Vorlesungsmanuskript zu

Analysis II

Werner Balser
Institut für Angewandte Analysis

Sommersemester 2009

Literaturverzeichnis

[1] **T. Apostol**, *Mathematical Analysis*, Addison & Wesley, Reading, 1979.

[2] **E. Behrends**, *Analysis. Vol. 2. A study book. (Analysis. Band 2. Ein Lernbuch.) 2nd revised ed.*, Wiesbaden: Vieweg. xiv, 376 p. EUR 24.90 , 2007.

[3] **I. N. Bronstein, K. A. Semendjajew, G. Musiol, und H. Mühlig**, *Taschenbuch der Mathematik*, Harri Deutsch, Frankfurt am Main, 1995. 80

[4] **K. Endl und W. Luh**, *Analysis II*, Aula Verlag, Wiesbaden, 1989.

[5] **F. Erwe**, *Differential- und Integralrechnung II*, BI Hochschultaschenbücher, Bibliographisches Institut, Mannheim, 1973.

[6] **H. Grauert und I. Lieb**, *Differential- und Integralrechnung II*, Heidelberger Taschenbücher, Springer, Berlin, 196.

[7] **H. Heuser**, *Lehrbuch der Analysis 2*, Teubner, Stuttgart, 1992. 60

[8] **W. Luh und M. W. 2**, *Aufgabensammlung Analysis*, Aula Verlag, Wiesbaden, 1992.

[9] **K. Meyberg und P. Vachenauer**, *Höhere Mathematik 1*, Springer, Berlin, 1999.

[10] ——, *Höhere Mathematik 2*, Springer, Berlin, 1999.

[11] **W. Walter**, *Analysis 2*, Springer, Berlin, 1992. 52, 63, 64, 84, 85

Inhaltsverzeichnis

Kapitel 1

Grundlagen

1.1 Ungleichungen

Die folgenden Ungleichungen spielen in der Mathematik eine zentrale Rolle:

Aufgabe 1.1.1 *Seien $p, q > 1$ mit $1/p + 1/q = 1$, und seien $a, b \in \mathbb{R}_+$. Zeige: Es ist stets*

$$a^{1/p}\, b^{1/q} \;\leq\; \frac{a}{p} \;+\; \frac{b}{q}\,,$$

und Gleichheit gilt genau dann, wenn $a = b$ ist. Anleitung: *Finde das Maximum der Funktion $f(x) = x^{1/p}\, b^{1/q}\,(x/p + b/q)^{-1}$, für $x \geq 0$.*

Lemma 1.1.2 (Höldersche Ungleichung) *Für $p, q > 1$ mit $1/p+1/q = 1$ sowie $a = (a_1, \ldots, a_n)^T, b = (b_1, \ldots, b_n)^T \in \mathbb{K}^n$ gilt*

$$\sum_{k=1}^{n} |a_k\, b_k| \;\leq\; \Big(\sum_{k=1}^{n} |a_k|^p\Big)^{1/p} \Big(\sum_{k=1}^{n} |b_k|^q\Big)^{1/q}.$$

Das Gleichheitszeichen gilt genau dann, wenn $\lambda, \mu \in \mathbb{R}$ existieren, welche nicht beide gleich 0 sind, mit $\lambda\, |a_k|^p = \mu\, |b_k|^q$ für alle k.

Beweis: O. B. d. A. seien $A = \sum_{k=1}^{n} |a_k|^p > 0$ und $B = \sum_{k=1}^{n} |b_k|^q > 0$. Aus Aufgabe 1.1.1 folgt dann

$$\frac{|a_k|}{A^{1/p}}\,\frac{|b_k|}{B^{1/q}} \;\leq\; \frac{|a_k|^p}{p\,A} \;+\; \frac{|b_k|^q}{q\,B} \qquad \forall\; k = 1, \ldots, n,$$

mit Gleichheit genau für $|a_k|^p/A = |b_k|^q/B$. Durch Summation über k folgt die Behauptung. □

Aufgabe 1.1.3 *Zeige mit der Hölderschen Ungleichung für $a_1, \ldots, a_n \in \mathbb{C}$ und $p, q > 1$ mit $1/p+1/q = 1$*

$$\sum_{k=1}^{n} |a_k| \;\leq n^{1/q} \left(\sum_{k=1}^{n} |a_k|^p\right)^{1/p},$$

und gib ein Beispiel, für welches das Gleichheitszeichen gilt.

