psi(y_i) \in G$ als auch einer Folge $\psi(y'_i) \in M \setminus G$. Also gilt $\psi(y) \in \partial G$ und somit $y \in K \cap \psi^{-1}(\partial G)$. Damit ist das Lemma bewiesen. \square

Wir können so vorgehen wie oben. Schneidet der Träger von $\omega \in \Omega^d(W)$ das Kartengebiet U in einer kompakten Menge und ist $\psi : V \rightarrow U$ eine glatte orientierte Parametrisierung, so hat die zurückgezogene d -Form $\psi^*\omega \in \Omega^d(V)$ kompakten Träger in V , und wir wählen wir ein Quadergebäude $B \subset V$ das den Träger von $\psi^*\omega$ enthält. Nach Lemma 7.4.4 ist der Rand der Menge $K \cap \psi^{-1}(G)$ eine Jordan-Nullmenge. Damit ist $K \cap \psi^{-1}(G)$ Jordan-messbar und $\psi^*\omega$ verschwindet ausserhalb dieser Menge. Wir definieren daher das Integral von ω über $G \cap U$ als das Integral der zurückgezogenen d -Form

$$\psi^*\omega =: h(y)dy_1 \wedge \cdots \wedge dy_d$$

über dem zurückgezogenen Gebiet $\psi^{-1}(G)$:

$$\int_{G \cap U} \omega := \int_{\psi^{-1}(G)} \psi^*\omega := \int_{K \cap \psi^{-1}(G)} h(y)dy_1 \cdots dy_d. \quad (7.4.6)$$

Wie bisher folgt aus der Transformationsformel in Satz 5.7.11, dass das Integral unabhängig von der Wahl der orientierten Parametrisierung ψ ist.

Ist $\omega \in \Omega^d(W)$ eine beliebige d -Form so wählen wir einen orientierten Atlas $\{U_\alpha, \phi_\alpha\}_{\alpha \in A}$ mit der Zusatzeigenschaft, dass U_α beschränkt ist und $\bar{U}_\alpha \subset M$ für jedes $\alpha \in A$ (wobei mit \bar{U}_α der Abschluss der Menge U_α im umgebenden Raum \mathbb{R}^n gemeint ist). Dann wählen wir eine Partition der Eins $\rho_\alpha : \mathbb{R}^n \rightarrow [0, 1]$, so dass $\text{supp}(\rho_\alpha)$ für jedes $\alpha \in A$ kompakt ist, ρ_α nur für endlich viele α von Null verschieden ist, und

$$\text{supp}(\rho_\alpha) \cap M \subset U_\alpha, \quad \sum_{\alpha} \rho_\alpha(x) = 1 \quad \forall x \in G.$$

Unter unseren Voraussetzungen ist $\text{supp}(\rho_\alpha) \cap M$ kompakt: Ist p_i eine Folge in $\text{supp}(\rho_\alpha) \cap M$ so konvergiert eine Teilfolge gegen ein $p \in \text{supp}(\rho_\alpha)$; da $p_i \in U_\alpha$ und $\bar{U}_\alpha \subset M$ ist folgt $p \in M$ und damit $p \in \text{supp}(\rho_\alpha) \cap M$. Zur Konstruktion der ρ_α gehen wir wie oben vor. Wir überdecken G zunächst durch endlich viele U_α , wählen dann für diese α beschränkte offene Mengen U'_α mit $U'_\alpha \cap M = U_\alpha$, und wenden schliesslich Satz 6.2.4 über die Existenz einer Partition der Eins auf die endliche Überdeckung der kompakten Menge G durch die U'_α an. Nun definieren wir das **Integral von ω über G** durch

$$\int_G \omega := \sum_{\alpha} \int_{G \cap U_\alpha} \rho_\alpha \omega. \quad (7.4.7)$$

Wie bisher ist dieses Integral unabhängig von der Wahl der U_α und ρ_α .

7.5 Der Satz von Stokes

In diesem Abschnitt formulieren und beweisen wir den Satz von Stokes für kompakten Teilmengen G einer glatten Untermannigfaltigkeit $M \subset \mathbb{R}^n$, die einen glatten Rand besitzen (vergleiche Definition 6.4.1).

Definition 7.5.1 (Glatter Rand). Sei $M \subset \mathbb{R}^n$ eine glatte d -dimensionale Untermannigfaltigkeit und sei G eine kompakte Teilmenge von M . Wir nennen G eine **kompakte Menge mit glattem Rand**, wenn ihr M -Rand ∂G eine glatte $(d-1)$ -dimensionale Untermannigfaltigkeit des \mathbb{R}^n ist, und wenn dieser M -Rand mit dem M -Rand des M -Inneren von G übereinstimmt (das heißt, wenn jede offene Menge $U \subset \mathbb{R}^n$ mit $U \cap \partial G \neq \emptyset$ sowohl Elemente aus $G \setminus \partial G$ als auch Elemente aus $M \setminus G$ enthält).

Definition 7.5.2 (Äußerer Einheitsnormalenvektor). Sei $M \subset \mathbb{R}^n$ eine glatte d -dimensionale Untermannigfaltigkeit, sei $G \subset M$ eine kompakte Teilmenge mit glattem Rand, und sei $p \in \partial G = \partial_M G$. Ein Vektor $v \in T_p M$ mit $\|v\| = 1$ heißt **äußerer Einheitsnormalenvektor von G an der Stelle p** wenn $v \perp T_p \partial G$ ist und eine glatte Kurve $\gamma : (-\delta, \delta) \rightarrow M$ existiert mit $\gamma(0) = p$, $\dot{\gamma}(0) = v$, und $\gamma(t) \notin G$ für $0 < t < \delta$.

Lemma 7.5.3. Sei $M \subset \mathbb{R}^n$ eine glatte orientierte d -dimensionale Untermannigfaltigkeit und sei $G \subset M$ eine kompakte Teilmenge mit glattem Rand. Dann gilt folgendes.

- (i) Ist $\phi : U \rightarrow V \subset \mathbb{R}^d$ eine Karte von M so ist $\phi(\partial G \cap U) \subset V$ eine glatte $(d-1)$ -dimensionale Untermannigfaltigkeit von \mathbb{R}^d .
- (ii) ∂G ist eine Jordansche Nullmenge bezüglich M .
- (iii) Für jedes Element $p \in \partial G$ existiert ein Kartengebiet $U \subset M$ mit $p \in U$, eine offene Teilmenge $V \subset \mathbb{R}^d$, und eine Karte $\phi : U \rightarrow V$ mit

$$\phi(U \cap G) = \{x \in V \mid x_d \leq 0\}. \quad (7.5.1)$$

(iv) Für jedes Element $p \in \partial G$ existiert genau ein äußerer Einheitsnormalenvektor $\nu(p) \in S^{n-1}$. Die dadurch definierte Abbildung $\nu : \partial G \rightarrow S^{n-1}$ heißt **äußeres Einheitsnormalenfeld von G** .

(v) Die Abbildung $\nu : \partial G \rightarrow S^{n-1}$ in (iv) läßt sich zu einer glatten Abbildung auf einer offenen Umgebung von ∂G mit Werten in S^{n-1} fortsetzen.

(vi) Es existiert eine Orientierung von ∂G , so dass eine Basis v_2, \dots, v_d von $T_p \partial G$ genau dann positiv ist, wenn $\nu(p), v_2, \dots, v_d$ eine positive Basis von $T_p M$ ist. Sie heißt **kanonische Rand-Orientierung von ∂G** .

Beweis. Wir beweisen (i). Sei $x_0 \in \phi(U \cap \partial G)$ und $p_0 := \phi^{-1}(x_0) \in U \cap \partial G$. Dann existiert ein Kartengebiet $U_0 \subset U \cap \partial G$ von ∂G mit $x_0 \in U_0$ und ein Diffeomorphismus $\phi_0 : U_0 \rightarrow V_0$ auf eine offene Teilmenge $V_0 \subset \mathbb{R}^{d-1}$. Insbesondere ist $U_0 \subset U \cap \partial G$ offen bezüglich der Relativtopologie von $U \cap \partial G$. Daher ist die Teilmenge $U'_0 := \phi(U_0) \subset \phi(U \cap \partial G)$ offen bezüglich der Relativtopologie von $\phi(U \cap \partial G) \subset \mathbb{R}^d$. Die Menge U'_0 enthält den Punkt $x_0 = \phi(p_0)$. Außerdem ist die Abbildung $\phi'_0 := \phi_0 \circ \phi^{-1} : U'_0 \rightarrow V_0 \subset \mathbb{R}^{d-1}$ ein Diffeomorphismus. Dies zeigt, dass Menge $\phi(U \cap \partial G)$ eine glatte $(d-1)$ -dimensionale Untermannigfaltigkeit von \mathbb{R}^d ist. Damit ist Teil (i) bewiesen.

Wir beweisen Teil (ii). Sei $U \subset M$ ein Kartengebiet, sei $V \subset \mathbb{R}^d$ eine offene Menge, und sei $\phi : U \rightarrow V$ eine Karte von M (Definition 6.1.9). Da ∂G eine M -abgeschlossene, Teilmenge von M ist, ist $U \cap \partial G$ eine abgeschlossene Teilmenge von U bezüglich der Relativtopologie von U (Lemma B.1.2). Da ϕ ein Homöomorphismus ist, folgt daraus, dass $\phi(U \cap \partial G)$ eine abgeschlossene Teilmenge von V bezüglich der Relativtopologie von V ist. Außerdem ist $\phi(U \cap \partial G)$ nach (i) eine glatte $(d-1)$ -dimensionale Untermannigfaltigkeit von \mathbb{R}^d . Daraus folgt, dass die Menge $\phi(U \cap \partial G) \cap K$, für jede kompakten Teilmenge $K \subset V$, eine kompakte Teilmenge einer $(d-1)$ -dimensionalen Untermannigfaltigkeit des \mathbb{R}^d , und damit nach Beispiel 5.6.13 eine Jordansche Nullmenge ist. Nach Definition 6.2.1 ist ∂G daher eine Jordansche Nullmenge bezüglich M . Damit ist Teil (ii) bewiesen.

Wir beweisen Teil (iii). Sei $p_0 \in \partial G$ gegeben. Wir zeigen in zwei Schritten, dass eine Karte $\phi : U \rightarrow V$ mit $p_0 \in U$ existiert, die (7.5.1) erfüllt.

Schritt 1. *Es existiert eine Karte $\phi : U \rightarrow V$ von M mit $p_0 \in U$, $V \subset \mathbb{R}^d$ offen und konvex, und $\phi(U \cap \partial G) = \{x \in V \mid x_d = 0\}$.*

Sei $U \subset M$ ein Kartengebiet mit $p_0 \in U$, sei $V \subset \mathbb{R}^d$ eine offene Menge, und sei $\phi : U \rightarrow V$ eine Karte von M . Nach (i) ist die Menge $\phi(U \cap \partial G)$ eine Untermannigfaltigkeit des \mathbb{R}^d und diese enthält den Punkt $x_0 := \phi(p_0)$. Daher existiert eine offene Teilmenge $V_0 \subset V$ mit $x_0 \in V_0$, eine offene Menge $V' \subset \mathbb{R}^d$ und ein Diffeomorphismus $\psi_0 : V_0 \rightarrow V'$ mit

$$\psi_0(V_0 \cap \phi(U \cap \partial G)) = \{x' \in V' \mid x'_d = 0\}.$$

Da $x_0 = \phi(p_0) \in V_0$ ist, ist die Menge $U' := \phi^{-1}(V_0) \subset U$ ein Kartengebiet in M mit $p_0 \in U'$ und die Abbildung $\phi' := \psi_0 \circ \phi : U' \rightarrow V'$ ist eine Karte von M mit $\phi'(U' \cap \partial G) = \psi_0(V_0 \cap \phi(U \cap \partial G)) = \{x' \in V' \mid x'_d = 0\}$. Durch Verkleinern von U' und V' , falls nötig, können wir annehmen, dass V' eine konvexe Teilmenge von \mathbb{R}^d ist. Damit ist Schritt 1 bewiesen.

Schritt 2. Es existiert ein Kartengebiet $U \subset M$ mit $p \in U$, eine offene Teilmenge $V \subset \mathbb{R}^n$, und eine Karte $\phi : U \rightarrow V$, die (7.5.1) erfüllt.

Sei $\phi : U \rightarrow V$ wie in Schritt 1. Weiterhin sei $\Omega := \overset{\circ}{G}$ das M -Innere von G . Dann gilt nach Voraussetzung $\partial\Omega = \partial G$ und daher erfüllt der Homöomorphismus $\phi : U \rightarrow V$ die Voraussetzungen von Lemma 6.4.4. Insbesondere gilt $U \cap \partial\Omega = U \cap \partial G = U^0$ nach Schritt 1 (in der Notation von Lemma 6.4.4). Sind $U^- \cap \Omega$ und $U^+ \cap \Omega$ beide leer, so gilt $U \cap \Omega = \emptyset$ und daher $p_0 \notin \partial\Omega$, im Widerspruch zu $\partial\Omega = \partial G$. Sind $U^- \cap \Omega$ und $U^+ \cap \Omega$ beide nichtleer, so ist $U \subset \bar{\Omega} = G$ nach Lemma 6.4.4 und daher $p_0 \in \overset{\circ}{G}$, im Widerspruch zu $p_0 \in \partial G$. Also ist genau eine der Mengen $U^\pm \cap \Omega$ nichtleer. Durch Vorzeichenwechsel von ϕ_d , falls nötig, können wir annehmen, dass $U^- \cap \Omega \neq \emptyset$ ist. Dann gilt $U^+ \cap G = U^+ \cap (\Omega \cup \partial\Omega) = \emptyset$, und $U^- \cup U^0 \subset \Omega \cup \partial\Omega = G$ nach Lemma 6.4.4. Damit sind Schritt 2 und Teil (iii) bewiesen.

Wir beweisen Teil (iv). Sei $p \in \partial G$ und sei $\phi : U \rightarrow V$ wie in Teil (iii). Dann gilt mit $\psi := \phi^{-1} : V \rightarrow U$ und $x := \phi(p)$

$$T_p M = \text{im} d\psi(x), \quad T_p \partial G = \left\{ \sum_{i=1}^{d-1} \frac{d\psi}{dx_i}(x) \xi_i \mid \xi_1, \dots, \xi_{d-1} \in \mathbb{R} \right\}.$$

Ist $\xi = (\xi_1, \dots, \xi_d) \in \mathbb{R}^d$ mit $\xi_d > 0$, so erfüllt $\gamma(t) := \psi(x + t\xi) \in M$ die Bedingung $\gamma(t) \notin G$ für kleine $t > 0$, da $U \cap G = \{\psi(x) \mid x \in V, x_d \leq 0\}$ ist nach (7.5.1). Ein solcher Vektor ξ mit $d\psi(x)\xi \perp T_p \partial G$ ist durch die Formel $\xi = (d\psi(x)^T d\psi(x))^{-1} e_d$ gegeben, wobei $e_d := (0, \dots, 0, 1) \in \mathbb{R}^d$ der standard Basisvektor ist. Daraus ergibt sich die Formel

$$\nu(p) = \frac{d\psi(x)(d\psi(x)^T d\psi(x))^{-1} e_d}{\|d\psi(x)(d\psi(x)^T d\psi(x))^{-1} e_d\|} \quad (7.5.2)$$

für den äußeren Einheitsnormalenvektor von G an der Stelle p . Dieser Vektor ist eindeutig, da der lineare Unterraum $T_p M \cap (T_p \partial G)^\perp$ eindimensional ist, und jede Kurve $\gamma : (-\delta, \delta) \rightarrow U$ mit $\gamma(0) = p$ und $\dot{\gamma}(0) = -\nu(p)$ die Bedingung $\gamma(t) \in G$ für kleine $t > 0$ erfüllt. Damit ist Teil (iv) bewiesen.

In Teil (v) folgt die lokale Fortsetzung des äußeren Einheitsnormalenfeldes $\nu : \partial G \rightarrow S^{n-1}$ zu einer glatten Funktion aus der Formel (7.5.2), indem man $x = \phi(p)$ setzt und die Tatsache verwendet, dass sich ϕ zu einer glatten Funktion von einer offenen Teilmenge des \mathbb{R}^n nach V fortsetzen läßt. Eine globale Fortsetzung von ν kann dann über eine Partition der Eins mittels Satz 6.2.4 konstruiert werden.

Teil (vi) folgt direkt aus Teil (v) und damit ist Lemma 7.5.3 bewiesen. \square

Teil (ii) von Lemma 7.5.3 zeigt, dass wir eine d -Form über G integrieren können, und Teil (vi) von Lemma 7.5.3 zeigt, dass wir $(d-1)$ -Form mit Hilfe der kanonischen Rand-Orientierung über ∂G integrieren können. Damit sind wir bereit für die Formulierung des **Satzes von Stokes**.

Satz 7.5.4 (Stokes). *Sei $M \subset \mathbb{R}^n$ eine glatte d -dimensionale orientierte Untermannigfaltigkeit, sei $G \subset M$ eine kompakte Teilmenge mit glattem Rand, und sei $\omega \in \Omega^{d-1}(W)$ eine $(d-1)$ -Form auf einer offenen Menge $W \subset \mathbb{R}^n$, die M enthält. Dann gilt*

$$\int_G d\omega = \int_{\partial G} \omega. \quad (7.5.3)$$

Beweis. Zunächst betrachten wir den Fall $d = n$. Dann ist $M = W \subset \mathbb{R}^d$ eine offene Teilmenge und $G \subset W$ eine kompakte Menge mit glattem Rand (Definition 6.4.1). Eine $(d-1)$ -Form auf W können wir dann in der Form

$$\omega = \sum_{i=1}^d (-1)^{j-1} f_j(x) dx_1 \wedge \cdots \wedge dx_{j-1} \wedge dx_{j+1} \wedge \cdots \wedge dx_d$$

schreiben, wobei

$$f = (f_1, \dots, f_d) : W \rightarrow \mathbb{R}^d$$

ein glattes Vektorfeld ist. Sei $\nu : \partial G \rightarrow \mathbb{R}^d$ das äußere Einheitsnormalenfeld. Wir werden die Gleichung

$$\int_{\partial G} \omega = \int_{\partial G} \langle f, \nu \rangle dS \quad (7.5.4)$$

beweisen. Mit

$$d\omega = \operatorname{div}(f) dx_1 \wedge \cdots \wedge dx_d$$

(Übung 7.2.11) folgt dann

$$\begin{aligned} \int_G d\omega &= \int_G \operatorname{div}(f) dx_1 \cdots dx_d \\ &= \int_{\partial G} \langle f, \nu \rangle dS \\ &= \int_{\partial G} \omega. \end{aligned}$$

Hier folgt die zweite Gleichung aus Satz 6.4.9.

Wir beweisen nun die Gleichung (7.5.4). Sei $V \subset \mathbb{R}^{d-1}$ offen und sei

$$\psi : V \rightarrow \partial G$$

eine glatte orientierte Parametrisierung eines Kartengebiets von ∂G . Wir zeigen zunächst, dass der äussere Einheitsnormalenvektor $\nu(\psi(y)) \in \mathbb{R}^d$ für alle $y \in V$ durch die Formel

$$\nu_j(\psi(y)) \sqrt{\det(d\psi(y)^T d\psi(y))} = (-1)^{j-1} \det(d\psi_{I_j}(y)) \quad (7.5.5)$$

für $j = 1, \dots, d$ gegeben ist, wobei $\psi_{I_j} : V \rightarrow \mathbb{R}^{d-1}$ die Abbildung bezeichnet, die sich aus ψ durch Fortlassen die j ten Koordinate ergibt, das heisst

$$\psi_{I_j}(y) := (\psi_1(y), \dots, \psi_{j-1}(y), \psi_{j+1}(y), \dots, \psi_d(y)).$$

Zum Beweis von (7.5.5) halten wir ein Element $y \in V$ fest und bezeichnen mit $\eta = (\eta_1, \dots, \eta_d) \in \mathbb{R}^d$ den durch die rechte Seite von (7.5.5) definierten Vektor, das heisst

$$\eta_j := (-1)^{j-1} \det(d\psi_{I_j}(y)), \quad j = 1, \dots, d.$$

Für $\xi = (\xi_1, \dots, \xi_d) \in \mathbb{R}^d$ betrachten wir die Matrix

$$A_\xi := \left(\xi, \frac{\partial \psi}{\partial y_1}(y), \dots, \frac{\partial \psi}{\partial y_{d-1}}(y) \right) \in \mathbb{R}^{d \times d}.$$

Entwickeln wir die Determinante von A_ξ nach der ersten Spalte so ergibt sich nach Teil (iv) von Satz D.3.5 die Gleichung

$$\det(A_\xi) = \sum_{j=1}^d (-1)^{j-1} \xi_j \det(d\psi_{I_j}(y)) = \langle \xi, \eta \rangle. \quad (7.5.6)$$

Mit $\xi \in \text{im } d\psi(y)$ verschwindet die linke Seite von (7.5.6). Also steht η senkrecht auf $\text{im } d\psi(y) = T_{\psi(y)}\partial G$. Für $\xi = \eta$ erhalten wir $\det(A_\eta) = \|\eta\|^2 > 0$. Also bilden die Vektoren

$$\eta, \frac{\partial \psi}{\partial y_1}, \dots, \frac{\partial \psi}{\partial y_{d-1}}$$

eine positive Basis von \mathbb{R}^d und, da die Vektoren $\partial \psi / \partial y_1, \dots, \partial \psi / \partial y_{d-1}$ nach Wahl von ψ eine positive Basis von $T_{\psi(y)}\partial G$ bilden, folgt dass der Vektor η nach aussen zeigt. Außerdem gilt

$$A_\eta^T A_\eta = \begin{pmatrix} \|\eta\|^2 & 0 \\ 0 & d\psi(y)^T d\psi(y) \end{pmatrix},$$

und daher $\|\eta\|^2 \det(d\psi(y)^T d\psi(y)) = \det(A_\eta^T A_\eta) = \det(A_\eta)^2 = \|\eta\|^4$. Also gilt $\|\eta\|^2 = \det(d\psi(y)^T d\psi(y))$ und daraus folgt (7.5.5).

Nach (7.5.5) erhalten wir im Fall $\text{supp}(f) \cap \partial G \subset \psi(V)$ die Gleichung

$$\begin{aligned} \int_{\partial G} \langle f, \nu \rangle dS &= \int_{\psi(V)} \langle f, \nu \rangle dS \\ &= \int_V \langle f(\psi(y)), \nu(\psi(y)) \rangle \sqrt{\det(d\psi(y)^T d\psi(y))} dy_1 \cdots dy_{d-1} \\ &= \int_V \sum_{j=1}^d (-1)^{j-1} f_j(\psi(y)) \det(d\psi_{I_j}(y)) dy_1 \cdots dy_{d-1} \\ &= \int_V \psi^* \omega = \int_{\psi(V)} \omega = \int_{\partial G} \omega. \end{aligned}$$

Hier folgt der zweite Schritt aus den Gleichungen (6.1.15) und (6.1.16) in Definition 6.1.10. Der dritte Schritt folgt dann aus (7.5.5) und der vierte Schritt aus (7.2.2). Damit haben wir (7.5.4) für $(d-1)$ -Formen mit kleinem Träger bewiesen. Der Fall einer beliebigen $(d-1)$ -Form ω auf W lässt sich auf den lokalen Fall mit Hilfe von Partitionen der Eins zurückführen. Damit ist der Satz von Stokes im Fall $d = n$ bewiesen.

Betrachten wir nun den Fall $d < n$. Wir wählen eine endliche Überdeckung von G durch Kartengebiete $U_\alpha \subset M$ so dass $\bar{U}_\alpha \subset M$ für alle α . Weiterhin wählen wir glatte orientierte Parametrisierungen $\psi_\alpha : V_\alpha \rightarrow U_\alpha$ auf offenen Teilmengen $V_\alpha \subset \mathbb{R}^d$ und eine Partition der Eins $\rho_\alpha : \mathbb{R}^n \rightarrow [0, 1]$ so dass so dass $\text{supp}(\rho_\alpha)$ für jedes $\alpha \in A$ kompakt ist und

$$\text{supp}(\rho_\alpha) \cap M \subset U_\alpha, \quad \sum_{\alpha} \rho_\alpha(x) = 1 \quad \forall x \in G.$$

Wir führen folgende Abkürzungen ein:

$$\omega_\alpha := \psi_\alpha^*(\rho_\alpha \omega) \in \Omega^{d-1}(V_\alpha), \quad G_\alpha := \psi_\alpha^{-1}(G), \quad \partial G_\alpha := \psi_\alpha^{-1}(\partial G).$$

Dann ist ∂G_α der Rand von G_α in der Relativtopologie von V_α . Da ω_α auf V_α kompakten Träger hat, schliessen wir aus dem Fall $d = n$, dass

$$\int_{G_\alpha} d\omega_\alpha = \int_{\partial G_\alpha} \omega_\alpha.$$

Dies ist nach Definition äquivalent zu der Gleichung

$$\int_{G \cap U_\alpha} d(\rho_\alpha \omega) = \int_{\partial G \cap U_\alpha} \rho_\alpha \omega.$$

Bilden wir nun die Summe über alle α so folgt hieraus Satz 7.5.4. \square

7.6 Beispiele und Anwendungen

Beispiel 7.6.1 (Flächeninhalt). Wir betrachten die 1-Form

$$\omega := \frac{1}{2}(xdy - ydx), \quad d\omega = dx \wedge dy.$$

Für ein Jordan-messbares Gebiet $G \subset \mathbb{R}^2$ ist das Integral von $d\omega$ über G gleich dem Jordan-mass, beziehungsweise dem Flächeninhalt, oder dem “2-dimensionalen Volumen”. Ist der Rand von G eine glatte 1-Mannigfaltigkeit, so gilt nach Stokes

$$\text{Vol}_2(G) = \int_G dx \wedge dy = \frac{1}{2} \int_{\partial G} (xdy - ydx).$$

Beispiel 7.6.2 (Umlaufzahl). Wir betrachten die 1-Form

$$\omega := \frac{1}{2\pi} \frac{xdy - ydx}{x^2 + y^2}$$

auf $\mathbb{R}^2 \setminus \{(0,0)\}$. Diese Form ist geschlossen (Übung). Sie ist jedoch nicht exakt. Andernfalls müsste ihr Integral über jeder geschlossenen Kurve

$$\Gamma \subset \mathbb{R}^2 \setminus \{0\}$$

nach dem Satz von Stokes verschwinden. Dies wäre der Fall $d = 1$ mit $\partial\Gamma = \emptyset$ und einer 1-Form df . Jedoch läßt sich leicht zeigen, dass das Integral von ω über S^1 (mit der Standardorientierung gegen den Uhrzeigersinn) gleich 1 ist. Allgemeiner, ist $\gamma : [0, 1] \rightarrow \mathbb{R}^2 \setminus \{(0,0)\}$ eine glatte Kurve mit

$$\gamma(0) = \gamma(1)$$

und schreiben wir $\gamma(t) = (x(t), y(t))$ so ist

$$\iota(\gamma) := \int_{[0,1]} \gamma^* \omega = \frac{1}{2\pi} \int_0^1 \frac{x(t)\dot{y}(t) - y(t)\dot{x}(t)}{x(t)^2 + y(t)^2} dt$$

eine ganze Zahl (Übung). Man kann $\iota(\gamma)$ als Umlaufzahl um den Nullpunkt interpretieren.

Beispiel 7.6.3 (Der klassische Satz von Stokes). Sei $M \subset \mathbb{R}^3$ eine glatte 2-dimensionale orientierte Untermannigfaltigkeit und $G \subset M$ eine kompakte Teilmenge mit glattem Rand. Dann ist der Rand ∂G (wie gehabt verstanden bezüglich der Relativtopologie von M) eine kompakte 1-dimensionale Untermannigfaltigkeit von \mathbb{R}^3 . Sei weiterhin $U \subset \mathbb{R}^3$ eine offene Teilmenge, die M enthält, und sei

$$f = (a, b, c) : U \rightarrow \mathbb{R}^3$$

ein glattes Vektorfeld auf U . Wir betrachten die 1-Form

$$\alpha_f := a dx + b dy + c dz \in \Omega^1(U).$$

Wie wir in Gleichung (7.2.8) in Beispiel 7.2.9 gesehen haben, ist

$$d\alpha_f = \omega_{\text{rot}(f)} = u dy \wedge dz + v dz \wedge dx + w dx \wedge dy, \quad (7.6.1)$$

wobei $u := \partial c / \partial y - \partial b / \partial z$, $v := \partial a / \partial z - \partial c / \partial x$, $w := \partial b / \partial x - \partial a / \partial y$, die Koordinaten des Vektorfeldes $\text{rot}(f) = (u, v, w) : U \rightarrow \mathbb{R}^3$ sind. Nach Satz 7.5.4 und (7.6.1) gilt

$$\int_G \omega_{\text{rot}(f)} = \int_{\partial G} \alpha_f. \quad (7.6.2)$$

Diese Gleichung läßt sich wie folgt umformulieren. Sei $\nu : M \rightarrow S^2$ das **positive Einheitsnormalenfeld**, so dass $\nu(p) \perp T_p M$ ist und zwei Vektoren $v, w \in T_p M$ genau dann eine positive Basis des Tangentialraumes bilden, wenn $\det(\nu(p), v, w) > 0$ ist. Weiterhin sei $\tau : \partial G \rightarrow S^2$ das **positive Einheitstangentenfeld**, so dass $\tau(p) \in T_p \partial G$ für jedes $p \in \partial G$ der positive Einheitstangentenvektor ist. Da der Buchstabe ν anderweitig verwendet wurde, bezeichnen wir in diesem Beispiel das **äußere Einheitsnormalenfeld** von G in Lemma 7.5.3 mit

$$\eta : \partial G \rightarrow S^2.$$

Für $p \in T_p \partial G$ ist $\eta(p) \in T_p M \cap (T_p \partial G)^\perp$ also der eindeutige Einheitsvektor orthogonal zu $\nu(p)$ und $\tau(p)$, der nach der in Lemma 7.5.3 eingeführten Konventionen für die kanonische Rand-Orientierung von ∂G die Gleichung

$$\det(\nu(p), \eta(p), \tau(p)) = 1. \quad (7.6.3)$$

erfüllt. Umgekehrt ist es die Gleichung (7.6.3), die über das vorgegebene äußere Einheitsnormalenfeld η von G die Beziehung herstellt zwischen der Orientierung von G via ν und der von ∂G via τ . (Siehe Abbildung 7.5.)

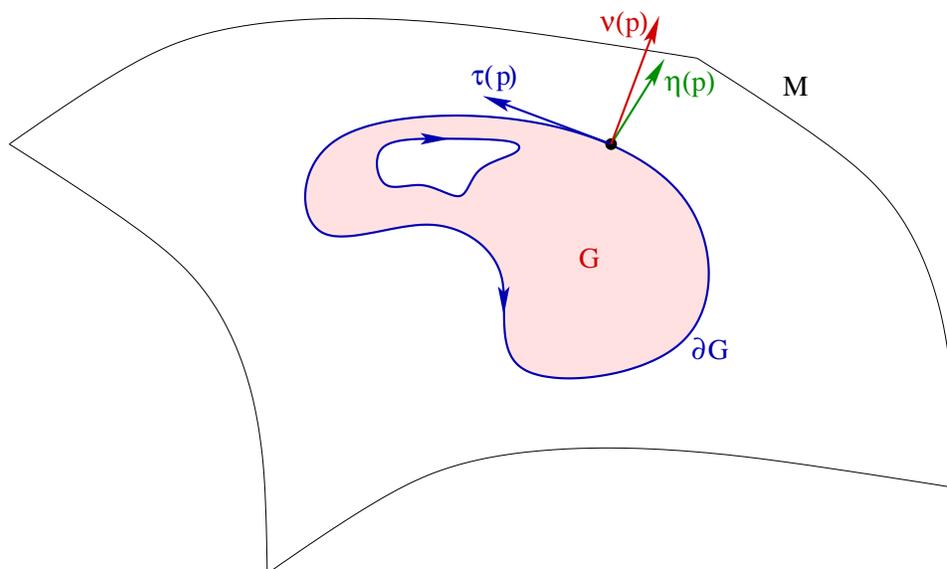


Abbildung 7.5: Der klassische Satz von Stokes.

Nun folgt aus der Gleichung (7.6.1) zusammen mit der Gleichung (7.5.4) im Beweis von Satz 7.5.4 (mit M statt ∂G und $\text{rot}(f)$ statt f) die Formel

$$\int_G \omega_{\text{rot}(f)} = \int_G \langle \text{rot}(f), \nu \rangle dS. \quad (7.6.4)$$

Wählen wir eine Zusammenhangskomponente $\Gamma \subset \partial G$ und eine positive Parametrisierung $\gamma : \mathbb{R} \rightarrow \Gamma$ nach der Bogenlänge, so gilt $\gamma(t+T) = \gamma(t)$ für alle $t \in \mathbb{R}$ und ein geeignetes $T > 0$ (die Länge von Γ), und $\dot{\gamma}(t) = \tau(\gamma(t))$ für alle $t \in \mathbb{R}$. Insbesondere ist $\|\dot{\gamma}(t)\| = 1$ für alle $t \in \mathbb{R}$ und daraus folgt

$$\int_\Gamma \alpha_f = \int_0^T \alpha_f(\gamma(t); \dot{\gamma}(t)) dt = \int_0^T \langle f(\gamma(t)), \tau(\gamma(t)) \rangle dt = \int_\Gamma \langle f, \tau \rangle dS.$$

Bilden wir nun die Summe über alle Zusammenhangskomponenten von ∂G so erhalten wir die Formel

$$\int_{\partial G} \alpha_f = \int_{\partial G} \langle f, \tau \rangle dS. \quad (7.6.5)$$

Mit den Gleichungen (7.6.4) und (7.6.5) wird aus (7.6.2) die Formel

$$\int_G \langle \text{rot}(f), \nu \rangle dS = \int_{\partial G} \langle f, \tau \rangle dS. \quad (7.6.6)$$

Dies ist der Satz von Stokes in der klassischen Form.

Beispiel 7.6.4. Wir betrachten die 2-Form

$$\omega := \frac{xdy \wedge dz + ydz \wedge dx + zdx \wedge dy}{(x^2 + y^2 + z^2)^{3/2}}$$

auf $\mathbb{R}^3 \setminus \{(0, 0, 0)\}$. Diese 2-Form ist geschlossen aber nicht exakt. Ihr Integral über der Einheitssphäre $S^2 \subset \mathbb{R}^3$ ist

$$\int_{S^2} \omega = \text{Vol}_2(S^2) = 4\pi.$$

(Übung: Diese Aussage mit dem Satz von Stokes beweisen.)

Der folgende Retraktionssatz ist eine interessante Anwendung des Satzes von Stokes. Da diese Anwendung den Fall $d = n$ betrifft, könnte man den Retraktionssatz natürlich auch direkt mit Hilfe des Divergenzsatzes von Gauß (Satz 6.4.9) beweisen. Jedoch ist der Beweis mit dem Formalismus der Differentialformen erheblich einfacher. Will man ihn direkt mit Hilfe der Divergenz eines geeigneten Vektorfeldes beweisen, erhält man recht komplizierte Formeln, die darauf hinauslaufen, die Argumente im Beweis des Satzes von Stokes (Satz 7.5.4) für den Fall $d = n$ nochmals durchzugehen.

Satz 7.6.5 (Retraktionssatz). *Sei $\Omega \subset \mathbb{R}^n$ eine beschränkte offene Teilmenge mit glattem Rand, und sei $U \subset \mathbb{R}^n$ eine offene Menge mit $\overline{\Omega} \subset U$. Dann existiert keine glatte Abbildung $\phi : U \rightarrow \mathbb{R}^n$ mit*

$$\phi(\overline{\Omega}) \subset \partial\Omega$$

und $\phi(x) = x$ für alle $x \in \partial\Omega$.

Beweis. Wir nehmen an, dass eine solche glatte Abbildung $\phi : U \rightarrow \mathbb{R}^n$ existiert. Dann gilt $\phi(x + t\xi) \in \partial\Omega$ für jedes $x \in \Omega$, jeden Vektor $\xi \in \mathbb{R}^n$, und jedes hinreichend kleine $t \in \mathbb{R}$. Daraus folgt

$$d\phi(x)\xi \in T_{\phi(x)}\partial\Omega$$

für alle $x \in \Omega$ und alle $\xi \in \mathbb{R}^n$. Also sind die Spalten der Jacobi-Matrix $d\phi(x)$ linear abhängig für alle $x \in \Omega$, und daher wegen der Stetigkeit der ersten Ableitungen von ϕ auch für alle $x \in \overline{\Omega}$. Daraus folgt die Gleichung

$$\det(d\phi(x)) = 0 \quad \text{für alle } x \in \overline{\Omega}. \quad (7.6.7)$$

Wir betrachten nun die $(n - 1)$ -Form

$$\omega := x_1 dx_2 \wedge \cdots \wedge dx_n.$$

Diese hat als Differential

$$d\omega = dx_1 \wedge dx_2 \wedge \cdots \wedge dx_n$$

die standard Volumenform auf dem \mathbb{R}^n . Daher folgt aus (7.6.7) und Teil (iv) von Lemma 7.1.12 mit $k = n$ die Formel

$$\phi^* d\omega = \det(d\phi) d\omega = 0.$$

Daher gilt, nach Satz 7.2.7 und Satz 7.5.4,

$$\begin{aligned} 0 &= \int_G \phi^* d\omega \\ &= \int_G d\phi^* \omega \\ &= \int_{\partial G} \phi^* \omega \\ &= \int_{\partial G} \omega \\ &= \int_G d\omega \\ &= \text{Vol}_n(G) \\ &> 0 \end{aligned}$$

Hier folgt die vierte Gleichheit aus der Annahme, dass die Abbildung

$$\phi|_{\partial G} : \partial G \rightarrow \partial G$$

die Identität ist. Wir haben also einen Widerspruch zu dieser Annahme erzeugt, welcher zeigt, dass eine glatte Abbildung $\phi : U \rightarrow \partial\Omega$ mit $\phi(x) = x$ für alle $x \in \partial\Omega$ nicht existieren kann. Damit ist Satz 7.6.5 bewiesen. \square

Der Retraktionsatz hat verschiedene interessante Konsequenzen. Die erste betrifft Fixpunkte einer glatten Abbildung $f : B \rightarrow B$, wobei

$$B := \{x \in \mathbb{R}^n \mid \|x\| \leq 1\}$$

den abgeschlossenen Einheitsball im \mathbb{R}^n mit der Euklidischen Norm bezeichnet.

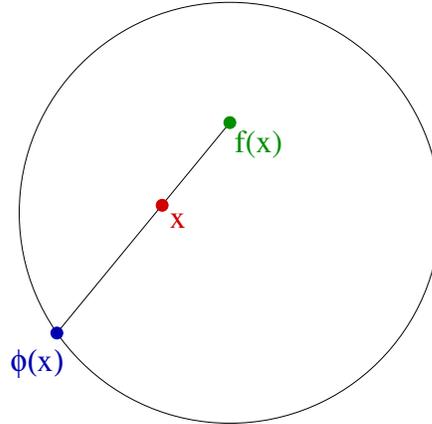


Abbildung 7.6: Der Brouwersche Fixpunktsatz.

Korollar 7.6.6. *Jede glatte Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit $f(B) \subset B$ besitzt einen Fixpunkt $x = f(x) \in B$.*

Beweis. Wir bezeichnen den Rand von B mit

$$S := \partial B = \{x \in \mathbb{R}^n \mid \|x\| = 1\} = S^{n-1}.$$

Nun nehmen wir an, es gibt eine glatte Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, die den Ball B auf sich selbst abbildet aber keinen Fixpunkt in B besitzt. Dann ist

$$U := \left\{ x \in \mathbb{R}^n \mid f(x) \neq x, \|x\|^2 < 1 + \frac{\langle x, x - f(x) \rangle^2}{\|x - f(x)\|^2} \right\}$$

eine offene Teilmenge von \mathbb{R}^n , die B enthält, und es existiert eine glatte Abbildung $\phi : U \rightarrow S$, die die Bedingung $\phi(x) = x$ für alle $x \in S$ erfüllt. Eine explizite Formel für eine solche Abbildung ist, für $x \in U$,

$$\begin{aligned} \phi(x) &:= x - t(x)(x - f(x)), \\ t(x) &:= \frac{1}{\|x - f(x)\|} \left(\sqrt{1 - \|x\|^2 + \frac{\langle x, x - f(x) \rangle^2}{\|x - f(x)\|^2}} - \frac{\langle x, x - f(x) \rangle}{\|x - f(x)\|} \right). \end{aligned}$$

Geometrisch erhält man diese Formel, indem man der Geraden von $f(x)$ in Richtung x folgt, solange bis man auf die Sphäre S trifft. Dieser Punkt auf der Sphäre ist der Bildpunkt $\phi(x)$ (siehe Abbildung 7.6). Da aber eine solche Abbildung $\phi : U \rightarrow S$ nach Satz 7.6.5 nicht existiert, muß also f entgegen unserer Annahme doch einen Fixpunkt in B haben, und damit ist Korollar 7.6.6 bewiesen. \square

Korollar 7.6.7 (Brouwerscher Fixpunktsatz).

Jede stetige Abbildung $f : B \rightarrow B$ besitzt einen Fixpunkt $x = f(x) \in B$.

Beweis. Der Beweis erfordert den Approximationssatz von Weierstraß, welcher besagt, dass für jede stetige Funktion $f : B \rightarrow \mathbb{R}$ eine Folge von Polynomen $p_i : \mathbb{R}^n \rightarrow \mathbb{R}$ (in n Variablen) existiert, die auf B gleichmäßig gegen f konvergiert. Dieses Resultat gilt natürlich auch für Abbildungen mit Werten in \mathbb{R}^n .

Nehmen wir nun an, dass f keinen Fixpunkt hat. Dann gilt

$$\varepsilon := \inf_{x \in B} \|x - f(x)\| > 0.$$

Nach dem Approximationssatz von Weierstraß existiert ein Polynom

$$p : \mathbb{R}^n \rightarrow \mathbb{R}^n,$$

das die Bedingung

$$\|p(x) - f(x)\| \leq \frac{\varepsilon}{2} \quad \text{für alle } x \in B$$

erfüllt. Nun definieren wir die Abbildung $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ durch

$$g(x) := \frac{p(x)}{1 + \varepsilon/2} \quad \text{für } x \in \mathbb{R}^n.$$

Diese Abbildung ist glatt und bildet den Einheitsball B auf sich selbst ab, da für jeden Punkt $x \in B$ die Ungleichung

$$\|p(x)\| \leq \|f(x)\| + \|p(x) - f(x)\| \leq 1 + \frac{\varepsilon}{2}$$

erfüllt ist. Für jede Punkt $x \in B$ gilt ebenfalls die Ungleichung

$$\begin{aligned} \|g(x) - x\| &= \left\| f(x) - x + \frac{p(x) - f(x)}{1 + \varepsilon/2} - \frac{\varepsilon/2}{1 + \varepsilon/2} f(x) \right\| \\ &\geq \|f(x) - x\| - \frac{\|p(x) - f(x)\| + (\varepsilon/2) \|f(x)\|}{1 + \varepsilon/2} \\ &\geq \varepsilon - \frac{\varepsilon/2 + \varepsilon/2}{1 + \varepsilon/2} \\ &> 0. \end{aligned}$$

Also besitzt g keinen Fixpunkt in B , im Widerspruch zu Korollar 7.6.6. Damit ist Korollar 7.6.7 bewiesen. \square

Korollar 7.6.7 ist ein Spezialfall des **Brouwerschen Fixpunktsatzes**. Dieser wurde von Luitzen Egbertus Jan Brouwer im Jahre 1910 bewiesen und besagt, dass jede stetige Abbildung von einer kompakten konvexen Teilmenge des \mathbb{R}^n auf sich selbst einen Fixpunkt besitzt. Der Brouwersche Fixpunktsatz wurde 1930 von Juliusz Schauder auf kompakte konvexe Teilmengen eines Banachraumes erweitert. Ein Satz, der unabhängig von Oskar Perron in 1907 und Georg Frobenius in 1912 bewiesen wurde, besagt, dass jede quadratische Matrix mit positiven Einträgen einen Eigenvektor mit nichtnegativen Einträgen besitzt. Er kann aus Korollar 7.6.7 hergeleitet werden.

Korollar 7.6.8 (Perron–Frobenius).