Aufgabe 1.1.4 (Höldersche Ungleichung für Integrale) *Sei $a < b$, und seien $f, g : [a,b] \longrightarrow \mathbb{K}$ integrierbar. Sei $\mathcal{Z} = \{x_0, \dots, x_N\}$ eine Zerlegung von $[a,b]$, und sei $\xi = (\xi_1, \dots, \xi_n)^T$ ein zugehöriger Zwischenpunktvektor. Zeige mit Hilfe der Hölderschen Ungleichung*

$$\sum_{k=1}^{N} |f(\xi_k)\, g(\xi_k)|\, (x_k - x_{k-1}) \quad \leq$$
$$\left(\sum_{k=1}^{N} |f(\xi_k)|^p (x_k - x_{k-1}) \right)^{1/p} \left(\sum_{k=1}^{N} |g(\xi_k)|^q (x_k - x_{k-1}) \right)^{1/q}$$

und schließe hieraus mit der Definition des Integrals auf die Gültigkeit der Ungleichung

$$\int_a^b |f(x)g(x)|\, dx \leq \left(\int_a^b |f(x)|^p\, dx \right)^{1/p} \left(\int_a^b |g(x)|^q\, dx \right)^{1/q}$$

mit $p, q > 1$ und $1/p + 1/q = 1$.

Lemma 1.1.5 (Minkowskische Ungleichung) *Für $a = (a_1, \dots, a_n)^T, b = (b_1, \dots, b_n)^T \in \mathbb{K}^n$ und $p \geq 1$ gilt*

$$\Big(\sum_{k=1}^{n} |a_k + b_k|^p \Big)^{1/p} \quad \leq \quad \Big(\sum_{k=1}^{n} |a_k|^p \Big)^{1/p} \quad + \quad \Big(\sum_{k=1}^{n} |b_k|^p \Big)^{1/p}.$$

Beweis: Für $p = 1$ ist die Behauptung klar wegen der Dreiecksungleichung. Sei jetzt $p > 1$, und sei o. B. d. A. angenommen, dass $A = \sum_{k=1}^{n} |a_k + b_k|^p > 0$. Dann folgt mit der Hölderschen Ungleichung und $1/q = 1 - 1/p$, also $(p-1)\, q = p$:

$$\begin{aligned}
\sum_{k=1}^{n} |a_k|\, |a_k + b_k|^{p-1} \quad &\leq \quad \Big(\sum_{k=1}^{n} |a_k|^p \Big)^{1/p} \Big(\sum_{k=1}^{n} |a_k + b_k|^{(p-1)q} \Big)^{1/q} \\
&= \quad A^{1/q} \Big(\sum_{k=1}^{n} |a_k|^p \Big)^{1/p},
\end{aligned}$$

und dasselbe gilt, wenn wir die a_k und b_k vertauschen. Daraus folgt wegen

$$A \;\leq \sum_{k=1}^{n} |a_k|\, |a_k + b_k|^{p-1} \; + \; \sum_{k=1}^{n} |b_k|\, |a_k + b_k|^{p-1}$$

die Ungleichung

$$A \;\leq A^{1/q} \left(\Big(\sum_{k=1}^{n} |a_k|^p \Big)^{1/p} \; + \; \Big(\sum_{k=1}^{n} |b_k|^p \Big)^{1/p} \right),$$

und Division durch $A^{1/q}$ ergibt die Behauptung. □

Aufgabe 1.1.6 (Minkowskische Ungleichung für Integrale)
Zeige für $p \geq 1$ und f, g so, dass die rechtsstehenden Integrale existieren:

$$\left(\int_a^b |f(x) + g(x)|^p dx \right)^{1/p} \leq \left(\int_a^b |f(x)|^p dx \right)^{1/p} + \left(\int_a^b |g(x)|^p dx \right)^{1/p}.$$

1.2 Äquivalenzrelationen

Definition 1.2.1 *Sei X eine beliebige nicht-leere Menge. Eine Relation auf X, also eine Teilmenge $R \subset X \times X$, heißt* eine Äquivalenzrelation *auf X, falls für beliebige $x, x_1, x_2, x_3 \in X$ gilt:*