Sei $A = (a_{ij})_{i,j=1}^n \in \mathbb{R}^{n \times n}$ eine Matrix mit $a_{ij} > 0$ für alle $i, j = 1, \dots, n$. Dann existiert ein Vektor $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ und eine reelle Zahl λ mit

$$x_i \geq 0 \text{ für } i = 1, \dots, n, \quad \sum_{i=1}^n x_i = 1, \quad \lambda > 0, \quad Ax = \lambda x \quad (7.6.8)$$

Beweis. Für $\xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$ sei $\|\xi\|_1 := \sum_i |\xi_i|$ und $\|\xi\|_2 := \sqrt{\sum_i \xi_i^2}$. Der standard Simplex ist die Menge

$$\Delta := \left\{ x \in \mathbb{R}^n \mid x_i \geq 0 \text{ für } i = 1, \dots, n, \sum_{i=1}^n x_i = 1 \right\}$$

und wir definieren die stetige Abbildung $f : \Delta \rightarrow \Delta$ durch

$$f(x) := \frac{Ax}{\|Ax\|_1} \quad \text{für } x \in \Delta.$$

Der Simplex Δ ist homöomorph zum Ball $B := \{y \in \mathbb{R}^{n-1} \mid \|y\|_2 \leq 1\}$. Ein expliziter Homöomorphismus $\phi = (\phi_1, \dots, \phi_n) : B \rightarrow \Delta$ ist durch

$$\begin{aligned} \phi_i(y) &:= \frac{1}{n} \left(1 - \frac{y_i \|y\|_2}{\max\{y_1, \dots, y_{n-1}, -\sum_{j=1}^{n-1} y_j\}} \right), \quad i = 1, \dots, n-1, \\ \phi_n(y) &:= \frac{1}{n} \left(1 + \frac{\sum_{j=1}^{n-1} y_j \|y\|_2}{\max\{y_1, \dots, y_{n-1}, -\sum_{j=1}^{n-1} y_j\}} \right) \end{aligned} \quad (7.6.9)$$

für $y \in B$ gegeben. Nun ist die Abbildung $g := \phi^{-1} \circ f \circ \phi : B \rightarrow B$ stetig und besitzt daher nach Korollar 7.6.7 einen Fixpunkt $y = g(y) \in B$. Daraus folgt, dass $x := \phi(y) \in \Delta$ ein Fixpunkt von f ist. Dies ist äquivalent zu der Gleichung $Ax = \|Ax\|_1 x$, und daher erfüllt x die Behauptung von Korollar 7.6.8 mit $\lambda := \|Ax\|_1 > 0$. \square

Anhang A

Abschluss und Inneres

A.1 Topologische Grundbegriffe

Sei (X, d) ein **metrischer Raum**, d.h. X ist eine Menge und $d : X \times X \rightarrow \mathbb{R}$ ist eine Funktion mit folgenden Eigenschaften

(d1) Für alle $x, y \in X$ gilt $d(x, y) \geq 0$ und $d(x, y) = 0 \iff x = y$.

(d2) Für alle $x, y \in X$ gilt $d(x, y) = d(y, x)$.

(d3) Für alle $x, y, z \in X$ gilt $d(x, z) \leq d(x, y) + d(y, z)$.

Die Ungleichung in (d3) heißt **Dreiecksungleichung**. Die Funktion d heißt **Abstandsfunktion**. Eine Teilmenge $U \subset X$ heißt **offen** wenn es für jedes Element $x \in U$ ein $\varepsilon > 0$ gibt so dass

$$B_\varepsilon(x) := \{y \in X \mid d(x, y) < \varepsilon\} \subset U.$$

Insbesondere ist die Menge $B_r(x)$ für jedes $x \in X$ und jedes $r > 0$ offen; sie wird der **offene Ball in X mit Mittelpunkt x und Radius r** genannt. Offene Mengen haben folgende Eigenschaften.

(O1) X und \emptyset sind offene Teilmengen von X .

(O2) Sind U_1, \dots, U_n offene Teilmengen von X so ist auch $\bigcap_{i=1}^n U_i$ offen.

(O3) Ist I eine Menge und $U_i \subset X$ eine offene Teilmenge für jedes $i \in I$ so ist auch $\bigcup_{i \in I} U_i$ offen.

Eine Folge $(x_n)_{n \in \mathbb{N}}$ in X **konvergiert** gegen $x \in X$ wenn für jedes $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert, so dass für alle $n \in \mathbb{N}$ folgendes gilt

$$n \geq n_0 \quad \implies \quad d(x_n, x) < \varepsilon.$$

Eine Teilmenge $A \subset X$ heißt **abgeschlossen** wenn für jede Folge $(x_n)_{n \in \mathbb{N}}$ in A , die gegen ein Element $x \in X$ konvergiert, ihr Grenzwert x ebenfalls ein Element von A ist.

Eine Teilmenge $A \subset X$ ist genau dann abgeschlossen wenn ihr Komplement $U := X \setminus A$ offen ist (siehe [6]). Abgeschlossene Mengen haben folgende Eigenschaften.

- (A1) Die leere Menge und der ganze Raum X sind abgeschlossene Teilmengen von X .
- (A2) Sind A_1, \dots, A_n abgeschlossene Teilmengen von X so ist auch $\bigcup_{i=1}^n A_i$ abgeschlossen.
- (A3) Ist I eine Menge und $A_i \subset X$ eine abgeschlossene Teilmenge für jedes $i \in I$ so ist auch $\bigcap_{i \in I} A_i$ abgeschlossen.

Seien (X, d_X) und (Y, d_Y) zwei metrische Räume. Eine Abbildung $f : X \rightarrow Y$ heißt **stetig** wenn für jedes $x_0 \in X$ und jedes $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass für alle $x \in X$ folgendes gilt

$$d_X(x, x_0) < \delta \quad \implies \quad d_Y(f(x), f(x_0)) < \varepsilon.$$

Für jede Abbildung $f : X \rightarrow Y$ sind folgende Aussagen äquivalent.

- (S1) f ist stetig.
- (S2) Ist $(x_n)_{n \in \mathbb{N}}$ eine Folge in X die gegen $x \in X$ konvergiert, so konvergiert auch $f(x_n)$ gegen $f(x) \in Y$.
- (S3) Ist V eine offene Teilmenge von Y so ist $f^{-1}(V) := \{x \in X \mid f(x) \in V\}$ eine offene Teilmenge von X .

(Siehe Analysis I [6].) Eine Abbildung $f : X \rightarrow Y$ heißt **Homöomorphismus** wenn sie bijektiv ist und f und f^{-1} stetig sind.

A.2 Die Begriffe von Abschluss und Innerem

Dieser Abschnitt behandelt die Begriffe des *Abschlusses* und des *Inneren* einer Teilmenge eines metrischen Raumes. Sei also (X, d) ein metrischer Raum. Das Standardbeispiel ist ein reeller normierter Vektorraum $(X, \|\cdot\|)$ mit der durch $d(x, y) := \|x - y\|$ für $x, y \in X$ definierten Abstandsfunktion. Jedoch darf (X, d) in diesem Abschnitt ein beliebiger metrischer Raum sein. Für $x \in X$ und $r > 0$ bezeichnen wir, wie üblich, den offenen Ball vom Radius r mit Mittelpunkt x mit $B_r(x) := \{y \in X \mid d(x, y) < r\}$.

Definition A.2.1. Gegeben sei eine Teilmenge $A \subset X$. Der **Abschluss von A** ist die Menge

$$\bar{A} := \{x \in X \mid \forall \varepsilon > 0 \text{ gilt } B_\varepsilon(x) \cap A \neq \emptyset\}.$$

Das **Innere von A** ist die Menge

$$\mathring{A} := \{x \in X \mid \exists \varepsilon > 0 \text{ so dass } B_\varepsilon(x) \subset A\}.$$

Der **Rand von A** ist die Menge

$$\partial A := \bar{A} \setminus \mathring{A} = \left\{ x \in X \mid \begin{array}{l} \forall \varepsilon > 0 \text{ gilt } B_\varepsilon(x) \cap A \neq \emptyset \\ \text{und } B_\varepsilon(x) \cap (X \setminus A) \neq \emptyset \end{array} \right\}.$$

Beispiel A.2.2. Sei $X = \mathbb{R}$ die Menge der reellen Zahlen mit der üblichen Abstandsfunktion $d(x, y) = |x - y|$. Seien $a < b$ zwei reelle Zahlen und sei A das halboffene Intervall

$$A := [a, b) := \{x \in \mathbb{R} \mid a \leq x < b\}.$$

Dann sind der Abschluss, das Innere, und der Rand von A durch

$$\bar{A} = [a, b], \quad \mathring{A} = (a, b), \quad \partial A = \{a, b\}$$

gegeben.

In jedem metrischen Raum (X, d) ist der Abschluss einer Teilmenge A stets eine abgeschlossene Menge, und das Innere von A ist stets eine offene Menge. Damit ist auch der Rand von A eine abgeschlossene Menge. Die folgende Proposition fasst die wichtigsten Eigenschaften des Abschlusses und des Inneren zusammen.

Proposition A.2.3. Sei (X, d) ein metrischer Raum und $A \subset X$.

(i) Sei $x \in X$. Dann ist $x \in \bar{A}$ genau dann, wenn eine Folge $x_n \in A$ existiert, die gegen x konvergiert.

(ii) $X \setminus \overset{\circ}{A} = \overline{X \setminus A}$, und $\partial A = \bar{A} \cap \overline{X \setminus A}$.

(iii) $\bar{A} = A \cup \partial A$ und $\overset{\circ}{A} = A \setminus \partial A$.

(iv) \bar{A} ist die kleinste abgeschlossene Teilmenge von X , die A enthält.

(v) $\overset{\circ}{A}$ ist die grösste offene Teilmenge von X , die in A enthalten ist.

(vi) A ist abgeschlossen genau dann, wenn $\bar{A} = A$ (bzw. $\partial A \subset A$) ist.

(vii) A ist offen genau dann, wenn $\overset{\circ}{A} = A$ (bzw. $\partial A \cap A = \emptyset$) ist.

Beweis. Wir beweisen (i). Sei $x_n \in A$ eine Folge die gegen ein Element $x \in X$ konvergiert. Sei $\varepsilon > 0$. Nach Definition der Konvergenz existiert eine natürliche Zahl $n_0 \in \mathbb{N}$ so dass, für alle $n \in \mathbb{N}$ gilt

$$n \geq n_0 \quad \implies \quad d(x, x_n) < \varepsilon.$$

Damit gilt $x_n \in B_\varepsilon(x)$ für jedes $n \geq n_0$ und daher ist $B_\varepsilon(x) \cap A \neq \emptyset$. Da dies für jedes $\varepsilon > 0$ gilt, folgt daraus $x \in \bar{A}$. Umgekehrt sei $x \in \bar{A}$ gegeben. Dann ist $B_{1/n}(x) \cap A \neq \emptyset$ für jede natürliche Zahl $n \in \mathbb{N}$. Nach dem abzählbaren Auswahlaxiom existiert also eine Folge $x_n \in A$ so dass für jedes $n \in \mathbb{N}$ die Ungleichung $d(x, x_n) < 1/n$ erfüllt ist. Daraus folgt $\lim_{n \rightarrow \infty} d(x, x_n) = 0$ und daher $x = \lim_{n \rightarrow \infty} x_n$. Damit ist (i) bewiesen.

Wir beweisen (ii). Nach Definition von Abschluss und Innerem gilt

$$\begin{aligned} \overline{X \setminus A} &= \{x \in X \mid \forall \varepsilon > 0 \text{ gilt } B_\varepsilon(x) \cap (X \setminus A) \neq \emptyset\} \\ &= \{x \in X \mid \forall \varepsilon > 0 \text{ gilt } B_\varepsilon(x) \not\subset A\} \\ &= X \setminus \overset{\circ}{A}. \end{aligned}$$

Daraus folgt

$$\partial A = \bar{A} \setminus \overset{\circ}{A} = \bar{A} \cap (X \setminus \overset{\circ}{A}) = \bar{A} \cap \overline{X \setminus A}.$$

Damit ist (ii) bewiesen.

Wir beweisen (iii). Es gilt $\overset{\circ}{A} \subset A \subset \bar{A}$. Nach Definition des Randes folgt daraus

$$A \cup \partial A = A \cup (\bar{A} \setminus \overset{\circ}{A}) = A \cup \bar{A} = \bar{A}$$

und

$$A \setminus \partial A = A \setminus (\bar{A} \setminus \overset{\circ}{A}) = A \setminus (A \setminus \overset{\circ}{A}) = \overset{\circ}{A}.$$

Damit ist (iii) bewiesen.

Wir beweisen (iv). Zunächst zeigen wir, dass \bar{A} abgeschlossen ist. Sei also $x_n \in \bar{A}$ eine Folge die gegen ein Element $x \in X$ konvergiert. Sei eine Zahl $\varepsilon > 0$ gegeben. Nach Definition der Konvergenz existiert ein $n \in \mathbb{N}$ mit

$$d(x, x_n) < \frac{\varepsilon}{2}.$$

Da $x_n \in \bar{A}$ ist, gilt $B_{\varepsilon/2}(x_n) \cap A \neq \emptyset$. Wähle ein Element $a \in B_{\varepsilon/2}(x_n) \cap A$. Dann gilt nach der Dreiecksungleichung

$$d(x, a) \leq d(x, x_n) + d(x_n, a) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Daraus folgt

$$B_\varepsilon(x) \cap A \neq \emptyset.$$

Da $\varepsilon > 0$ beliebig gewählt war, folgt daraus $x \in \bar{A}$. Damit haben wir gezeigt dass \bar{A} eine abgeschlossene Teilmenge von X ist. Sei nun $B \subset X$ eine beliebige abgeschlossene Menge die A enthält. Es bleibt zu zeigen, dass $\bar{A} \subset B$ ist. Sei also $x \in \bar{A}$. Dann gibt es nach (i) eine Folge $x_n \in A$, die gegen x konvergiert. Also ist x_n eine Folge in B und, da B abgeschlossen ist, folgt daraus, dass ihr Grenzwert x ebenfalls ein Element von B ist. Damit ist (iv) bewiesen.

Die Aussage (v) folgt aus (ii) und (iv). Die Aussage (vi) folgt unmittelbar aus (iii) und (iv). Die Aussage (vii) folgt unmittelbar aus (iii) und (v). Damit ist Proposition A.2.3 bewiesen. \square

Übung A.2.4. In einem metrischen Raum (X, d) bezeichnen wir oft den abgeschlossenen Ball mit Mittelpunkt $x \in X$ und Radius $r > 0$ mit

$$\bar{B}_r(x) := \{y \in X \mid d(x, y) \leq r\}.$$

Diese Menge ist stets abgeschlossen.

(i) Ist $(X, \|\cdot\|)$ ein normierter Vektorraum mit der durch

$$d(x, y) := \|x - y\|$$

für $x, y \in X$ definierten Abstandsfunktion, zeigen Sie, dass $\bar{B}_r(x)$ der Abschluss von $B_r(x)$ ist.

(ii) Finden Sie ein Beispiel eines metrischen Raumes (X, d) , eines Punktes $x \in X$, und einer Zahl $r > 0$, so dass $\bar{B}_r(x)$ nicht der Abschluss von $B_r(x)$ ist, das heißt $\bar{B}_r(x) \subsetneq B_r(x)$.

Übung A.2.5. Sei (X, d) ein metrischer Raum und seien $A, B \subset X$. Dann gilt

$$\overline{A \cup B} = \overline{A} \cup \overline{B}, \quad \overline{A \cap B} \subset \overline{A} \cap \overline{B}. \quad (\text{A.2.1})$$

und

$$\overset{\circ}{A} \cup \overset{\circ}{B} \subset A \overset{\circ}{\cup} B, \quad \overset{\circ}{A} \cap \overset{\circ}{B} = A \overset{\circ}{\cap} B. \quad (\text{A.2.2})$$

Für $X := \mathbb{R}$ und $A := \mathbb{Q}$ und $B := \mathbb{R} \setminus \mathbb{Q}$ erhalten wir $\overset{\circ}{A} = \overset{\circ}{B} = A \cap B = \emptyset$ und $\overline{A} = \overline{B} = A \cup B = X$.

Übung A.2.6. Sei (X, d) ein metrischer Raum und sei $A \subset X$. Dann gilt

$$\overset{\circ}{A} \subset \overset{\circ}{\overline{A}}, \quad \partial \overline{A} \subset \partial A \quad (\text{A.2.3})$$

und

$$\overline{\overset{\circ}{A}} \subset \overline{A}, \quad \partial \overset{\circ}{A} \subset \partial A. \quad (\text{A.2.4})$$

Für $A := \mathbb{Q}$ und $X := \mathbb{R}$ sind dies echte Inklusionen mit $\overset{\circ}{A} = \partial \overset{\circ}{A} = \partial \overline{A} = \emptyset$ und $\overline{A} = \partial A = X$. Für Intervalle in \mathbb{R} gilt Gleichheit in (A.2.3) und (A.2.4).

Übung A.2.7. Seien (X, d_X) und (Y, d_Y) metrische Räume, und $f : X \rightarrow Y$ eine stetige Abbildung. Dann gilt

$$f(\overline{A}) \subset \overline{f(A)}, \quad (\text{A.2.5})$$

für $A \subset X$ und

$$\overline{f^{-1}(B)} \subset f^{-1}(\overline{B}), \quad f^{-1}(\overset{\circ}{B}) \subset f^{-1}(\overset{\circ}{B}) \quad (\text{A.2.6})$$

für $B \subset Y$. Für $A = X = Y = \mathbb{R}$ und $f = \tan^{-1}$ ist (A.2.5) eine echte Inklusion. Für $X = Y = \mathbb{R}$, $f(x) := x^2$, und $A = \{-1\} \cup (-1/2, 1) \cup (1, 2)$ erhalten wir $f(A) = [0, 4)$, $f(\overset{\circ}{A}) = (0, 4)$, und $f(\overset{\circ}{A}) = [0, 1) \cup (1, 4)$; in diesem Beispiel ist also das Innere der Menge $f(A)$ weder in $f(\overset{\circ}{A})$ enthalten, noch enthält es die Menge $f(\overset{\circ}{A})$. Mit anderen Worten, für das Innere gibt es keine zu der Inklusion (A.2.5) analoge Beziehung.

Anhang B

Zusammenhängende Räume

B.1 Die Relativtopologie

Sei (X, d) ein metrischer Raum und $Y \subset X$ eine beliebige Teilmenge. Dann ist die Restriktion der Abstandsfunktion $d : X \times X \rightarrow \mathbb{R}$ auf $Y \times Y$ wieder eine Abstandsfunktion, und wir bezeichnen sie mit

$$d_Y := d|_{Y \times Y} : Y \times Y \rightarrow \mathbb{R}.$$

Damit wird (Y, d_Y) zu einem metrischen Raum.

Definition B.1.1. *Die durch d_Y induzierte Topologie auf Y wird auch die **Relativtopologie** auf Y genannt. Eine Teilmenge $V \subset Y$ heißt **Y -offen** wenn sie bezüglich der Relativtopologie offen ist. Eine Teilmenge $B \subset Y$ heißt **Y -abgeschlossen** wenn sie bezüglich der Relativtopologie abgeschlossen ist.*

Das folgende Lemma charakterisiert die Y -offenen und Y -abgeschlossenen Teilmengen von Y mit Hilfe der offenen und abgeschlossenen Teilmengen von X .

Lemma B.1.2. *Sei (X, d) ein metrischer Raum und $Y \subset X$ eine Teilmenge.*

(i) *Eine Teilmenge $V \subset Y$ ist genau dann Y -offen wenn es eine offenen Teilmenge $U \subset X$ gibt so dass $V = U \cap Y$ ist.*

(ii) *Eine Teilmenge $B \subset Y$ ist genau dann Y -abgeschlossen wenn es eine abgeschlossene Teilmenge $A \subset X$ gibt so dass $B = A \cap Y$ ist.*

Proof. Wir beweisen (i). Ist $V \subset Y$ offen bezüglich d_Y so gibt es für jedes $y \in V$ ein $\varepsilon > 0$ so dass

$$B_\varepsilon(y; Y) := \{x \in Y \mid d(x, y) < \varepsilon\} \subset V.$$

Für $y \in V$ sei $\varepsilon(y) := \sup\{\varepsilon \in (0, 1] \mid B_\varepsilon(y; Y) \subset V\}$. Dann ist $B_{\varepsilon(y)}(y; Y) \subset V$ für jedes $y \in V$. Damit ist die Menge

$$U := \bigcup_{y \in V} B_{\varepsilon(y)}(y; X)$$

offen in X und es gilt

$$U \cap Y = \bigcup_{y \in V} (B_{\varepsilon(y)}(y; X) \cap Y) = \bigcup_{y \in V} B_{\varepsilon(y)}(y; Y) = V.$$

Gibt es andererseits eine offene Teilmenge $U \subset X$ mit $U \cap Y = V$ und ist $y \in V$, so ist auch $y \in U$, also existiert ein $\varepsilon > 0$ mit $B_\varepsilon(y; X) \subset U$, und daher gilt $B_\varepsilon(y; Y) = B_\varepsilon(y; X) \cap Y \subset U \cap Y = V$. Also ist V offen bezüglich d_Y . Damit ist Teil (i) bewiesen.

Um (ii) zu zeigen, nehmen wir zunächst an, dass $B \subset Y$ abgeschlossen ist bezüglich d_Y . Dann ist

$$V := Y \setminus B$$

offen bezüglich d_Y . Nach (i) gibt es also eine offene Menge $U \subset X$ so dass $U \cap Y = V$. Damit ist

$$A := X \setminus U$$

eine abgeschlossen Teilmenge von X und es gilt

$$A \cap Y = (X \setminus U) \cap Y = Y \setminus (U \cap Y) = Y \setminus V = B.$$

Ist andererseits $A \subset X$ abgeschlossen und $B = A \cap Y$, so ist

$$U := X \setminus A$$

eine offene Teilmenge von X und es gilt

$$U \cap Y = (X \setminus A) \cap Y = Y \setminus (A \cap Y) = Y \setminus B.$$

Nach (i) ist daher die Menge $Y \setminus B$ offen bezüglich d_Y , und folglich ist B abgeschlossen bezüglich d_Y . Damit ist Lemma B.1.2 bewiesen. \square

Beispiel B.1.3. Sei $X = \mathbb{R}$ mit der Standardmetrik

$$d(x, y) := |x - y|$$

für $x, y \in \mathbb{R}$.

(i) Ist

$$Y := [0, 1]$$

so ist $[0, b)$ für $0 < b \leq 1$ eine Y -offene Teilmenge von Y . Hingegen ist eine Teilmenge $B \subset Y$ genau dann Y -abgeschlossen wenn sie abgeschlossen ist (da Y selbst eine abgeschlossene Teilmenge von X ist).

(ii) Ist

$$Y := (0, 1)$$

so ist $(0, b]$ für $0 < b < 1$ eine Y -abgeschlossene Teilmenge von Y . Hingegen ist eine Teilmenge $V \subset Y$ genau dann Y -offen wenn sie offen ist (da Y selbst eine offene Teilmenge von X ist).

B.2 Der Zusammenhangsbegriff

Definition B.2.1. Ein metrischer Raum (X, d) heißt **zusammenhängend** wenn er sich nicht als disjunkte Vereinigung zweier nichtleerer offener Teilmengen darstellen läßt, d.h. wenn für je zwei offene Teilmengen $U, V \subset X$ folgendes gilt:

$$U \cup V = X, \quad U \cap V = \emptyset \quad \implies \quad U = \emptyset \quad \text{oder} \quad V = \emptyset;$$

das heißt, dass die leere Menge und der ganze Raum die einzigen Teilmengen von X sind, die sowohl offen als auch abgeschlossen sind.

Eine Teilmenge $Y \subset X$ eines metrischen Raumes (X, d) heißt **zusammenhängend**, wenn sie bezüglich der induzierten Metrik d_Y zusammenhängend ist. Nach Lemma B.1.2 heißt das, dass für je zwei offene Teilmengen $U, V \subset X$ folgende Aussage gilt:

$$A \subset U \cup V, \quad A \cap U \cap V = \emptyset \quad \implies \quad A \cap U = \emptyset \quad \text{oder} \quad A \cap V = \emptyset.$$

Der nächste Satz zeigt, dass die zusammenhängenden Teilmengen von \mathbb{R} gerade die Intervalle sind.

Satz B.2.2. *Sei $X = \mathbb{R}$ mit der Standardmetrik*

$$d(x, y) := |x - y|.$$

Eine Teilmenge $I \subset \mathbb{R}$ ist genau dann zusammenhängend wenn sie ein Intervall ist.

Beweis. Ist $I \subset \mathbb{R}$ kein Intervall, so gibt es drei reelle Zahlen $a, b, c \in \mathbb{R}$ mit

$$a, b \in I, \quad c \notin I, \quad a < c < b.$$

Definieren wir

$$U := \{x \in \mathbb{R} \mid x < c\}, \quad V := \{x \in \mathbb{R} \mid x > c\}$$

so sind dies offene Mengen mit

$$I \subset U \cup V, \quad I \cap U \neq \emptyset, \quad I \cap V \neq \emptyset.$$

Also ist I nicht zusammenhängend.

Nehmen wir andererseits an, I sei ein Intervall und $A \subset I$ sei eine Teilmenge die sowohl offen als auch abgeschlossen ist (bezüglich der Metrik d_I) und die weder leer noch gleich dem gesamten Intervall I ist. Dann ist auch

$$B := I \setminus A$$

nicht leer und wir wählen $a \in A$ und $b \in B$. Wir können ohne Einschränkung der Allgemeinheit annehmen, dass $a < b$ ist. Sei

$$c := \sup(A \cap [a, b]).$$

Dann gilt $c \in [a, b] \subset I$ und es gibt eine Folge $a_k \in A$ mit $a_k \leq c$ die gegen c konvergiert. Da A abgeschlossen bezüglich d_I ist, folgt hieraus dass $c \in A$ und damit $c < b$ ist. Damit ist das halboffene Intervall $(c, b]$ in B enthalten. Also gilt $c + 1/n \in B$ für jede hinreichend grosse natürliche Zahl n und damit

$$c = \lim_{n \rightarrow \infty} (c + 1/n) \in B,$$

da auch B abgeschlossen bezüglich d_I ist. Wir haben daher gezeigt, dass c sowohl ein Element von A als auch ein Element von B ist, im Widerspruch zu $A \cap B = \emptyset$. Damit ist Satz B.2.2 bewiesen. \square

Satz B.2.3. *Seien (X, d_X) und (Y, d_Y) metrische Räume und $f : X \rightarrow Y$ eine stetige Abbildung; ist $A \subset X$ zusammenhängend so ist auch $f(A) \subset Y$ zusammenhängend.*

Beweis. Seien $U, V \subset Y$ zwei offene Mengen so dass

$$f(A) \subset U \cup V, \quad f(A) \cap U \cap V = \emptyset.$$

Zu zeigen ist, dass eine der Mengen $f(A) \cap U$ oder $f(A) \cap V$ leer ist.

Da f stetig ist, sind $f^{-1}(U)$ und $f^{-1}(V)$ offenen Teilmengen von X . Wir zeigen nun, dass diese Teilmengen folgende Eigenschaften haben:

$$A \subset f^{-1}(U) \cup f^{-1}(V), \quad A \cap f^{-1}(U) \cap f^{-1}(V) = \emptyset. \quad (\text{B.2.1})$$

Ist nämlich $a \in A$ so ist $f(a) \in f(A) \subset U \cup V$, und daher $f(a) \in U$ oder $f(a) \in V$; im ersten Fall gilt $a \in f^{-1}(U)$ und im zweiten Fall gilt $a \in f^{-1}(V)$. Dies beweist die erste Aussage in (B.2.1). Die zweite Aussage in (B.2.1) zeigt man am besten indirekt. Gibt es ein Element $a \in A \cap f^{-1}(U) \cap f^{-1}(V)$ so gilt $f(a) \in f(A) \cap U \cap V$, im Widerspruch zu unserer Annahme über U und V . Damit sind beide Aussagen in (B.2.1) bewiesen.

Da A zusammenhängend ist, folgt aus (B.2.1), dass mindestens eine der Mengen $A \cap f^{-1}(U)$ oder $A \cap f^{-1}(V)$ leer ist. Wenn aber $A \cap f^{-1}(U) = \emptyset$ ist, so heißt das, dass kein Element von A unter f auf U abgebildet wird, und damit ist auch $f(A) \cap U = \emptyset$ ist; ebenso mit V statt U . Damit ist mindestens eine der Mengen $f(A) \cap U$ oder $f(A) \cap V$ die leere Menge, wie behauptet. Damit ist Satz B.2.3 bewiesen. \square

Satz B.2.4. (Zwischenwertsatz)

Sei (X, d) ein zusammenhängender metrischer Raum, sei $f : X \rightarrow \mathbb{R}$ eine stetige Funktion, und seien zwei Punkte $x, y \in X$ und eine reelle Zahl $c \in \mathbb{R}$ gegeben, welche die Ungleichungen

$$f(x) < c < f(y)$$

erfüllen. Dann gibt es ein Element $z \in X$ mit $f(z) = c$.

Beweis. Da X zusammenhängend ist, folgt aus Satz B.2.3 dass auch $f(X)$ zusammenhängend ist. Nach Satz B.2.2 ist also $f(X)$ ein Intervall. Da $f(x)$ und $f(y)$ Elemente vom $f(X)$ sind, folgt aus der Definition eines Intervalls dass $c \in f(X)$ ist. Damit ist Satz B.2.4 bewiesen. \square

B.3 Weg-zusammenhängende Mengen

Sei (X, d) ein metrischer Raum. Eine Teilmenge $A \subset X$ heißt **weg-zusammenhängend** wenn für je zwei Elemente $x_0, x_1 \in A$ eine stetige Abbildung

$$\gamma : [0, 1] \rightarrow A$$

existiert, welche die Punkte

$$\gamma(0) = x_0, \quad \gamma(1) = x_1$$

miteinander verbindet.

Satz B.3.1. (i) *Jede weg-zusammenhängende Teilmenge eines metrischen Raumes (X, d) ist zusammenhängend.*

(ii) *Jede offene zusammenhängende Teilmenge eines normierten Vektorraumes $(X, \|\cdot\|)$ ist weg-zusammenhängend.*

Beweis. Wir beweisen (i). Sei (X, d) ein metrischer Raum und $A \subset X$ eine weg-zusammenhängende Teilmenge. Es ist zu zeigen, dass A zusammenhängend ist. Wir nehmen an, dies sei nicht der Fall. Dann existieren offene Mengen $U_0, U_1 \subset X$ so dass

$$A \subset U_0 \cup U_1, \quad A \cap U_0 \cap U_1 = \emptyset, \quad A \cap U_0 \neq \emptyset, \quad A \cap U_1 \neq \emptyset.$$

Sei $x_0 \in A \cap U_0$ und $x_1 \in A \cap U_1$. Da A weg-zusammenhängend existiert eine stetige Abbildung $\gamma : [0, 1] \rightarrow A$ mit $\gamma(0) = x_0$ und $\gamma(1) = x_1$. Das Intervall $I := [0, 1]$ ist die disjunkte Vereinigung der beiden I -offenen Teilmengen

$$I_0 := \gamma^{-1}(U_0) = \{t \in I \mid \gamma(t) \in U_0\}$$

und

$$I_1 := \gamma^{-1}(U_1) = \{t \in I \mid \gamma(t) \in U_1\}.$$

Diese sind beide nichtleer da $0 \in I_0$ und $1 \in I_1$ ist. Dies steht aber im Widerspruch zu der Tatsache, dass das Intervall $I = [0, 1]$ nach Satz B.2.2 zusammenhängend ist. Dieser Widerspruch zeigt, dass die Menge A entgegen unserer ursprünglichen Annahme zusammenhängend sein muss, und damit ist Teil (i) bewiesen.

Wir beweisen (ii). Sei also $(X, \|\cdot\|)$ ein normierter reeller Vektorraum und sei $U \subset X$ eine zusammenhängende offene Teilmenge. Sei $x_0 \in U$ und sei $U_0 \subset U$ die Menge aller Punkte $x \in U$, die sich mit x_0 durch einen stetigen Weg in U verbinden lassen, das heißt

$$U_0 := \left\{ x \in U \mid \begin{array}{l} \text{es existiert eine stetige Abbildung} \\ \gamma : [0, 1] \rightarrow U \text{ mit } \gamma(0) = x_0 \text{ und } \gamma(1) = x \end{array} \right\}.$$

Diese Menge ist nichtleer, da $x_0 \in U_0$ ist.

Wir zeigen, dass U_0 eine offene Teilmenge von X ist. Sei $x \in U_0$. Da U offen ist, gibt es ein $\varepsilon > 0$ so dass $B_\varepsilon(x) = \{y \in X \mid \|x - y\| < \varepsilon\} \subset U$ ist. Da $x \in U_0$ ist, existiert eine stetige Abbildung $\beta : [0, 1] \rightarrow U$ mit

$$\beta(0) = x_0, \quad \beta(1) = x.$$

Sei $y \in B_\varepsilon(x)$ und definiere die Abbildung $\gamma : [0, 1] \rightarrow X$ durch

$$\gamma(t) := \begin{cases} \beta(2t), & \text{für } 0 \leq t \leq 1/2, \\ (2 - 2t)x + (2t - 1)y, & \text{für } 1/2 < t \leq 1. \end{cases}$$

Diese Abbildung ist stetig, da $\beta(1) = x$ ist; sie nimmt Werte in U an, da $B_\varepsilon(x) \subset U$ ist; und sie erfüllt $\gamma(0) = x_0$ und $\gamma(1) = y$. Damit ist $y \in U_0$ für jedes $y \in B_\varepsilon(x)$. Also haben wir gezeigt, dass für jedes $x \in U_0$ ein $\varepsilon > 0$ mit $B_\varepsilon(x) \subset U_0$ existiert. Also ist U_0 eine offene Teilmenge von X .

Wir zeigen, dass U_0 eine abgeschlossene Teilmenge von U bezüglich der Relativtopologie ist. Sei also $(x_k)_{k \in \mathbb{N}}$ eine Folge in U_0 die gegen ein Element $x \in U$ konvergiert. Wähle $\varepsilon > 0$ so, dass $B_\varepsilon(x) \subset U$ ist. Dann existiert ein $k \in \mathbb{N}$ mit $\|x_k - x\| < \varepsilon$. Da $x_k \in U_0$ ist, existiert eine stetige Abbildung $\beta : [0, 1] \rightarrow U$ mit

$$\beta(0) = x_0, \quad \beta(1) = x_k.$$

Definiere $\gamma : [0, 1] \rightarrow X$ durch

$$\gamma(t) := \begin{cases} \beta(2t), & \text{für } 0 \leq t \leq 1/2, \\ (2 - 2t)x_k + (2t - 1)x, & \text{für } 1/2 < t \leq 1. \end{cases}$$

Diese Abbildung ist stetig, da $\beta(1) = x_k$ ist; sie nimmt Werte in U an, da $B_\varepsilon(x) \subset U$ ist; und sie erfüllt $\gamma(0) = x_0$ und $\gamma(1) = x$. Daher ist $x \in U_0$. Damit ist gezeigt, dass U_0 eine nichtleere, U -offene, und U -abgeschlossene Teilmenge von U ist. Da U eine zusammenhängende Teilmenge von X ist folgt daraus, dass $U_0 = U$ ist. Da $x_0 \in U$ beliebig gewählt war, heißt das, dass U weg-zusammenhängend ist. Damit ist Satz B.3.1 bewiesen. \square

B.4 Zusammenhangskomponenten

Lemma B.4.1. *Sei (X, d) ein metrischer Raum, sei I eine beliebige Indexmenge, und sei $A_i \subset X$ eine zusammenhängende Teilmenge für jedes $i \in I$. Ist*

$$\bigcap_{i \in I} A_i \neq \emptyset,$$

so ist die Menge

$$A := \bigcup_{i \in I} A_i$$

zusammenhängend.

Beweis. Seien $U, V \subset X$ zwei offene Teilmengen so dass

$$A \subset U \cup V, \quad A \cap U \cap V = \emptyset.$$

Sei ein Element

$$x_0 \in \bigcap_{i \in I} A_i$$

fest gewählt. Dann gilt entweder $x_0 \in U$ oder $x_0 \in V$. Nehmen wir an $x_0 \in U$ und sei $i \in I$. Da $A_i \subset U \cup V$ und $A_i \cap U \cap V = \emptyset$ und A_i zusammenhängend ist, gilt entweder $A_i \cap U = \emptyset$ oder $A_i \cap V = \emptyset$. Da $x_0 \in A_i \cap U$ ist, folgt hieraus $A_i \cap V = \emptyset$ und damit $A_i \subset U$. Also gilt $A_i \subset U$ für jedes $i \in I$ und damit $A \subset U$ und $A \cap V = \emptyset$. Daher ist A zusammenhängend. \square

Sei (X, d) ein metrischer Raum. Wir definieren auf X folgende Äquivalenzrelation: zwei Elemente $x, y \in X$ heissen **äquivalent** (Notation $x \sim y$) wenn es eine zusammenhängende Teilmenge $A \subset X$ gibt mit $x, y \in A$. Nach Lemma B.4.1 ist dies in der Tat eine Äquivalenzrelation. Darüber hinaus folgt aus Lemma B.4.1, dass die Äquivalenzklassen dieser Äquivalenzrelation zusammenhängend sind; sie heißen **Zusammenhangskomponenten** von X . Ist $x_0 \in X$ so heißt die Teilmenge

$$A_0 := \{x \in X \mid x \sim x_0\}$$

die **Zusammenhangskomponente** von x_0 ; dies ist die grösste zusammenhängende Teilmenge von X die x_0 enthält.

B.5 Beispiele

Beispiel B.5.1. Sei $X := \mathbb{R}^n$ mit der euklidischen Metrik und sei $K \subset \mathbb{R}^n$ eine konvexe Teilmenge. Diese Teilmenge ist weg-zusammenhängend, denn für $x_0, x_1 \in K$ definiert die Formel $\gamma(t) := (1-t)x_0 + tx_1$ eine stetige Abbildung $\gamma : [0, 1] \rightarrow K$ mit $\gamma(0) = x_0$ und $\gamma(1) = x_1$. Also ist K nach Satz B.3.1 zusammenhängend.

Beispiel B.5.2. Sei $X := \mathbb{Q}$ die Menge der rationalen Zahlen mit der Standardmetrik $d(x, y) := |x - y|$. Dann besteht jede Zusammenhangskomponente nur aus einem einzigen Element. Einen solchen Raum nennt man auch **total unzusammenhängend**.

Beispiel B.5.3. Sei $X = \mathbb{R}^n$ mit der euklidischen Metrik. Dieser Raum ist zusammenhängend. Das Komplement $Y := \mathbb{R}^n \setminus \{0\}$ des Nullvektors ist zusammenhängend für $n \geq 2$ aber nicht für $n = 1$. **Übung:** Beweisen Sie diese Behauptungen. Schliessen Sie daraus, dass für $n > 1$ kein Homöomorphismus von \mathbb{R} nach \mathbb{R}^n existieren kann.

Beispiel B.5.4. Sei $X = \mathbb{R}^n$ mit der euklidischen Metrik. Dann ist die Einheitskugel $S^{n-1} := \{x \in \mathbb{R}^n \mid \|x\|_2 = 1\}$ zusammenhängend. Ihr Komplement $Y := \mathbb{R}^n \setminus S^{n-1}$ hat zwei Zusammenhangskomponenten für $n > 1$, und drei für $n = 1$. **Übung:** Beweisen Sie diese Behauptungen. **Hinweis:** Verwenden Sie den Begriff “weg-zusammenhängend”.

Beispiel B.5.5. Die Menge

$$A := \{(x, y) \mid x > 0, y = \sin(1/x)\} \cup \{(0, y) \mid -1 \leq y \leq 1\}$$

ist eine abgeschlossene und zusammenhängende Teilmenge von \mathbb{R}^2 , ist aber nicht weg-zusammenhängend. **Übung:** Beweisen Sie diese Behauptungen.

Beispiel B.5.6. Die Gruppe

$$\mathrm{GL}^+(n, \mathbb{R}) := \{A \in \mathbb{R}^{n \times n} \mid \det(A) > 0\}$$

für jede natürliche Zahl n zusammenhängend. (Ein eleganter Beweis dieser Aussage findet sich in [2, Seite 36/37]; siehe auch Satz D.4.1)) Hieraus folgt, dass die Menge der $n \times n$ -Matrizen mit negativer Determinante ebenfalls zusammenhängend ist. Daher hat die Gruppe $\mathrm{GL}(n, \mathbb{R})$ genau zwei Zusammenhangskomponenten.

Anhang C

Kompakte metrische Räume

Der erste Abschnitt dieses Anhangs charakterisiert die folgenkompakten Teilmengen eines metrischen Raumes. Der zweite Teil beweist den Satz von Arzelà–Ascoli.

C.1 Der Kompaktheitsbegriff

Sei X eine Menge und sei $K \subset X$. Eine **Überdeckung** von K ist eine, durch eine Menge I indizierte, Ansammlung $\mathcal{U} = \{U_i\}_{i \in I}$ von Teilmengen $U_i \subset X$ so dass die Vereinigung dieser Teilmengen die Menge K enthält, das heißt $K \subset \bigcup_{i \in I} U_i$. Man kann eine Überdeckung auch als eine Menge $\mathcal{U} \subset 2^X$ von Teilmengen von X beschreiben mit der Eigenschaft, dass deren Vereinigung K enthält, das heißt $K \subset \bigcup_{U \in \mathcal{U}} U$. Die Beziehung zwischen diesen beiden Formulierungen desselben Phänomens ist, dass eine Teilmenge $U \subset X$ genau dann ein Element von \mathcal{U} ist wenn es ein $i \in I$ gibt so dass $U_i = U$ ist. Wenn die indizierten Mengen U_i paarweise verschieden sind (das heißt $U_i \neq U_j$ für $i \neq j$), so ist die Abbildung $I \rightarrow \mathcal{U} : i \mapsto U_i$ bijektiv.

Eine Überdeckung \mathcal{U} (bzw $\{U_i\}_{i \in I}$) heißt **endlich**, wenn die Menge \mathcal{U} (bzw die Indexmenge I) endlich ist.

Eine **Teilüberdeckung** von \mathcal{U} (bzw $\{U_i\}_{i \in I}$) ist eine Teilmenge von \mathcal{U} die immer noch eine Überdeckung ist (bzw eine Teilmenge $J \subset I$ so dass die Ansammlung $\{U_i\}_{i \in J}$ immer noch eine Überdeckung ist).

Ist (X, d) ein metrischer Raum, so ist eine **offene Überdeckung** einer Teilmenge $K \subset X$ eine Überdeckung, die nur aus offenen Mengen besteht; im Fall $\{U_i\}_{i \in I}$ heißt das, dass die Menge U_i für jedes $i \in I$ offen ist.

Definition C.1.1. Sei (X, d) ein metrischer Raum und sei $K \subset X$.

(i) Die Teilmenge K heißt **total beschränkt**, wenn sie entweder leer ist oder für jedes $\varepsilon > 0$ endlich viele Elemente $\xi_1, \dots, \xi_m \in K$ existieren, so dass

$$K \subset \bigcup_{i=1}^m B_\varepsilon(\xi_i)$$

ist. Hier bezeichnen wir mit $B_\varepsilon(\xi) := \{x \in X \mid d(x, \xi) < \varepsilon\}$ den **offenen Ball** in (X, d) vom Radius ε mit Mittelpunkt ξ .

(ii) Die Teilmenge K heißt **vollständig**, wenn jede Cauchy-Folge in K gegen ein Element von K konvergiert

(iii) Die Teilmenge K heißt **kompakt**, wenn jede Folge in K eine konvergente Teilfolge besitzt, die gegen ein Element von K konvergiert

Der folgende Satz zeigt, dass die kompakten Teilmengen eines metrischen Raumes nur durch die Topologie des Raumes (also die Gesamtheit der offenen Mengen) bestimmt werden. In allgemeinen topologischen Räumen wird Bedingung (ii) in Satz C.1.2 zur Definition des Kompaktheitsbegriffes benutzt.

Satz C.1.2. Sei (X, d) ein metrischer Raum und sei $K \subset X$. Dann sind folgende Aussagen äquivalent.

(i) K ist kompakt.

(ii) Jede offene Überdeckung von K besitzt eine endliche Teilüberdeckung.

(iii) K ist vollständig und total beschränkt.

Beweis. Die leere Menge $K = \emptyset$ erfüllt alle drei Bedingungen. Wir nehmen daher an, K sei nichtleer. Damit keine Verwirrung entsteht, bezeichnen wir in diesem Beweis die Menge aller offenen Teilmengen von X mit $\mathcal{T}(X, d) \subset 2^X$. (\mathcal{T} wie in “Topologie”.)

“(i) \implies (ii)”. Wir nehmen an, dass jede Folge in K eine Teilfolge besitzt, die gegen ein Element von K konvergiert. Sei

$$\mathcal{U} \subset \mathcal{T}(X, d)$$

eine offene Überdeckung von X . Wir beweisen in zwei Schritten, dass \mathcal{U} eine endliche Teilüberdeckung besitzt.

Schritt 1. Es gibt ein $\varepsilon > 0$ so dass für jedes $x \in K$ ein $U \in \mathcal{U}$ existiert mit $B_\varepsilon(x) \subset U$.

Als logische Formel kann man diese Aussage wie folgt schreiben:

$$(\exists \varepsilon > 0)(\forall x \in K)(\exists U \in \mathcal{U})(B_\varepsilon(x) \subset U).$$

Wir beweisen dies indirekt und nehmen an das Gegenteil sei der Fall, das heißt

$$(\forall \varepsilon > 0)(\exists x \in K)(\forall U \in \mathcal{U})(B_\varepsilon(x) \not\subset U).$$

Wähle $\varepsilon := 1/n$ mit $n \in \mathbb{N}$. Dann existiert nach dem abzählbaren Auswahlaxiom eine Folge $(x_n)_{n \in \mathbb{N}}$ in K , so dass für jedes $n \in \mathbb{N}$ folgendes gilt:

$$B_{1/n}(x_n) \not\subset U \quad \forall U \in \mathcal{U}.$$

Da K kompakt ist, besitzt die Folge $(x_n)_{n \in \mathbb{N}}$ eine Teilfolge $(x_{n_i})_{i \in \mathbb{N}}$, die gegen ein Element $x_0 := \lim_{i \rightarrow \infty} x_{n_i} \in K$ konvergiert. Sei $U \in \mathcal{U}$ mit $x_0 \in U$. Da U offen ist, gibt es ein $\varepsilon > 0$ mit $B_\varepsilon(x_0) \subset U$. Da x_{n_i} gegen x_0 konvergiert, gibt es eine natürliche Zahl $N \in \mathbb{N}$ so dass für alle $i \in \mathbb{N}$ folgendes gilt:

$$i \geq N \quad \implies \quad d(x_{n_i}, x_0) < \frac{\varepsilon}{2}.$$

Sei $i > N$ mit $1/n_i < \varepsilon/2$. Dann gilt $B_{1/n_i}(x_{n_i}) \subset B_{\varepsilon/2}(x_{n_i}) \subset B_\varepsilon(x_0) \subset U$, im Widerspruch zur Konstruktion der Folge (x_n) . Damit ist Schritt 1 bewiesen.

Schritt 2. \mathcal{U} besitzt eine endliche Teilüberdeckung.

Wir nehmen an \mathcal{U} habe keine endliche Teilüberdeckung. Sei $\varepsilon > 0$ eine Zahl für die die Behauptung von Schritt 1 gilt. Wir werden induktiv eine Folge $(x_n)_{n \in \mathbb{N}}$ in K und eine Folge $(U_n)_{n \in \mathbb{N}}$ in \mathcal{U} konstruieren, so dass

$$B_\varepsilon(x_1) \subset U_1$$

ist und

$$B_\varepsilon(x_n) \subset U_n, \quad x_n \notin U_1 \cup \dots \cup U_{n-1}$$

für alle $n \in \mathbb{N}$ mit $n \geq 2$. Zunächst wählen wir ein beliebiges Element $x_1 \in K$. Dann gibt es nach Schritt 1 ein Element $U_1 \in \mathcal{U}$ mit $B_\varepsilon(x_1) \subset U_1$. Nehmen wir nun an wir hätten x_1, \dots, x_k und U_1, \dots, U_k so konstruiert, dass unsere Bedingungen für $n = 1, \dots, k$ erfüllt sind. Da \mathcal{U} keine endliche Teilüberdeckung besitzt, muss es ein Element $x_{k+1} \in K$ geben das in keiner der Teilmengen U_1, \dots, U_k enthalten ist. Nach Schritt 1 gibt es ein $U_{k+1} \in \mathcal{U}$ mit $B_\varepsilon(x_{k+1}) \subset U_{k+1}$. Damit ist das Induktionsargument beendet und wir haben die gewünschten Folgen konstruiert.

Nach Konstruktion der Folgen gilt $x_n \notin U_m$ für alle $m, n \in \mathbb{N}$ mit $n > m$. Da $B_\varepsilon(x_m) \subset U_m$ ist, folgt daraus, dass $x_n \notin B_\varepsilon(x_m)$ und somit $d(x_n, x_m) \geq \varepsilon$ ist für alle $m, n \in \mathbb{N}$ mit $n > m$. Vertauschen wir n und m , so erhalten wir die Ungleichung $d(x_n, x_m) \geq \varepsilon$ für alle $m, n \in \mathbb{N}$ mit $n \neq m$. Also hat die Folge $(x_n)_{n \in \mathbb{N}}$ keine konvergente Teilfolge, im Widerspruch zu (i). Damit ist gezeigt, dass (ii) aus (i) folgt.

“(ii) \implies (iii)”. Wir nehmen jetzt an, dass jede offene Überdeckung von K eine endliche Teilüberdeckung besitzt. Daß K total beschränkt ist, folgt dann durch die Wahl der Überdeckung $\mathcal{U} := \{B_\varepsilon(x) \mid x \in K\}$. Nach (ii) existieren Elemente $\xi_1, \dots, \xi_N \in K$ mit $K \subset \bigcup_{i=1}^N B_\varepsilon(\xi_i)$. Damit ist K total beschränkt.

Wir zeigen nun, dass K vollständig ist. Sei also $(x_n)_{n \in \mathbb{N}}$ eine Cauchy-folge in K . Nehmen wir an, diese Folge konvergiere nicht gegen ein Element von K . Dann konvergiert nach einem Resultat aus der Analysis I auch keine Teilfolge von $(x_n)_{n \in \mathbb{N}}$ gegen ein Element von K . Damit besitzt die Folge $(x_n)_{n \in \mathbb{N}}$ auch keine Häufungspunkte in K (siehe Analysis I). Das heißt folgendes: Für jedes $\xi \in K$ gibt es ein $\varepsilon(\xi) > 0$ so dass der Ball $B_{\varepsilon(\xi)}(\xi)$ nur endlich viele Glieder der Folge $(x_n)_{n \in \mathbb{N}}$ enthält. Daraus folgt, dass die offene Überdeckung

$$\mathcal{U} := \{B_{\varepsilon(\xi)}(\xi) \mid \xi \in K\}$$

von K keine endliche Teilüberdeckung enthalten kann. Dies steht im Widerspruch zu (ii), und damit ist gezeigt, dass (iii) aus (ii) folgt.

“(iii) \implies (i)”. Wir nehmen nun an, dass K vollständig und total beschränkt ist. Sei $(x_n)_{n \in \mathbb{N}}$ eine Folge in K . Wir müssen eine konvergente Teilfolge von (x_n) finden, die gegen ein Element von K konvergiert. Dazu werden wir induktiv eine Folge unendlicher Teilmengen $T_k \subset \mathbb{N}$, $k = 0, 1, 2, 3, \dots$, konstruieren mit $\mathbb{N} \supset T_0 \supset T_1 \supset T_2 \supset \dots$ und

$$n, m \in T_k \quad \implies \quad d(x_n, x_m) \leq 2^{-k}.$$

Da (X, d) total beschränkt ist läßt sich K durch endlich viele Bälle $B_{1/2}(\xi_i)$ mit $i = 1, \dots, m$ überdecken. Einer dieser Bälle muss unendlich viele Glieder unserer Folge enthalten. Sei dies der Ball $B_{1/2}(\xi_i)$ und definiere

$$T_0 := \{n \in \mathbb{N} \mid x_n \in B_{1/2}(\xi_i)\}.$$

Nach Konstruktion ist dies eine unendliche Teilmenge von \mathbb{N} und es gilt

$$d(x_n, x_m) \leq d(x_n, \xi_i) + d(\xi_i, x_m) < 1 \quad \text{für alle } n, m \in T_0.$$

Sei nun $k \geq 1$ und seien $T_0 \supset T_1 \supset \dots \supset T_{k-1}$ bereits konstruiert. Da (X, d) total beschränkt ist können wir X durch endlich viele Bälle $B_{2^{-k-1}}(\xi_i)$ mit $i = 1, \dots, m$ überdecken. Einer dieser Bälle muss unendlich viele der Folgenglieder x_n mit $n \in T_{k-1}$ enthalten. Sei dies der Ball $B_{2^{-k-1}}(\xi_i)$ und definiere die Menge $T_k := \{n \in T_{k-1} \mid x_n \in B_{2^{-k-1}}(\xi_i)\}$. Dann ist T_k eine unendliche Teilmenge von T_{k-1} und es gilt $d(x_n, x_m) \leq d(x_n, \xi_i) + d(\xi_i, x_m) < 2^{-k}$ für alle $n, m \in T_k$. Damit sind die Mengen T_k für alle $k \in \mathbb{N}$ konstruiert.

Da jede der Teilmengen T_k unendlich ist, existiert eine Folge $n_k \in T_k$, so dass $n_k < n_{k+1}$ ist für alle $k \in \mathbb{N}$. Dann gilt $n_\ell \in T_k \subset T_\ell$ für $\ell \geq k$ und daher erfüllen alle $k, \ell \in \mathbb{N}$ mit $\ell \geq k$ die Ungleichung $d(x_k, x_\ell) \leq 2^{-k}$. Damit ist die Folge $(x_{n_k})_{k \in \mathbb{N}}$ eine Cauchy-Folge und, da K vollständig ist, konvergiert sie gegen ein Element von K . Wir haben also gezeigt, dass jede Folge in X eine konvergente Teilfolge besitzt. Also folgt (i) aus (iii). Damit ist Satz C.1.2 bewiesen. \square

C.2 Der Satz von Arzelà–Ascoli

Nach dem Satz von Heine–Borel ist eine Teilmenge des \mathbb{R}^n (mit der Standardmetrik, also der Euklidischen) genau dann kompakt wenn sie abgeschlossen und beschränkt ist. Insbesondere ist also der abgeschlossene Einheitsball im \mathbb{R}^n kompakt. Dies gilt auch für jeden beliebigen endlichdimensionalen Vektorraum, oder äquivalenterweise für jede beliebige Norm auf dem \mathbb{R}^n . Im Gegenzug zeigt die folgende Übung, dass diese Eigenschaft genau die endlichdimensionalen normierten Vektorräume charakterisiert.

Übung C.2.1. Sei V ein normierter reeller Vektorraum mit kompaktem Einheitsball $B := \{x \in V \mid \|x\| \leq 1\}$. Dann ist V endlichdimensional.

Hinweise: 1. Jeder endlichdimensionale lineare Unterraum $W \subset V$ ist abgeschlossen.

2. Ist $W \subset V$ ein abgeschlossener linearer Unterraum und $v \in V \setminus W$, so gilt $d(v, W) := \inf_{w \in W} \|v - w\| > 0$.

3. Ist $W \subsetneq V$ ein abgeschlossener linearer Unterraum so gibt es einen Vektor $v \in V$ mit $\|v\| = 1$ und $d(v, W) \geq 1/2$. (Sei $v_0 \in V \setminus W$ und wähle $w_0 \in W$ so dass $\|v_0 - w_0\| \leq 2d(v_0, W)$; definiere $v := \|v_0 - w_0\|^{-1}(v_0 - w_0)$).

4. Ist V unendlichdimensional so gibt es eine Folge $(v_n)_{n \in \mathbb{N}}$ in V mit $\|v_n\| = 1$ und $\|v_n - v_m\| \geq 1/2$ für alle $n, m \in \mathbb{N}$ mit $n \neq m$. Eine solche Folge hat keine konvergente Teilfolge.

Übung C.2.1 gibt uns ein wichtiges Kriterium für die Endlichdimensionalität eines normierten Vektorraumes. Sie zeigt auch, dass das Heine-Borel-Prinzip in unendlichdimensionalen normierten Vektorräumen nicht gilt: es gibt dort abgeschlossene und beschränkte Teilmengen die nicht kompakt sind. Damit stellt sich die Frage, welche Teilmengen eines solchen unendlichdimensionalen Raumes denn kompakt sind. Eine Antwort auf diese Frage ist in vielen Anwendungen von grundlegender Bedeutung. Ein besonders wichtiges Kompaktheitskriterium für Teilmengen des Raumes der stetigen Funktionen gibt uns der Satz von Arzelà-Ascoli.

Sei (X, d) ein kompakter metrischer Raum und $\mathcal{C}(X)$ der Raum der stetigen Funktionen $f : X \rightarrow \mathbb{R}$. In der Vorlesung haben wir gezeigt, dass dieser Raum mit der Supremumsnorm

$$\|f\| := \sup_{x \in X} |f(x)|$$

zu einem Banachraum, das heißt zu einem vollständigen normierten Vektorraum, wird (siehe [5, Satz 7.1]). Wir haben auch gezeigt, dass jede stetige Funktion $f : X \rightarrow \mathbb{R}$ auf einem kompakten metrischen Raum (X, d) **gleichmäßig stetig** ist, das heißt, dass für jedes $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass alle $x, y \in X$ mit $d(x, y) < \delta$ die Ungleichung $|f(x) - f(y)| < \varepsilon$ erfüllen. Wir nennen eine Teilmenge $\mathcal{F} \subset \mathcal{C}(X)$ **gleichgradig stetig** wenn die Zahl $\delta > 0$ in dieser Definition unabhängig von der Funktion $f \in \mathcal{F}$ gewählt werden kann.

Definition C.2.2. Eine Teilmenge $\mathcal{F} \subset \mathcal{C}(X)$ heißt **gleichgradig stetig**, wenn es für jedes $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für alle $f \in \mathcal{C}(X)$ und alle $x, y \in X$ gilt:

$$f \in \mathcal{F}, \quad d(x, y) < \delta \quad \implies \quad |f(x) - f(y)| < \varepsilon.$$

Übung C.2.3. Jede endliche Teilmenge von $\mathcal{C}(X)$ ist gleichgradig stetig. Seien c und μ zwei positive reelle Zahlen. Dann ist die Teilmenge

$$\mathcal{F} := \left\{ f \in \mathcal{C}(X) \mid \|f\| \leq c, \sup_{x \neq y} \frac{|f(x) - f(y)|}{d(x, y)^\mu} \leq c \right\}$$

abgeschlossen, beschränkt, und gleichgradig stetig. Finden Sie eine Folge in $\mathcal{C}([0, 1])$ die beschränkt aber nicht gleichgradig stetig ist.

Satz C.2.4 (Arzelà–Ascoli). *Sei (X, d) ein kompakter metrischer Raum. Eine Teilmenge $\mathcal{F} \subset \mathcal{C}(X)$ des Raumes der stetigen reellwertigen Funktionen auf X ist genau dann kompakt bezüglich der Supremumsnorm wenn sie folgende Eigenschaften hat:*

- (i) \mathcal{F} ist abgeschlossen.
- (ii) \mathcal{F} ist beschränkt.
- (iii) \mathcal{F} ist gleichgradig stetig.

Beweis. Wir nehmen zunächst an, dass \mathcal{F} eine kompakte Teilmenge von $\mathcal{C}(X)$ ist. Dann ist \mathcal{F} abgeschlossen, nach einem Resultat in Analysis I [6]. Außerdem ist \mathcal{F} beschränkt, da eine Folge $f_n \in \mathcal{F}$ mit $\|f_n\| \rightarrow \infty$ keine konvergente Teilfolge hätte. Es bleibt also zu zeigen, dass die Menge \mathcal{F} gleichgradig stetig ist. Sei $\varepsilon > 0$. Da die Menge $\mathcal{F} \subset \mathcal{C}(X)$ nach Satz C.1.2 total beschränkt ist, gibt es endlich viele Funktionen $f_1, \dots, f_m \in \mathcal{F}$ mit

$$\mathcal{F} \subset \bigcup_{i=1}^m B_{\varepsilon/3}(f_i; \mathcal{C}(X)).$$

Da (X, d) kompakt ist, ist jede der Funktionen $f_i : X \rightarrow \mathbb{R}$ gleichmäßig stetig nach einem Satz aus Analysis I [6]. Daher gibt es für jedes $i \in \{1, \dots, m\}$ eine Konstante $\delta_i > 0$ so dass für alle $x, y \in X$ gilt:

$$d(x, y) < \delta_i \quad \implies \quad |f_i(x) - f_i(y)| < \varepsilon/3.$$

Wähle

$$\delta := \min\{\delta_1, \dots, \delta_m\} > 0.$$

Sei $f \in \mathcal{F}$. Dann existiert ein Index $i \in \{1, \dots, m\}$ mit $\|f - f_i\| < \varepsilon/3$. Daher erfüllen alle $x, y \in X$ mit $d(x, y) < \delta \leq \delta_i$ die Ungleichung

$$|f(x) - f(y)| \leq |f(x) - f_i(x)| + |(f_i(x) - f_i(y))| + |f_i(y) - f(y)| < \varepsilon.$$

Also ist die Menge \mathcal{F} gleichgradig stetig.

Umgekehrt nehmen wir nun an, die Menge $\mathcal{F} \subset \mathcal{C}(X)$ sei abgeschlossen, beschränkt und gleichgradig stetig. Wir werden in vier Schritten beweisen, dass \mathcal{F} kompakt ist.

Schritt 1. Es gibt eine Folge $(x_k)_{k \in \mathbb{N}}$ in X mit der folgenden Eigenschaft. Für jedes $\delta > 0$ gibt es ein $m = m(\delta) \in \mathbb{N}$ so dass

$$X = \bigcup_{k=1}^{m(\delta)} B_\delta(x_k).$$

Wir konstruieren die Folge induktiv. Zunächst wissen wir, nach Satz C.1.2, dass der metrische Raum (X, d) total beschränkt ist. Also existieren, für $\delta = 1$, endlich viele Punkte $x_1, \dots, x_{m_1} \in X$ so dass

$$X = \bigcup_{k=1}^{m_1} B_1(x_k).$$

Ebenso gibt es für $\delta = 1/2$ endlich viele Punkte $x_{m_1+1}, \dots, x_{m_2} \in X$ so dass

$$X = \bigcup_{k=m_1+1}^{m_2} B_{1/2}(x_k) = \bigcup_{k=1}^{m_2} B_{1/2}(x_k).$$

Wenn $x_1, \dots, x_{m_{n-1}}$ konstruiert sind, wählen wir $x_{m_{n-1}+1}, \dots, x_{m_n} \in X$ so dass

$$X = \bigcup_{k=m_{n-1}+1}^{m_n} B_{1/n}(x_k) = \bigcup_{k=1}^{m_n} B_{1/n}(x_k).$$

Damit gilt die Behauptung von Schritt 1 mit $m(\delta) = m_n$, wobei n so gewählt wird dass $1/n < \delta$.

Schritt 2. Sei $(f_n)_{n \in \mathbb{N}}$ eine Folge in \mathcal{F} . Dann existiert eine Teilfolge $(f_{n_i})_{i \in \mathbb{N}}$ so dass der Grenzwert

$$y_k := \lim_{i \rightarrow \infty} f_{n_i}(x_k)$$

für jedes $k \in \mathbb{N}$ existiert.

Dies ist ein typisches Diagonalfolgen-Argument. Zunächst sei $k = 1$, dann ist die Folge $(f_n(x_1))_{n \in \mathbb{N}}$ reeller Zahlen beschränkt und hat daher, nach dem Satz von Bolzano–Weierstraß, eine konvergente Teilfolge. Mit anderen Worten, es existiert eine strikt monoton wachsende Funktion $g_1 : \mathbb{N} \rightarrow \mathbb{N}$ so dass der Grenzwert $y_1 := \lim_{i \rightarrow \infty} f_{g_1(i)}(x_1)$ existiert. Nun ist die Folge $(f_{g_1(i)}(x_2))_{n \in \mathbb{N}}$ ebenfalls beschränkt und besitzt somit auch eine konvergente Teilfolge. Also gibt es eine strikt monoton wachsende Funktion $g_2 : \mathbb{N} \rightarrow \mathbb{N}$ so dass

der Grenzwert $y_2 := \lim_{i \rightarrow \infty} f_{g_1 \circ g_2(i)}(x_2)$ existiert. Mit vollständiger Induktion finden wir nun eine Folge strikt monoton wachsender Funktionen $g_k : \mathbb{N} \rightarrow \mathbb{N}$ so dass der Grenzwert

$$y_k := \lim_{i \rightarrow \infty} f_{g_1 \circ \dots \circ g_k(i)}(x_k)$$

für jedes $k \in \mathbb{N}$ existiert. Nun sei

$$n_i := g_1 \circ \dots \circ g_i(i)$$

für $i \in \mathbb{N}$. Dann ist $(f_{n_i(i)}(x_k))_{i \geq k}$ eine Teilfolge von $(f_{g_1 \circ \dots \circ g_k(i)}(x_k))_{k \in \mathbb{N}}$ und konvergiert somit gegen y_k . Da dies für jedes $k \in \mathbb{N}$ gilt, ist Schritt 2 bewiesen.

Schritt 3. $(f_{n_i})_{i \in \mathbb{N}}$ ist eine Cauchy-Folge in $\mathcal{C}(X)$.

Sei $\varepsilon > 0$. Da \mathcal{F} gleichgradig stetig ist, gibt es ein $\delta > 0$ so dass für alle $f \in \mathcal{C}(X)$ und alle $x, y \in X$ folgendes gilt:

$$f \in \mathcal{F}, \quad d_X(x, y) < \delta \quad \implies \quad |f(x) - f(y)| < \varepsilon/3. \quad (\text{C.2.1})$$

Nach Schritt 1 gibt es eine natürliche Zahl $m = m(\delta) \in \mathbb{N}$ so dass

$$X = \bigcup_{k=1}^m B_\delta(x_k). \quad (\text{C.2.2})$$

Nach Schritt 2 gibt es ein $N \in \mathbb{N}$ so dass für alle $i, j, k \in \mathbb{N}$ folgendes gilt:

$$k \leq m, \quad i, j \geq N \quad \implies \quad |f_{n_i}(x_k) - f_{n_j}(x_k)| < \varepsilon/3. \quad (\text{C.2.3})$$

Wir behaupten, dass mit dieser Wahl von N die Ungleichung $\|f_{n_i} - f_{n_j}\| \leq \varepsilon$ für alle $i, j \geq N$ gilt. Sei also $x \in X$ und wähle eine Zahl $k \in \{1, \dots, m\}$ so dass $d(x, x_k) < \delta$ ist (siehe (C.2.2)). Dann folgt aus (C.2.1) die Ungleichung

$$|f_{n_i}(x_k) - f_{n_i}(x)| < \varepsilon/3 \quad (\text{C.2.4})$$

für alle $i \in \mathbb{N}$. Aus (C.2.3) und (C.2.4) folgt die Ungleichung

$$\begin{aligned} |f_{n_i}(x) - f_{n_j}(x)| &\leq |f_{n_i}(x) - f_{n_i}(x_k)| + |f_{n_i}(x_k) - f_{n_j}(x_k)| \\ &\quad + |f_{n_j}(x_k) - f_{n_j}(x)| \\ &< \varepsilon \end{aligned}$$

für alle $i, j \geq N$. Damit haben wir gezeigt, dass die in Schritt 2 konstruierte Teilfolge $(f_{n_i})_{i \in \mathbb{N}}$ eine Cauchy-Folge in $\mathcal{C}(X)$ ist.

Schritt 4. Die Folge $(f_{n_i})_{i \in \mathbb{N}}$ konvergiert gegen eine Funktion $f \in \mathcal{F}$.

Da $\mathcal{C}(X)$ ein Banachraum ist, konvergiert die Folge $(f_{n_i})_{i \in \mathbb{N}}$ in Schritt 3 gegen ein $f \in \mathcal{C}(X)$. Da \mathcal{F} abgeschlossen ist, gehört der Grenzwert zu \mathcal{F} . Also haben wir gezeigt, dass jede Folge in \mathcal{F} eine konvergente Teilfolge mit Grenzwert in \mathcal{F} besitzt. Also ist die Menge \mathcal{F} kompakt und damit ist Satz C.2.4 bewiesen. \square

Der Satz von Arzelà-Ascoli und sein Beweis lassen sich Wort für Wort auf den Raum $\mathcal{C}(X, V)$ der stetigen Funktionen auf X mit Werten in einem endlichdimensionalen Banachraum V übertragen. Im unendlichdimensionalen Fall bedarf es der Zusatzbedingung, dass die Menge $\{f(x) \mid f \in \mathcal{F}\}$ für jedes $x \in X$ eine kompakte Teilmenge von V ist (siehe [1, Cor 1.1.12]).

Übung C.2.5. Finden Sie eine Teilmenge des Raumes $\mathcal{BC}(\mathbb{R})$ (der beschränkten stetigen Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$), welche zwar abgeschlossen, beschränkt und gleichgradig stetig, aber nicht kompakt ist.

Anhang D

Die Determinante

D.1 Gruppen

Definition D.1.1. Eine **Gruppe** ist ein Tripel $(G, \cdot, \mathbb{1})$, bestehend aus einer Menge G , einer Abbildung

$$G \times G \rightarrow G : (g, h) \mapsto g \cdot h$$

(der **Gruppenoperation**), und einem Element $\mathbb{1} \in G$ (dem **neutralen Element**), welche folgende Bedingungen erfüllen.

- (i) Die Gruppenoperation ist assoziativ, d.h. es gilt $g_1 \cdot (g_2 \cdot g_3) = (g_1 \cdot g_2) \cdot g_3$ für alle $g_1, g_2, g_3 \in G$.
- (ii) Es gilt $\mathbb{1} \cdot g = g \cdot \mathbb{1} = g$ für alle $g \in G$.
- (iii) Für jedes $g \in G$ gibt es ein Element $h \in G$ so dass $g \cdot h = h \cdot g = \mathbb{1}$.

Das Element h in (iii) ist eindeutig durch g bestimmt. Es heißt **Inverse von g** und wird mit g^{-1} bezeichnet. Eine Gruppe $(G, \cdot, \mathbb{1})$ heißt **kommutativ** (oder **abelsch**) wenn

$$g \cdot h = h \cdot g \quad \forall g, h \in G.$$

Beispiel D.1.2. Die ganzen Zahlen bilden eine Gruppe \mathbb{Z} mit der Addition als Gruppenoperation und der Zahl 0 als neutralem Element. Gleiches gilt für die rationalen, die reellen und die komplexen Zahlen.

Beispiel D.1.3. Die von Null verschiedenen rationalen Zahlen bilden eine Gruppe mit der Multiplikation als Gruppenoperation und der Zahl 1 als neutralem Element. Gleiches gilt für die reellen und die komplexen Zahlen.

Beispiel D.1.4. Die zwei-elementige Menge $\mathbb{Z}_2 := \{0, 1\}$ ist eine Gruppe mit der Gruppenoperation $0 + 0 := 0$, $0 + 1 := 1$, $1 + 0 := 1$ und $1 + 1 := 0$. Die 0 ist das neutrale Element.

Beispiel D.1.5. Die zwei-elementige Menge $\{\pm 1\}$ ist eine Gruppe mit der Gruppenoperation $1 \cdot 1 := 1$, $(-1) \cdot (-1) := 1$, $1 \cdot (-1) := -1$ und $(-1) \cdot 1 := -1$. Die 1 ist das neutrale Element.

Beispiel D.1.6. Die Menge $S^1 := \{z \in \mathbb{C} \mid |z| = 1\}$ aller komplexen Zahlen vom Betrag eins ist eine Gruppe mit der Multiplikation als Gruppenoperation und der Zahl 1 als neutralem Element.

Beispiel D.1.7. Sei X eine Menge. Die bijektiven Abbildungen $f : X \rightarrow X$ bilden eine Gruppe mit der Komposition als Gruppenoperation und der Identität als neutralem Element.

Beispiel D.1.8. Sei $X := \{1, \dots, n\}$. Eine **Permutation (von X)** ist eine bijektive Abbildung $\sigma : X \rightarrow X$. Die Menge aller Permutationen von X bezeichnen wir mit S_n . Nach Beispiel D.1.6 ist S_n eine Gruppe. Für $n \geq 3$ ist diese Gruppe nicht kommutativ.

Definition D.1.9. Seien G und H zwei Gruppen. (Die beiden Gruppenoperationen bezeichnen wir mit \cdot und die beiden neutralen Elemente mit $\mathbb{1}_G$ und $\mathbb{1}_H$.) Eine Abbildung

$$\rho : G \rightarrow H$$

heißt **Gruppenhomomorphismus** (oder **Homomorphismus**) wenn sie für alle $a, b \in G$ die Gleichung

$$\rho(a \cdot b) = \rho(a) \cdot \rho(b), \quad \rho(\mathbb{1}_G) = \mathbb{1}_H,$$

erfüllt. Ein **Gruppenisomorphismus** (oder **Isomorphismus**) ist ein bijektiver Gruppenhomomorphismus.

Beispiel D.1.10. Die Abbildung

$$\{0, 1\} \rightarrow \{\pm 1\} : a \mapsto (-1)^a$$

ist ein Isomorphismus.

Beispiel D.1.11. Die Abbildung

$$\mathbb{R} \rightarrow S^1 : t \mapsto \exp(it)$$

ist ein Gruppenhomomorphismus.

Beispiel D.1.12. Die Exponentialfunktion

$$\exp : \mathbb{R} \rightarrow (0, \infty)$$

ist ein Isomorphismus von der Additiven Gruppe der reellen Zahlen in die multiplikative Gruppe der positiven reellen Zahlen. Die Umkehrabbildung

$$\log : (0, \infty) \rightarrow \mathbb{R}$$

ist ebenfalls ein Gruppenisomorphismus.

Beispiel D.1.13. Sei $(G, \cdot, \mathbb{1})$ eine Gruppe, X eine Menge, und $\mathcal{G}(X)$ die Gruppe der bijektiven Abbildungen $f : X \rightarrow X$. Eine **Gruppenaktion von G auf X** ist eine Abbildung

$$\phi : G \times X \rightarrow X$$

so dass für alle $x \in X$ und alle $a, b \in G$ die Gleichungen

$$\phi(a \cdot b, x) = \phi(a, \phi(b, x)), \quad \phi(\mathbb{1}, x) = x$$

gelten. Ist ϕ eine solche Gruppenaktion und definieren wir die Abbildungen $\phi_a : X \rightarrow X$ für $a \in G$ durch

$$\phi_a(x) := \phi(a, x) \quad \text{für } x \in X,$$

so ist ϕ_a für jedes $a \in G$ bijektiv (mit Umkehrabbildung $\phi_{a^{-1}}$) und die Abbildung

$$G \rightarrow \mathcal{G}(X) : a \mapsto \phi_a$$

ist ein Gruppenhomomorphismus, denn es gilt

$$\phi_{a \cdot b} = \phi_a \circ \phi_b, \quad \phi_{\mathbb{1}} = \text{id},$$

für alle $a, b \in G$. Umgekehrt definiert jeder Gruppenhomomorphismus von G nach $\mathcal{G}(X)$ eine Gruppenaktion von G auf X .

Beispiel D.1.14. Wir ordnen einer Permutation $\sigma \in S_n$ die natürliche Zahl

$$N(\sigma) := \#\{(i, j) \in \mathbb{N} \times \mathbb{N} \mid 1 \leq i < j \leq n, \sigma(i) > \sigma(j)\} \quad (\text{D.1.1})$$

zu. Die **Parität** von σ ist definiert durch

$$\varepsilon(\sigma) := (-1)^{N(\sigma)}. \quad (\text{D.1.2})$$

Das folgende Lemma zeigt, dass die Abbildung $\varepsilon : S_n \rightarrow \{\pm 1\}$ ein Gruppenhomomorphismus ist.

Lemma D.1.15. *Es gilt $\varepsilon(\text{id}) = 1$ und, für alle $\sigma, \tau \in S_n$,*

$$\varepsilon(\tau \circ \sigma) = \varepsilon(\tau)\varepsilon(\sigma).$$

Beweis. Wir definieren

$$N(\sigma, \tau) := \#\{(i, j) \mid i < j, \sigma(i) > \sigma(j), \tau(\sigma(i)) < \tau(\sigma(j))\}.$$

Dann gilt

$$\begin{aligned} N(\tau \circ \sigma) &= \#\{(i, j) \mid i < j, \tau(\sigma(i)) > \tau(\sigma(j))\} \\ &= \#\{(i, j) \mid i < j, \sigma(i) > \sigma(j), \tau(\sigma(i)) > \tau(\sigma(j))\} \\ &\quad + \#\{(i, j) \mid i < j, \sigma(i) < \sigma(j), \tau(\sigma(i)) > \tau(\sigma(j))\} \\ &= \#\{(i, j) \mid i < j, \sigma(i) > \sigma(j)\} \\ &\quad - \#\{(i, j) \mid i < j, \sigma(i) > \sigma(j), \tau(\sigma(i)) < \tau(\sigma(j))\} \\ &\quad + \#\{(i, j) \mid \sigma(i) < \sigma(j), \tau(\sigma(i)) > \tau(\sigma(j))\} \\ &\quad - \#\{(i, j) \mid i > j, \sigma(i) < \sigma(j), \tau(\sigma(i)) > \tau(\sigma(j))\} \\ &= N(\sigma) + N(\tau) - 2N(\sigma, \tau). \end{aligned}$$

Hieraus folgt sofort die Behauptung. □

Übung D.1.16. Für $i, j \in \{1, \dots, n\}$ mit $i \neq j$ definieren wir die Permutation $\sigma_{ij} \in S_n$ als die Bijektion welche i und j vertauscht:

$$\sigma_{ij}(\nu) := \begin{cases} \nu, & \text{wenn } \nu \neq i \text{ und } \nu \neq j, \\ i, & \text{wenn } \nu = j, \\ j, & \text{wenn } \nu = i. \end{cases}$$

Eine solche Permutation nennt man **Transposition**. Beweisen Sie, dass jede Transposition Parität -1 hat und dass sich jede Permutation als Komposition von Transpositionen schreiben lässt.

D.2 Vektorräume

Definition D.2.1. Ein **reeller Vektorraum** besteht aus einer abelschen Gruppe $(V, +, 0)$ und einer Abbildung

$$\mathbb{R} \times V \rightarrow V : (\lambda, v) \mapsto \lambda v$$

(der **skalaren Multiplikation**) mit folgenden Eigenschaften.

(i) Für alle $\lambda, \mu \in \mathbb{R}$ und $v \in V$ gilt

$$\lambda \cdot (\mu \cdot v) = (\lambda\mu) \cdot v, \quad 1 \cdot v = v.$$

(ii) Für alle $\lambda, \mu \in \mathbb{R}$ und $v, w \in V$ gilt

$$(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v, \quad \lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w.$$

Die Elemente von V nennen wir **Vektoren**.

Bemerkung D.2.2. Um Verwirrungen zu vermeiden kann es manchmal nützlich sein, für den Nullvektor in V ein anderes Symbol zu verwenden, zum Beispiel 0_V . Mit dieser Bezeichnungsweise gelten die Regeln

$$\lambda \cdot 0_V = 0_V, \quad 0 \cdot v = 0_V$$

für alle $\lambda \in \mathbb{R}$ und $v \in V$. (Beweisen Sie diese Gleichungen!) Im allgemeinen bezeichnen wir den Nullvektor in V jedoch mit dem gleichen Symbol 0 wie die reelle Zahl Null. Wir werden im folgenden auch oft den Punkt “ \cdot ” weglassen und λv statt $\lambda \cdot v$ schreiben.

Definition D.2.3. Seien V und W zwei Vektorräume. Eine Abbildung $T : V \rightarrow W$ heißt **linear** wenn sie die Gleichungen

$$T(v_1 + v_2) = Tv_1 + Tv_2, \quad T(\lambda \cdot v) = \lambda \cdot Tv$$

für alle $\lambda \in \mathbb{R}$ und $v, v_1, v_2 \in V$ erfüllt. Eine lineare Abbildung wird auch **Vektorraumhomomorphismus** genannt. Ein **Vektorraumisomorphismus** (oder einfach **Isomorphismus**) ist ein bijektiver Vektorraumhomomorphismus.

Übung D.2.4. Ist $T : V \rightarrow W$ eine bijektive lineare Abbildung so ist die Umkehrabbildung $T^{-1} : W \rightarrow V$ ebenfalls linear.

Definition D.2.5. Sei V ein reeller Vektorraum. Eine endliche Folge

$$v_1, \dots, v_n$$

von Vektoren in V heißt **linear unabhängig** wenn für alle $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ folgendes gilt:

$$\lambda_1 v_1 + \dots + \lambda_n v_n = 0 \quad \implies \quad \lambda_1 = \lambda_2 = \dots = \lambda_n = 0.$$

Den Ausdruck $\lambda_1 v_1 + \dots + \lambda_n v_n$ nennen wir eine **Linearkombination** der Vektoren v_1, \dots, v_n . Wir nennen die Vektoren v_1, \dots, v_n ein **(endliches) Erzeugendensystem** von V wenn sich jeder Vektor in V als Linearkombination dieser Vektoren darstellen läßt, das heißt

$$\forall v \in V \exists \lambda_1, \dots, \lambda_n \in \mathbb{R} : v = \lambda_1 v_1 + \dots + \lambda_n v_n.$$

Eine **endliche Basis** von V ist ein endliches linear unabhängiges Erzeugendensystem v_1, \dots, v_n . Ein reeller Vektorraum V heißt **endlichdimensional** wenn er eine endliche Basis besitzt und **unendlichdimensional** wenn er keine endliche Basis besitzt.

Bemerkung D.2.6. Man kann den Begriff einer Basis auch auf unendliche Systeme $\{v_i\}_{i \in I}$ von Vektoren in V erweitern. Hier ist I eine (möglicherweise unendliche) Menge und $I \rightarrow V : i \mapsto v_i$ eine Abbildung. Ein solches System heißt **linear unabhängig** wenn je endlich viele Vektoren v_{i_1}, \dots, v_{i_N} (mit $i_1, \dots, i_N \in I$ paarweise verschieden) linear unabhängig sind. Es heißt **Erzeugendensystem** wenn sich jeder Vektor $v \in V$ als Linearkombination endlich vieler v_i darstellen läßt. Eine **Basis** ist nun ein linear unabhängiges Erzeugendensystem. Mit diesem verallgemeinerten Begriff läßt sich zeigen, dass jeder Vektorraum eine Basis besitzt. (Der Beweis beruht auf dem sogenannten *Lemma von Zorn* welches sich wiederum aus dem *Auswahlaxiom* herleiten läßt.) Wir werden uns hier jedoch nur für endliche Basen interessieren. (Unendlichdimensionale Vektorräume spielen eine wichtige Rolle in der Mathematik, nicht aber ihre Basen.)

Beispiel D.2.7. Der Raum \mathbb{R}^n aller n -Tupel $x = (x_1, \dots, x_n)$ reeller Zahlen ist ein Vektorraum. Die Vektoren $e_i := (0, \dots, 0, 1, 0, \dots, 0)$ (mit 1 an der i ten Stelle) bilden eine Basis von \mathbb{R}^n , welche die **Standardbasis** genannt wird. Ein Vektor $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ hat die Form

$$x = \sum_{i=1}^n x_i e_i.$$

Bemerkung D.2.8. Sei V ein reeller Vektorraum und $v_1, \dots, v_n \in V$. Dann ist die durch

$$Tx := \sum_{i=1}^n x_i v_i, \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n,$$

definierte Abbildung $T : \mathbb{R}^n \rightarrow V$ linear. Die Abbildung T ist genau dann injektiv wenn die Vektoren v_1, \dots, v_n linear unabhängig sind; sie ist genau dann surjektiv wenn die Vektoren v_1, \dots, v_n ein Erzeugendensystem von V bilden. Also ist T genau dann ein Vektorraum-Isomorphismus wenn die Vektoren v_1, \dots, v_n eine Basis von V bilden. Jeder Vektorraum-Isomorphismus $T : \mathbb{R}^n \rightarrow V$ hat diese Form mit $v_i := Te_i$.

Satz D.2.9. Sei V ein reeller Vektorraum mit einer Basis e_1, \dots, e_n . Seien v_1, \dots, v_m Vektoren in V . Dann gilt folgendes.

- (i) Ist v_1, \dots, v_m eine Basis von V so gilt $m = n$.
- (ii) Bilden die Vektoren v_1, \dots, v_m ein Erzeugendensystem so gilt $m \geq n$.
- (iii) Sind die Vektoren v_1, \dots, v_m linear unabhängig so gilt $m \leq n$.

Die Zahl n heißt **Dimension** von V und wird mit $\dim V$ bezeichnet.

Beweis. Nach der vorherigen Bemerkung bestimmt die Basis e_1, \dots, e_n von V einen Isomorphismus $T : \mathbb{R}^n \rightarrow V$. Daher genügt es, die Behauptungen für den Fall $V = \mathbb{R}^n$ mit der Standardbasis e_1, \dots, e_n zu beweisen. Seien also $v_1, \dots, v_m \in \mathbb{R}^m$.

Schritt 1. Bilden die Vektoren v_1, \dots, v_m eine Basis des \mathbb{R}^n so ist $m = n$.

Wir beweisen die Aussage durch vollständige Induktion. Für $n = 1$ folgt die Behauptung aus den folgenden elementaren Fakten.

- (a) Jede von Null verschiedene reelle Zahl bildet eine Basis von \mathbb{R} .
- (b) Jedes System von mehr als einem Element von \mathbb{R} ist linear abhängig.

Sei nun $n \geq 2$. Wir nehmen an, dass das Resultat für den \mathbb{R}^{n-1} gilt und dass die Vektoren v_1, \dots, v_m eine Basis des \mathbb{R}^n bilden. Wir schreiben $v_\nu =: (v_{\nu 1}, \dots, v_{\nu n})$ für $\nu = 1, \dots, m$. Da der Vektor e_n sich als Linearkombination der Vektoren v_1, \dots, v_m schreiben läßt, gibt es mindestens ein $\nu \in \{1, \dots, m\}$ so dass $v_{\nu n} \neq 0$. Wir können o.B.d.A. annehmen dass dies $\nu = m$ ist. Sei

$$w_\nu := v_\nu - \frac{v_{\nu n}}{v_{mn}} v_m, \quad \nu = 1, \dots, m-1.$$

Dann bilden die Vektoren w_1, \dots, w_{m-1}, v_m eine Basis des \mathbb{R}^n und es gilt $w_{\nu n} = 0$ ist für $\nu = 1, \dots, m-1$. Es folgt dass die Vektoren w_1, \dots, w_{m-1} (nach Fortlassen der letzten Koordinate) eine Basis des \mathbb{R}^{n-1} bilden. Nach Induktionsannahme gilt also $m-1 = n-1$ und daher $m = n$.

Schritt 2. Sind die Vektoren v_1, \dots, v_m linear unabhängig so gilt $m \leq n$.

Wir können aus den Vektoren v_1, \dots, v_m durch hinzufügen von höchstens n weiteren Vektoren e_{i_1}, \dots, e_{i_k} (die wir aus den Elementen der Standardbasis auswählen) eine Basis $v_1, \dots, v_m, e_{i_1}, \dots, e_{i_k}$ bilden. (Dies folgt durch Induktion: Wenn die Vektoren v_1, \dots, v_m keine Basis bilden, so gibt es einen Index $i_1 \in \{1, \dots, n\}$, so dass die Vektoren v_1, \dots, v_m, e_{i_1} linear unabhängig sind. Wenn diese Vektoren immer noch keine Basis bilden, können wir weitere Elemente der Standardbasis hinzufügen ohne die lineare Unabhängigkeit zu verletzen.) Die neue Basis besteht aus $m+k$ Elementen mit $k \geq 0$. Es gilt daher $m+k = n$, nach Schritt 1, und deshalb $m \leq n$.

Schritt 3. Bilden die Vektoren v_1, \dots, v_m ein Erzeugendensystem so gilt $m \geq n$.

Wir erhalten eine Basis des \mathbb{R}^n durch Fortlassen von geeigneten Vektoren. Die resultierende Basis hat dann höchstens m Elemente. Also folgt die Behauptung aus Schritt 1. \square

D.3 Die Determinante

Sei V ein reeller Vektorraum. Eine Abbildung

$$\phi : V^n = V \times V \times \dots \times V \rightarrow \mathbb{R}$$

heißt **multilinear** wenn ϕ in jeder Variablen linear ist, das heißt es gelten die Gleichungen

$$\begin{aligned} & \phi(v_1, \dots, v_{i-1}, v_i + w_i, v_{i+1}, \dots, v_n) \\ &= \phi(v_1, \dots, v_n) + \phi(v_1, \dots, v_{i-1}, w_i, v_{i+1}, \dots, v_n) \end{aligned}$$

und

$$\phi(v_1, \dots, v_{i-1}, \lambda v_i, v_{i+1}, \dots, v_n) = \lambda \phi(v_1, \dots, v_n)$$

für alle $i \in \{1, \dots, n\}$, $v_1, \dots, v_n \in V$, $w_i \in V$, und $\lambda \in \mathbb{R}$. Im Fall $n = 2$ nennen wir eine solche Abbildung **bilinear**. Im Allgemeinen nennen wir eine multilineare Abbildung $\phi : V^n \rightarrow \mathbb{R}$ auch **n -linear**.

Bemerkung D.3.1. Der Raum V^n ist selbst ein Vektorraum. Es macht also auch Sinn, von linearen Abbildungen $\phi : V^n \rightarrow \mathbb{R}$ zu sprechen. Dies ist aber etwas völlig anderes als eine multilineare Abbildung. Betrachten wir zum Beispiel den Fall $V = \mathbb{R}$ und $n = 2$. Dann ist die Abbildung

$$\mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R} : (x, y) \mapsto xy$$

bilinear, aber nicht linear. Hingegen ist die Abbildung

$$\mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R} : (x, y) \mapsto x + y$$

linear, aber nicht bilinear. (Ist $\phi : V^n \rightarrow \mathbb{R}$ eine multilineare Abbildung so gilt $\phi(v_1, \dots, v_n) = 0$ sobald einer der Vektoren v_1, \dots, v_n gleich Null ist.)

Übung D.3.2. Sei V ein reeller Vektorraum und sei $\phi : V^n \rightarrow \mathbb{R}$ sowohl linear als auch multilinear. Ist $n \geq 2$, so ist $\phi \equiv 0$.

Satz D.3.3. Für jedes $n \in \mathbb{N}$ gibt es genau eine n -lineare Abbildung

$$\det : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \rightarrow \mathbb{R},$$

genannt **Determinante**, mit folgenden Eigenschaften.

(i) Für jede Permutation $\sigma \in S_n$ und je n Vektoren $a_1, \dots, a_n \in \mathbb{R}^n$ gilt

$$\det(a_{\sigma(1)}, \dots, a_{\sigma(n)}) = \varepsilon(\sigma) \det(a_1, \dots, a_n).$$

(ii) Ist e_1, \dots, e_n die Standardbasis des \mathbb{R}^n so gilt

$$\det(e_1, \dots, e_n) = 1.$$

Schreiben wir

$$a_i = (a_{i1}, \dots, a_{in}) \in \mathbb{R}^n, \quad i = 1, \dots, n,$$

mit $a_{ij} \in \mathbb{R}$, so ist die Determinante der Vektoren a_1, \dots, a_n durch die Formel

$$\det(a_1, \dots, a_n) := \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} = \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^n a_{i\sigma(i)} \quad (\text{D.3.1})$$

gegeben. Hier bezeichnet $\varepsilon(\sigma) \in \{\pm 1\}$ die durch (D.1.1) und (D.1.2) definierte Parität der Permutation $\sigma \in S_n$.

Beweis. Wir nehmen zunächst an, dass die Abbildung $\det : (\mathbb{R}^n)^n \rightarrow \mathbb{R}$ durch (D.3.1) definiert ist und zeigen, dass sie die Bedingungen (i) und (ii) erfüllt. Ist $a_i = e_i$ so gilt

$$a_{ij} = \delta_{ij} = \begin{cases} 1, & \text{falls } i = j, \\ 0, & \text{falls } i \neq j. \end{cases}$$

Damit ist der einzige von Null verschiedene Summand auf der rechten Seite von (D.3.1) der zu $\sigma = \text{id}$ gehörige und daher gilt (ii). Außerdem gilt

$$\begin{aligned} \det(a_{\tau(1)}, \dots, a_{\tau(n)}) &= \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^n a_{\tau(i)\sigma(i)} \\ &= \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{j=1}^n a_{j\sigma(\tau^{-1}(j))} \\ &= \varepsilon(\tau) \sum_{\sigma \in S_n} \varepsilon(\sigma \circ \tau^{-1}) \prod_{j=1}^n a_{j\sigma \circ \tau^{-1}(j)} \\ &= \varepsilon(\tau) \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{j=1}^n a_{j\sigma(j)} \\ &= \varepsilon(\tau) \det(a_1, \dots, a_n) \end{aligned}$$

für $\tau \in S_n$ und $a_i = (a_{i1}, \dots, a_{in}) \in \mathbb{R}^n$. Also erfüllt \det auch Bedingung (i).

Wir beweisen die Eindeutigkeit. Sei also $\det : (\mathbb{R}^n)^n \rightarrow \mathbb{R}$ durch (D.3.1) gegeben und sei $\phi : (\mathbb{R}^n)^n \rightarrow \mathbb{R}$ eine beliebige n -lineare Abbildung, welche die Bedingungen (i) und (ii) erfüllt. Dann ist die Differenz

$$\psi := \phi - \det : (R^n)^n \rightarrow \mathbb{R}$$

eine n -lineare Abbildung, welche die Bedingungen (i) und $\psi(e_1, \dots, e_n) = 0$ erfüllt. Seien $a_i = (a_{i1}, \dots, a_{in}) = \sum_{j=1}^n a_{ij} e_j \in \mathbb{R}^n$ gegeben für $i = 1, \dots, n$. Dann gilt

$$\begin{aligned} \psi(a_1, \dots, a_n) &= \psi \left(\sum_{j_1=1}^n a_{1j_1} e_{j_1}, \sum_{j_2=1}^n a_{2j_2} e_{j_2}, \dots, \sum_{j_n=1}^n a_{nj_n} e_{j_n} \right) \\ &= \sum_{j_1=1}^n \sum_{j_2=1}^n \cdots \sum_{j_n=1}^n a_{1j_1} a_{2j_2} \cdots a_{nj_n} \psi(e_{j_1}, \dots, e_{j_n}) = 0. \end{aligned}$$

Hier benutzen wir die Gleichung. $\psi(e_{j_1}, \dots, e_{j_n}) = 0$ für alle j_1, \dots, j_n . Also ist $\psi \equiv 0$ und damit $\phi \equiv \det$. Damit ist Satz D.3.3 bewiesen. \square

Eine n -lineare Abbildung $\phi : V \times \cdots \times V \rightarrow \mathbb{R}$, heißt **alternierende n -Form** (auf V) wenn sie der Bedingung (i) in Satz D.3.3 genügt. Es gibt also auf dem \mathbb{R}^n genau eine alternierende n -Form

$$\det : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \rightarrow \mathbb{R},$$

die die Bedingung $\det(e_1, \dots, e_n) = 1$ erfüllt, und diese heißt **Determinante**. Betrachten wir die n Vektoren

$$a_i =: (a_{i1}, \dots, a_{in}) \in \mathbb{R}^n$$

für $i = 1, \dots, n$ als Zeilen einer Matrix

$$A = (a_{ij})_{i,j=1}^n \in \mathbb{R}^{n \times n}$$

so können wir die Determinante auch als Abbildung

$$\det : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$$

auf dem Raum der quadratischen Matrizen auffassen.

Beispiel D.3.4. Ist $n = 1$ so ist $\mathbb{R}^{1 \times 1} = \mathbb{R}$ und $\det : \mathbb{R}^{1 \times 1} \rightarrow \mathbb{R}$ is die Identität. Für $n = 2$ haben wir

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

Für $n = 3$ und

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

ist

$$\begin{aligned} \det(A) &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &\quad - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}. \end{aligned}$$

Der folgende Satz faßt einige wichtige Eigenschaften der Determinante zusammen.

Satz D.3.5. (i) Für jede Matrix $A \in \mathbb{R}^{n \times n}$ gilt

$$\det(A^T) = \det(A). \quad (\text{D.3.2})$$

(ii) Für alle $A, B \in \mathbb{R}^{n \times n}$ gilt

$$\det(AB) = \det(A) \det(B), \quad \det(\mathbf{1}) = 1. \quad (\text{D.3.3})$$

(iii) Für alle $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{m \times m}$ gilt

$$\det \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = \det(A) \det(C). \quad (\text{D.3.4})$$

(iv) Sei $A = (a_{ij})_{i,j=1,\dots,n} \in \mathbb{R}^{n \times n}$ und für $k = 1, \dots, n$ sei

$$A_k := (a_{ij})_{\substack{i \in \{1, \dots, n\} \setminus \{k\} \\ j=2, \dots, n}} \in \mathbb{R}^{(n-1) \times (n-1)}$$

die Matrix, die aus A durch das Fortlassen der ersten Spalte und der k -ten Zeile hervorgeht. Dann gilt

$$\det(A) = \sum_{k=1}^n (-1)^{k-1} a_{k1} \det(A_k). \quad (\text{D.3.5})$$

Beweis. Wir beweisen Teil (i). Sei $A = (a_{ij})_{i,j=1}^n$ und

$$A^T := (a_{ji})_{i,j=1}^n$$

die transponierte Matrix. Dann gilt

$$\begin{aligned} \det(A^T) &= \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n} \\ &= \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma^{-1}(1)} \cdots a_{n\sigma^{-1}(n)} \\ &= \sum_{\tau \in S_n} \varepsilon(\tau) a_{1\tau(1)} \cdots a_{n\tau(n)} = \det(A). \end{aligned}$$

Hier folgt die vorletzte Gleichung aus der Umbenennung $\tau := \sigma^{-1}$ und der Tatsache dass $\varepsilon(\sigma) = \varepsilon(\sigma^{-1})$ ist. Damit ist Teil (i) bewiesen.

Wir beweisen Teil (ii). Seien

$$A = (a_{ij})_{i,j=1}^n, \quad B = (b_{jk})_{j,k=1}^n$$

zwei Matrizen in $\mathbb{R}^{n \times n}$. Den durch die Einträge der j te Zeile der Matrix B definierten Vektor bezeichnen wir mit

$$b_j := (b_{j1}, \dots, b_{jn}) \in \mathbb{R}^n.$$

Dann gilt nach (D.3.1)

$$\det(b_{j_1}, \dots, b_{j_n}) = \sum_{\tau \in S_n} \varepsilon(\tau) b_{j_1 \tau(1)} \cdots b_{j_n \tau(n)}$$

für $j_1, \dots, j_n \in \{1, \dots, n\}$. Nach Definition der Determinante verschwindet dieser Ausdruck sobald zwei der Indizes j_i übereinstimmen. (Siehe Bedingung (i) in Satz D.3.3.) Damit erfüllen A und B die Gleichung

$$\begin{aligned} \det(AB) &= \sum_{\tau \in S_n} \varepsilon(\tau) \left(\sum_{j_1=1}^n a_{1j_1} b_{j_1 \tau(1)} \right) \cdots \left(\sum_{j_n=1}^n a_{nj_n} b_{j_n \tau(n)} \right) \\ &= \sum_{j_1=1}^n \cdots \sum_{j_n=1}^n \sum_{\tau \in S_n} \varepsilon(\tau) a_{1j_1} b_{j_1 \tau(1)} \cdots a_{nj_n} b_{j_n \tau(n)} \\ &= \sum_{j_1=1}^n \cdots \sum_{j_n=1}^n a_{1j_1} \cdots a_{nj_n} \sum_{\tau \in S_n} \varepsilon(\tau) b_{j_1 \tau(1)} \cdots b_{j_n \tau(n)} \\ &= \sum_{j_1=1}^n \cdots \sum_{j_n=1}^n a_{1j_1} \cdots a_{nj_n} \det(b_{j_1}, \dots, b_{j_n}) \\ &= \sum_{\sigma \in S_n} a_{1\sigma(1)} \cdots a_{n\sigma(n)} \det(b_{\sigma(1)}, \dots, b_{\sigma(n)}) \\ &= \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} \det(b_1, \dots, b_n) \\ &= \det(A) \det(B). \end{aligned}$$

Hier benutzen wir die Tatsache, dass in der vierten Zeile nur diejenigen Summanden von Null verschieden sein können für die j_1, \dots, j_n paarweise verschieden sind. In diesem Fall definieren die Indizes eine Permutation $\sigma \in S_n$ durch die Formel $\sigma(i) := j_i$ für $i = 1, \dots, n$. Damit ist Teil (ii) bewiesen.

Wir beweisen Teil (iii). Jeder Summand in (D.3.1) enthält genau einen Faktor aus jeder Zeile und aus jeder Spalte von A . Für die Matrix

$$\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$$

heißt dies, dass jeder Summand, der einen Faktor aus dem oberen rechten Block B enthält, auch einen Faktor aus dem unteren linken Block enthalten muß und damit verschwindet. Anders gesagt, jede Permutation $\rho \in S_{n+m}$, die ein Element $i \in \{1, \dots, n\}$ auf ein $\rho(i) \in \{n+1, \dots, n+m\}$ abbildet, muß auch ein Element $j \in \{n+1, \dots, n+m\}$ auf ein $\rho(j) \in \{1, \dots, n\}$ abbilden. Damit brauchen wir nur Permutationen $\rho \in S_{n+m}$ zu betrachten, welche die beiden Teilmengen $\{1, \dots, n\}$ und $\{n+1, \dots, n+m\}$ auf sich selbst abbilden. Jede solche Permutation hat die Form $\rho = \sigma \# \tau$ für $\sigma \in S_n$ und $\tau \in S_m$, mit $\sigma \# \tau(i) := \sigma(i)$ für $1 \leq i \leq n$ und $\sigma \# \tau(n+j) = n + \tau(j)$ für $1 \leq j \leq m$. Diese Permutation hat die Parität $\varepsilon(\sigma \# \tau) = \varepsilon(\sigma)\varepsilon(\tau)$. Daraus folgt sofort Teil (iii).

Wir beweisen Teil (iv). Wir nehmen zunächst an, dass ein $k \in \{1, \dots, n\}$ existiert, so dass $a_{i1} = 0$ ist für alle $i \in \{1, \dots, n\} \setminus \{k\}$. Unter dieser Annahme definieren die Permutation $\sigma_k \in S_n$ durch

$$\sigma_k(i) := \begin{cases} k, & \text{für } i = 1, \\ i - 1, & \text{für } i = 1, \dots, k - 1, \\ i, & \text{für } i = k, \dots, n. \end{cases}$$

Diese Permutation hat die Parität

$$\varepsilon(\sigma_k) = (-1)^{k-1} \tag{D.3.6}$$

und es gilt

$$B := (a_{\sigma_k(i)j})_{i,j=1,\dots,n} = \begin{pmatrix} a_{k1} & * \\ 0 & A_k \end{pmatrix} \tag{D.3.7}$$

Hier steht das Symbol $*$ für die restlichen Einträge der k -ten Zeile der Matrix A . Aus (D.3.6) und (D.3.7) folgt die Gleichung

$$\det(A) = \varepsilon(\sigma_k) \det(B) = (-1)^{k-1} \det(B) = (-1)^{k-1} a_{k1} \det(A_k).$$

Hier folgt die erste Gleichheit aus Teil (i), die zweite Gleichheit aus (D.3.6), und die dritte Gleichheit aus (D.3.7) und Teil (iii). Nun folgt Teil (iv) aus der Multilinearität der Determinante. Damit ist Satz D.3.5 bewiesen. \square

Übung D.3.6. Die Determinante ist durch die Eigenschaften (ii) und (iii) in Satz D.3.5 eindeutig bestimmt.

Satz D.3.7. Für jede Matrix $A \in \mathbb{R}^{n \times n}$ sind folgende Aussagen äquivalent.

(i) $\det(A) \neq 0$.

(ii) Für alle $x \in \mathbb{R}^n$ gilt $Ax = 0 \implies x = 0$.

(iii) Für jedes $y \in \mathbb{R}^n$ gibt es ein $x \in \mathbb{R}^n$ so dass $Ax = y$.

Beweis. Nach Satz D.2.9 ist ein System von n Vektoren $a_1, \dots, a_n \in \mathbb{R}^n$ genau dann linear unabhängig wenn es ein Erzeugendensystem vom \mathbb{R}^n ist. Angewandt auf die n Spalten der Matrix A heißt dies, dass (ii) äquivalent zu (iii) ist. Es bleibt also zu zeigen, dass (i) äquivalent zu (ii) ist.

Wir zeigen, dass (ii) aus (i) folgt. Wir argumentieren indirekt und nehmen an dass es einen Vektor $x = (x_1, \dots, x_n) \in \mathbb{R}^n \setminus \{0\}$ gibt so dass $Ax = 0$ ist. Bezeichnen wir die Spalten der Matrix A mit $a_1, \dots, a_n \in \mathbb{R}^n$, so heißt dies

$$x_1 a_1 + \dots + x_n a_n = 0.$$

Da $x \neq 0$ ist, gibt es ein $i \in \{1, \dots, n\}$ mit $x_i \neq 0$. Ohne Beschränkung der Allgemeinheit nehmen wir an, dass $x_1 \neq 0$ ist. Nach Satz D.3.5 (i) gilt dann

$$\begin{aligned} \det(A) &= x_1^{-1} \det(x_1 a_1, a_2, \dots, a_n) \\ &= x_1^{-1} \sum_{i=1}^n \det(x_i a_i, a_2, \dots, a_n) \\ &= x_1^{-1} \det\left(\sum_{i=1}^n x_i a_i, a_2, \dots, a_n\right) \\ &= x_1^{-1} \det(0, a_2, \dots, a_n) \\ &= 0. \end{aligned}$$

Wir zeigen, dass (i) aus (ii) folgt. Wenn (ii) erfüllt ist, so sind die Spalten a_1, \dots, a_n von A linear unabhängig und bilden damit, nach Satz D.2.9, eine Basis. Hieraus folgt, dass die durch

$$T_A(x) := Ax = \sum_{i=1}^n x_i a_i \quad \text{für } x = (x_1, \dots, x_n) \in \mathbb{R}^n$$

definierte lineare Abbildung $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ bijektiv ist. Sei $S : \mathbb{R}^n \rightarrow \mathbb{R}^n$ die Umkehrabbildung, das heißt $T_A \circ S = S \circ T_A = \text{id}$. Dann ist S linear und daher existiert eine Matrix $B \in \mathbb{R}^{n \times n}$ so dass $S = T_B$ ist. Hieraus folgt

$$AB = BA = \mathbb{1} \tag{D.3.8}$$

und daher, nach Satz D.3.5 (ii), $\det(A) \det(B) = 1$. Also ist $\det(A) \neq 0$. \square

Definition D.3.8. Der Beweis von Satz D.3.7 zeigt dass es zu jeder $(n \times n)$ -Matrix A mit $\det(A) \neq 0$ eine $(n \times n)$ -Matrix B gibt, welche die Gleichung (D.3.8) erfüllt. Diese Matrix ist durch A eindeutig bestimmt; sie heißt **Inverse von A** , beziehungsweise **inverse Matrix von A** , und wird mit A^{-1} bezeichnet. Nach Satz D.3.5 gilt $\det(A^{-1}) = \det(A)^{-1}$.

Bemerkung D.3.9. Alle Definitionen und Resultate dieses Kapitels gelten unverändert und mit denselben Beweisen für komplexe Matrizen. Es genügt, überall \mathbb{R} durch \mathbb{C} zu ersetzen.

Definition D.3.10. Sei $A \in \mathbb{R}^{n \times n}$ und sei $\lambda \in \mathbb{C}$. Die komplexe Zahl λ heißt **Eigenwert** der reellen Matrix A , wenn eine komplexer Vektor $x \in \mathbb{C}^n$ existiert, der den Bedingungen

$$Ax = \lambda x, \quad x \neq 0,$$

genügt. Nach Satz D.3.7 ist die Zahl λ genau dann ein Eigenwert von A wenn sie die Gleichung $\det(\lambda \mathbb{1} - A) = 0$ erfüllt. Das durch die Formel

$$p_A(z) := \det(z\mathbb{1} - A)$$

für $z \in \mathbb{C}$ definierte Polynom $p_A : \mathbb{C} \rightarrow \mathbb{C}$ heißt **charakteristisches Polynom der Matrix A** .

Übung D.3.11. Das charakteristische Polynom einer Matrix $A \in \mathbb{R}^{n \times n}$ hat den Grad n . **Hinweis:** Gleichung (D.3.1) für die Determinante.

Übung D.3.12. Ist $\lambda \in \mathbb{C}$ ein Eigenwert einer Matrix $A \in \mathbb{R}^{n \times n}$ so ist auch $\bar{\lambda}$ ein Eigenwert von A .

Korollar D.3.13. Die Determinante einer Matrix $A \in \mathbb{R}^{n \times n}$ ist das Produkt ihrer Eigenwerte mit Multiplizität, das heißt, wenn das charakteristische Polynom in der Form

$$p_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)$$

mit $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{C}$ geschrieben wird, so gilt

$$\det(A) = \lambda_1 \lambda_2 \cdots \lambda_n.$$

Beweis. Nach Definition gilt $p_A(0) = (-1)^n \det(A)$ und hieraus folgt sofort die Behauptung. \square

Beispiel D.3.14. Seien $a_1, \dots, a_n \in \mathbb{R}$ und sei $A \in \mathbb{R}^{(n+1) \times n}$ die Matrix

$$A := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \\ a_1 & \cdots & \cdots & a_n \end{pmatrix}. \quad (\text{D.3.9})$$

Dann ist die Determinante der quadratischen Matrix $A^T A \in \mathbb{R}^{n \times n}$ durch

$$\det(A^T A) = 1 + a_1^2 + a_2^2 + \cdots + a_n^2 \quad (\text{D.3.10})$$

gegeben. **Beweis:** Seien $P, Q \in \mathbb{R}^{(n+1) \times (n+1)}$ die Matrizen

$$P := \begin{pmatrix} 1 & 0 & \cdots & 0 & a_1 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & a_n \\ -a_1 & \cdots & \cdots & -a_n & 1 \end{pmatrix},$$

$$Q := \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & 0 \\ a_1 & \cdots & \cdots & a_n & 1 \end{pmatrix}.$$

Dann ist $\det(Q) = 1$ und $PQ \in \mathbb{R}^{(n+1) \times (n+1)}$ ist die Block-Matrix

$$PQ = \begin{pmatrix} A^T A & a \\ 0 & 1 \end{pmatrix}, \quad a := \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

Nach Teil (ii) und (iii) von Satz D.3.5 gilt daher

$$\det(A^T A) = \det(PQ) = \det(P) = 1 + a_1^2 + \cdots + a_n^2.$$

Hier zeigt man die letzte Gleichheit durch elementare Zeilenoperationen, indem man für $i = 1, \dots, n$ das Produkt der i -ten Zeile von P mit a_i zur letzten Zeile addiert, und dann nochmals Teil (iii) von Satz D.3.5 anwendet.

D.4 Die allgemeine lineare Gruppe

Die Gruppe der invertierbaren reellen $(n \times n)$ -Matrizen heißt **allgemeine lineare Gruppe** und wird mit

$$\mathrm{GL}(n, \mathbb{R}) := \{A \in \mathbb{R}^{n \times n} \mid \det(A) \neq 0\}$$

bezeichnet. Daß dies eine Gruppe ist folgt aus Satz D.3.5 und Satz D.3.7. Die Gruppenoperation ist Matrix-Multiplikation und das neutrale Element ist die Einheits-Matrix $\mathbb{1}$. Die Gruppe $\mathrm{GL}(n, \mathbb{R})$ kann mit der Gruppe der bijektiven linearen Abbildungen $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ identifiziert werden; solche Abbildungen werden auch **Automorphismen** des \mathbb{R}^n genannt. Die Korrespondenz ordnet jeder Matrix $A \in \mathrm{GL}(n, \mathbb{R})$ den durch $T_A(x) := Ax$ definierten Automorphismus des \mathbb{R}^n zu. Satz D.3.7 zeigt, dass die lineare Abbildung $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ genau dann bijektiv ist wenn $\det(A) \neq 0$ ist. Nach Satz D.3.5 ist

$$\mathrm{GL}^+(n, \mathbb{R}) := \{A \in \mathrm{GL}(n, \mathbb{R}) \mid \det(A) > 0\}$$

eine Untergruppe von $\mathrm{GL}(n, \mathbb{R})$. Die Abbildung

$$\det : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$$

ist stetig. Daher sind $\mathrm{GL}(n, \mathbb{R})$ und $\mathrm{GL}^+(n, \mathbb{R})$ offene Teilmengen von $\mathbb{R}^{n \times n}$ (dies sind die Urbilder der offenen Teilmengen $\mathbb{R} \setminus \{0\}$ und $(0, \infty)$ von \mathbb{R} unter der stetigen Abbildung $\det : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$). Die Restriktion der Abbildung $\det : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ auf $\mathrm{GL}(n, \mathbb{R})$ definiert einen Gruppenhomomorphismus von $\mathrm{GL}(n, \mathbb{R})$ auf die von Null verschiedenen reellen Zahlen (siehe Beispiel D.1.5 in Abschnitt D.1).

Für den folgenden Satz benötigen wir den Begriff eines weg-zusammenhängenden metrischen Raumes, der im Anhang B ausführlicher diskutiert wird. Hier sei nur erwähnt, dass ein metrischer Raum (X, d) **weg-zusammenhängend** genannt wird, wenn für alle $x_0, x_1 \in X$ eine stetige Abbildung $\gamma : [0, 1] \rightarrow X$ existiert, welche die Punkte $\gamma(0) = x_0$ und $\gamma(1) = x_1$ miteinander verbindet.

Satz D.4.1. *Die Gruppe $\mathrm{GL}^+(n, \mathbb{R})$ ist weg-zusammenhängend.*

Beweis. Siehe Seite 363. □

Der hier erbrachte Beweis beruht auf den folgenden drei Lemmata. Er geht auf Thomas Honold zurück und ist direkt aus [2, Seite 36] übernommen.

Lemma D.4.2. Für jede Matrix $A \in \text{GL}(n, \mathbb{R})$ existiert ein $\ell \in \{0, 1, \dots, n\}$, eine Matrix $U \in \text{GL}^+(n - \ell, \mathbb{R})$ ohne negative Eigenwerte, eine Matrix

$$\Lambda = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & \lambda_\ell \end{pmatrix} \in \text{GL}(\ell, \mathbb{R}), \quad \lambda_i < 0,$$

und eine Matrix $P \in \text{GL}(n, \mathbb{R})$, welche die Gleichung

$$P^{-1}AP = \begin{pmatrix} \Lambda & * \\ 0 & U \end{pmatrix}$$

erfüllen

Beweis. Wir beweisen dieses Lemma durch vollständige Induktion über n . Für $n = 1$ ist die Aussage offensichtlich wahr mit $P = 1$ und entweder $\ell = 0$, $U = A$ (im Falle $A > 0$) oder $\ell = 1$, $\Lambda = A$ (im Falle $A < 0$).

Sei nun $n > 1$ und nehmen wir an, die Behauptung sei für alle Matrizen in $\text{GL}(n - 1, \mathbb{R})$ bewiesen. Sei $A \in \text{GL}(n, \mathbb{R})$ gegeben. Hat A keine negativen Eigenwerte, so gilt die Behauptung mit $P := \mathbb{1}$, $\ell := 0$, und $U := A$. Hat A einen negativen Eigenwert $\lambda < 0$ so gibt es einen Vektor $v \in \mathbb{R}^n$ so dass

$$Av = \lambda v, \quad v \neq 0.$$

Sei $Q \in \mathbb{R}^{n \times n}$ irgendeine invertierbare Matrix deren erste Spalte gleich v ist. Dann gibt es eine Matrix $A_0 \in \text{GL}(n - 1, \mathbb{R})$ so dass

$$Q^{-1}AQ = \begin{pmatrix} \lambda & * \\ 0 & A_0 \end{pmatrix}.$$

Nach Induktionsannahme gibt es Matrizen $P_0 \in \text{GL}(n - 1, \mathbb{R})$ und Λ_0, U_0 wie in der Behauptung so dass

$$P_0^{-1}A_0P_0 = \begin{pmatrix} \Lambda_0 & * \\ 0 & U_0 \end{pmatrix}.$$

Also gilt die Behauptung für A mit

$$P := Q \begin{pmatrix} 1 & 0 \\ 0 & P_0 \end{pmatrix}, \quad \Lambda := \begin{pmatrix} \lambda & * \\ 0 & \Lambda_0 \end{pmatrix}, \quad U := U_0.$$

Damit ist das Lemma bewiesen. \square

Lemma D.4.3. Sei $A \in \text{GL}^+(n, \mathbb{R})$ gegeben. Dann existieren es zwei Matrizen $B, C \in \text{GL}^+(n, \mathbb{R})$, welche keine negativen Eigenwerte haben und die Gleichung

$$A = B^2C$$

erfüllen.

Beweis. Seien $P \in \text{GL}(n, \mathbb{R})$, $\Lambda \in \text{GL}(\ell, \mathbb{R})$, und $U \in \text{GL}(n - \ell, \mathbb{R})$ wie in Lemma D.4.2. Nach Satz D.3.5 gilt

$$\det(A) = \det(\Lambda) \det(U) = \lambda_1 \cdots \lambda_\ell \det(U).$$

Da U keine negativen Eigenwerte hat ist $\det(U) > 0$. Da $\det(A) > 0$ ist, ist also $\ell = 2k$ eine gerade Zahl. Wir definieren

$$S := \begin{pmatrix} H & 0 & \cdots & 0 & 0 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & H & 0 \\ 0 & \cdots & \cdots & 0 & \mathbb{1}_{n-2k} \end{pmatrix}, \quad H := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Diese Matrix hat keine negativen Eigenwerte (die Eigenwerte sind $\pm i$ und 1). Außerdem gilt

$$S^{-2} = S^2 = \begin{pmatrix} -\mathbb{1}_{2k} & 0 \\ 0 & \mathbb{1}_{n-2k} \end{pmatrix}$$

und damit

$$T := S^{-2}P^{-1}AP = \begin{pmatrix} -\mathbb{1}_{2k} & 0 \\ 0 & \mathbb{1}_{n-2k} \end{pmatrix} \begin{pmatrix} \Lambda & * \\ 0 & U \end{pmatrix} = \begin{pmatrix} -\Lambda & * \\ 0 & U \end{pmatrix}.$$

Diese Matrix hat ebenfalls keine negativen Eigenwerte. Also gilt

$$A = P(S^2T)P^{-1} = B^2C,$$

wobei die Matrizen

$$B := PSP^{-1}, \quad C := PTP^{-1}$$

keine negativen Eigenwerte haben. Damit ist das Lemma D.4.3 bewiesen. \square

Lemma D.4.4. Sei $A \in \text{GL}^+(n, \mathbb{R})$ eine Matrix ohne negative Eigenwerte. Dann gilt $\det((1-t)\mathbb{1} + tA) > 0$ für alle $t \in [0, 1]$.

Beweis. Wir definieren die Funktion $f : [0, 1] \rightarrow \mathbb{R}$ durch

$$f(t) := \det((1-t)\mathbb{1} + tA).$$

Diese Funktion ist stetig und es gilt $f(0) > 0$ und $f(1) > 0$. Die Behauptung des Lemmas sagt dass $f(t) > 0$ ist für alle $t \in [0, 1]$. Wenn dies nicht gilt, gibt es eine Zahl $t \in [0, 1]$ mit $f(t) \leq 0$. Nach dem Zwischenwertsatz, gibt es dann auch eine Zahl $t_0 \in [0, 1]$ mit $f(t_0) = 0$. Diese Zahl kann nicht 0 oder 1 sein. Dann ist aber die Zahl

$$\lambda := -\frac{1-t_0}{t_0}$$

ein negativer Eigenwert von A , im Widerspruch zu unserer Voraussetzung. Dieser Widerspruch beweist Lemma D.4.4. \square

Beweis von Satz D.4.1. Sei $A \in \text{GL}^+(n, \mathbb{R})$ und wähle $B, C \in \text{GL}^+(n, \mathbb{R})$ wie in Lemma D.4.3. Für $t \in [0, 1]$ definiere

$$A_t := ((1-t)\mathbb{1} + tB)^2((1-t)\mathbb{1} + tC).$$

Dann gilt $A_0 = \mathbb{1}$ und $A_1 = A$. Da die Matrizen B und C keine negativen Eigenwerte haben, folgt aus Lemma D.4.4 dass

$$\det(A_t) = \det((1-t)\mathbb{1} + tB)^2 \det((1-t)\mathbb{1} + tC) > 0$$

für alle $t \in [0, 1]$. Also ist die Abbildung $[0, 1] \rightarrow \text{GL}^+(n, \mathbb{R}) : t \mapsto A_t$ ein stetiger Weg der die Matrizen $\mathbb{1}$ und A miteinander verbindet. Damit ist $\text{GL}^+(n, \mathbb{R})$ weg-zusammenhängend. \square

Übung D.4.5. Die Gruppe $\text{GL}(n, \mathbb{R})$ ist die disjunkte Vereinigung der weg-zusammenhängenden Teilmengen $\text{GL}^+(n, \mathbb{R})$ und

$$\text{GL}^-(n, \mathbb{R}) := \{A \in \text{GL}(n, \mathbb{R}) \mid \det(A) < 0\}.$$

Die Gruppe $\text{GL}(n, \mathbb{R})$ ist selbst nicht weg-zusammenhängend.

Übung D.4.6. Jede Matrix $A \in \mathbb{R}^{n \times n}$ erfüllt die Gleichung

$$\det(e^A) = e^{\text{spur}(A)}, \quad (\text{D.4.1})$$

wobei $\text{spur}(A)$ die **Spur** der Matrix A , das heißt die Summe der Diagonaleinträge, ist.

Anhang E

Der Banachsche Fixpunktsatz

Sei (X, d) ein metrischer Raum und sei $f : X \rightarrow X$ eine Abbildung von X auf sich selbst. Die Abbildung f heißt **Kontraktion**, wenn eine Konstante

$$0 \leq \alpha < 1$$

existiert, so dass die Ungleichung

$$d(f(x), f(y)) \leq \alpha d(x, y) \tag{E.0.1}$$

für alle $x, y \in X$ erfüllt ist. Ein **Fixpunkt von f** ist ein Element $x \in X$, welches die Gleichung $f(x) = x$ erfüllt.

Satz E.0.1 (Banachscher Fixpunktsatz). *Sei (X, d) ein nichtleerer vollständiger metrischer Raum und sei $f : X \rightarrow X$ eine Kontraktion. Dann besitzt f genau einen Fixpunkt.*

Beweis. Wir beweisen zunächst die Eindeutigkeit. Seien $x, y \in X$ Fixpunkte von f . Dann gilt $f(x) = x$ und $f(y) = y$ und daher

$$d(x, y) = d(f(x), f(y)) \leq \alpha d(x, y)$$

nach (E.0.1). Daraus folgt die Ungleichung

$$(1 - \alpha)d(x, y) \leq 0.$$

Da $\alpha < 1$ ist folgt daraus $d(x, y) = 0$ und daher $x = y$.

Zum Beweis der Existenz wählen wir zunächst ein Element $x_0 \in X$. Ein solches Element existiert, da X nichtleer ist. Nun definieren wir induktiv eine Folge $(x_n)_{n \in \mathbb{N}}$ in X durch

$$x_1 := f(x_0), \quad x_{n+1} := f(x_n) \quad (\text{E.0.2})$$

für $n \in \mathbb{N}$. Dann gilt

$$d(x_n, x_{n+1}) = d(f(x_{n-1}), f(x_n)) \leq \alpha d(x_{n-1}, x_n)$$

für alle $n \in \mathbb{N}$ nach (E.0.1). Daraus ergibt sich durch Vollständige Induktion die Ungleichung

$$d(x_n, x_{n+1}) \leq \alpha^n d(x_0, x_1)$$

für alle $n \in \mathbb{N}$ und daher, für je zwei ganze Zahlen $0 < n < m$,

$$\begin{aligned} d(x_n, x_m) &\leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \cdots + d(x_{m-1}, x_m) \\ &\leq (\alpha^n + \alpha^{n+1} + \cdots + \alpha^{m-1}) d(x_0, x_1) \\ &\leq \frac{\alpha^n}{1 - \alpha} d(x_0, x_1). \end{aligned}$$

Dies zeigt, dass die Folge $(x_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge ist. Da (X, d) vollständig ist, konvergiert also die Folge $(x_n)_{n \in \mathbb{N}}$ gegen ein Element $x \in X$, und es gilt

$$\begin{aligned} x &= \lim_{n \rightarrow \infty} x_n \\ &= \lim_{n \rightarrow \infty} x_{n+1} \\ &= \lim_{n \rightarrow \infty} f(x_n) \\ &= f\left(\lim_{n \rightarrow \infty} x_n\right) \\ &= f(x). \end{aligned}$$

Also ist x ein Fixpunkt von f und damit ist Satz E.0.1 bewiesen. \square

Literaturverzeichnis

- [1] T. Bühler, D.A. Salamon, *Functional Analysis*. ETHZ, 2016.
<https://people.math.ethz.ch/~salamon/PREPRINTS/funcana.pdf>
- [2] K. Königsberger, *Analysis 2*. 5. Auflage, Springer Verlag, 2003.
- [3] John Milnor, *Topology from the Differentiable Viewpoint*. Based on notes by D. Weaver. Revised reprint of the 1965 original. *Princeton Landmarks in Mathematics*, Princeton University Press.
- [4] D.A. Salamon, *Summierbare Familien*, ETH Zürich, 2016.
<http://www.math.ethz.ch/~salamon/PREPRINTS/ana1-sum.pdf>
- [5] D.A. Salamon, *Das Riemannsche Integral*, Analysis I, ETH Zürich, 2016.
<http://www.math.ethz.ch/~salamon/PREPRINTS/ana1-int.pdf>
- [6] D.A. Salamon, *Analysis I*, ETHZ, Herbstsemester 2014. Videolink:
<http://www.video.ethz.ch/lectures/d-math/2014/autumn/401-1261-07L.html>
- [7] D.A. Salamon, *Analysis II*, ETHZ, Frühjahrssemester 2015. Videolink:
<http://www.video.ethz.ch/lectures/d-math/2015/spring/401-1262-07L.html>

Index

- Abbildung
 - differenzierbar, 35
 - partiell differenzierbar, 39
 - zweimal, 67
 - stetig, 318
 - stetig differenzierbar, 45
 - ℓ -mal, 72
 - zweimal, 72
- abgeschlossene Menge, 318
- Ableitung, 36
 - partiell, 39
 - höhere, 72
 - zweite, 67
- Abschluss, 319
- absolut summierbar, 250
- Abstandsfunktion, 317
- analytische Funktion, 85
- Anfangswertproblem, 133
- Atlas, 116, 294
- äußerer Einheitsnormalenvektor, 303
- äußeres Einheitsnormalenfeld, 303

- Banachalgebra, 14
- Banachraum, 2
- Banachscher Fixpunktsatz, 365
- Basis, 348
- Brouwerscher Fixpunktsatz, 315

- Cauchy-Folge, 2

- Determinante, 351

- Diffeomorphiesatz, 108
- Diffeomorphismus, 100
 - volumentreu, 224
- Differentialform, 283
 - exakt, 289
 - geschlossen, 289
- differenzierbar, 35
 - partiell, 39
 - zweimal, 67
 - stetig, 45
 - ℓ -mal, 72
 - zweimal, 72
- diskrete
 - Teilmenge, 250
 - Teimenge, 117
 - Topologie, 250
- Divergenz, 227
- divergenzfrei, 227

- Einheitsnormalenfeld, 266
- Euler–Lagrange-Gleichung, 115
- Existenzintervall
 - maximal, 142
- Exponentialabbildung, 23, 156
- Exponentialfunktion, 345

- Fixpunkt, 365
- Fluss eines Vektorfeldes, 142
 - Differenzierbarkeit, 149
 - Stetigkeit, 145
- Funktion

- eigentlich, 197
- konvex, 94
- strikt konvex, 94
- Gauß-Abbildung, 266
- glatte Abbildung, 73, 100
- glatter Rand, 260, 303
- Gradient, 54
- Gronwall
 - Lemma von, 145
- Gruppe, 343
 - abelsch, 343
- Gruppenaktion, 345
- Hamiltonsche Gleichung, 115
- Harmonische Funktion, 271
- Hessematrix, 90
- Hilbertraum, 2
- Homomorphismus, 344, 347
- Inneres, 319
- inneres Produkt, 1
- Integral
 - über ein Kartengebiet, 238
 - über einem Quadergebäude, 163
 - über einer
 - Jordan-messbaren Menge, 185
 - über einer Mannigfaltigkeit, 246
 - Riemannsch, 163
- Isomorphismus, 344, 347
- Jacobi-Matrix, 44
- Jordan-messbar, 185
 - bezüglich M , 240
- Jordansche Nullmenge, 165
 - bezüglich M , 240
- Karte, 116
- Kartengebiet, 116, 236, 293
- Keplerproblem, 56
- Kettenregel, 49
- kompakte Menge, 334
- Kontraktion, 365
- konvexe Funktion, 94
- konvexe Menge, 94
- kritischer Punkt, 88
 - nichtdegeneriert, 93, 108, 113
- Länge einer Kurve, 230
- Lagrange-Multiplikator, 128
- Laplace-Operator, 271
- Legendre-Transformation, 113
- Leibnizregel, 49
- lineare Gruppe, 360
- lokales Maximum, 87
- lokales Minimum, 87
- Matrixnorm, 10, 16
- maximales Existenzintervall, 142
- metrischer Raum, 317
 - kompakt, 334
 - total beschränkt, 334
 - vollständig, 334
 - weg-zusammenhängend, 328
 - zusammenhängend, 325
- Multiindex, 73
 - Ordnung, 73
- Newtonsche Gleichung, 56, 115
- normale Ableitung, 271
- normierter Vektorraum, 1
- Nullmenge
 - Jordansch, 165
- offene Menge, 317
- offener Ball, 317, 334
- Offenheitssatz, 108
- Operatornorm, 8

- Parametrisierung
 - eines Kartengebiets, 236
- partiell differenzierbar, 39
 - zweimal, 67
- partielle Ableitung, 39
 - der Ordnung ℓ , 72
 - zweite, 67
- Partition, 158
- Partition der Eins, 243
- Permutation, 344
- Perron–Frobenius, 316
- Polynom, 76
- Potenzreihe
 - in einer Banachalgebra, 18
 - Komposition, 27
 - von Matrizen, 58
- Quadergebäude, 159
- Rand, 319
 - bezüglich M , 240
 - kanonische Orientierung, 303
- reelle Algebra, 14
- regulärer Wert, 120
- rektifizierbar, 230
- Relativtopologie, 323
- Retraktionssatz, 312
- Richtungsableitung, 39
- Riemann-integrierbar, 163
 - lokal, 218
 - uneigentlich, 218, 249
- Riemannsche Summe, 171
 - verallgemeinerte, 201
- Römerfläche, 124
- Satz von
 - Arzelà–Ascoli, 339
 - der impliziten Funktion, 111
 - der Umkehrabbildung, 103
 - Fubini, 180, 187
 - für Polarkoordinaten, 256
 - Gauß, 266
 - Perron–Frobenius, 316
 - Sard, 121, 193, 211
 - Schwarz, 69
 - Stokes, 306
 - klassisch, 310
- Schrankensatz, 60
- singulärer Wert, 120
- Spur, 363
- stetig, 318
 - gleichgradig, 338
- stetig differenzierbar, 45
 - ℓ -mal, 72
 - zweimal, 72
- strikt konvexe Funktion, 94
- Tangentialraum, 125
- Tangentialvektor, 125
- Taylorpolynom, 77
- Taylorreihe, 82
- Torus, 90
- Träger, 243
- Transformationsformel, 211
 - linear, 206
- Tubenumgebung, 132
- Übergangsabbildung, 294
- Untermannigfaltigkeit, 116, 293
 - orientiert, 296
- Vektorfeld, 133
 - divergenzfrei, 227
- Vektorraum, 347
 - normiert, 1
 - orientiert, 292
- vollständig, 2
- Volumen

- der Einheitssphäre S^{n-1} , 254, 257
- des Einheitsballes B^n , 197
- des Einheitssimplex Δ^n , 184
- einer Mannigfaltigkeit, 246
- eines Kartengebiets, 238
- eines Kegels, 188
- eines Quadergebäudes, 164
- eines Quaders, 157
- eines Zylinders, 188
- volumentreu, 224

- zusammenhängend, 325
- Zusammenhangskomponente, 330