(R) $\quad (x, x) \in R$ (Reflexivität)

(S) $\quad (x_1, x_2) \in R \implies (x_2, x_1) \in R$ (Symmetrie)

(T) $\quad (x_1, x_2) \in R$ und $(x_2, x_3) \in R \implies (x_1, x_3) \in R$ (Transitivität)

Statt $(x_1, x_2) \in R$ schreiben wir auch $x_1 \sim x_2$ und sagen in Worten: x_1 ist äquivalent zu *x_2. Für $x \in X$ sei $A_x = \{\tilde{x} \in X : \quad x \sim \tilde{x}\ \}$. Wir nennen ein solches A_x eine* Äquivalenzklasse*. Ein beliebiges Element einer Äquivalenzklasse heißt auch* ein Repräsentant *dieser Äquivalenzklasse.*

Proposition 1.2.2 *Sei X eine beliebige nicht-leere Menge mit einer Äquivalenzrelation auf X. Seien A_{x_1} und A_{x_2} zwei Äquivalenzklassen mit $A_{x_1} \cap A_{x_2} \neq \emptyset$. Dann gilt bereits $A_{x_1} = A_{x_2}$.*

Beweis: Sei $x \in A_{x_1} \cap A_{x_2}$. Dann ist $x \sim x_1$ und $x \sim x_2$, und aus der Symmetrie und Transitivität folgt dann $x_1 \sim x_2$, also $x_2 \in A_{x_1}$. Sei jetzt $x \in A_{x_2}$, also $x_2 \sim x$. Dann folgt aber aus der Transitivität, dass $x_1 \sim x$, also $x \in A_{x_1}$ ist. Somit ist $A_{x_2} \subset A_{x_1}$. Die Umkehrung gilt aber ebenso. □

Korollar zu Proposition 1.2.2 (Zerlegung in Äquivalenzklassen) *Sei X eine beliebige nicht-leere Menge mit einer Äquivalenzrelation auf X. Dann bilden die Äquivalenzklassen eine Zerlegung von X, d. h., X ist die Vereinigung aller Äquivalenzklassen, und je zwei verschiedene Äquivalenzklassen sind disjunkt.*

Beweis: Da jedes $x \in X$ in der Äquivalenzklasse A_x liegt, folgt $X = \cup_{x \in X} A_x$. Dass verschiedene Äquivalenzklassen disjunkt sind, ist äquivalent zur Proposition. □

Aufgabe 1.2.3 *Sei $n \in \mathbb{N}$. Zeige: Auf $\mathbb{Z}$ wird durch*

$$p \sim q \quad \Longleftrightarrow \quad \exists\, m \in \mathbb{Z} : \quad p - q = n\, m$$

eine Äquivalenzrelation definiert. Formuliere in Worten, wann zwei ganze Zahlen in diesem Sinn äquivalent sind, und bestimme die Anzahl der Äquivalenzklassen.

Aufgabe 1.2.4 *Seien $a < b$, und sei $p \geq 1$. Zeige: Auf der Menge $R[a, b]$ aller über $[a, b]$ Riemann-integrierbaren Funktionen wird durch*

$$f \ \sim\ g \quad \Longleftrightarrow \quad \int_a^b |f(x) - g(x)|^p \, dx = 0$$

eine Äquivalenzrelation definiert. Zeige weiter, dass die Anzahl der Äquivalenzklassen unendlich ist, und dass jede Äquivalenzklasse überabzählbar ist.

Aufgabe 1.2.5 *Zeige mit den Bezeichnungen und Voraussetzungen der vorigen Aufgabe:*

$$f_1 \ \sim\ g_1\,, \quad f_2 \ \sim\ g_2 \quad \Longrightarrow \quad f_1 + f_2 \ \sim\ g_1 + g_2\,, \quad f_1\, f_2 \ \sim\ g_1\, g_2\,.$$

Schließe hieraus, dass es sinnvoll ist, von der Summe und dem Produkt zweier Äquivalenzklassen zu sprechen.

1.3 Normierte Räume

Definition 1.3.1 *Sei X ein beliebiger Vektorraum über $\mathbb{K}$. Eine Abbildung*

$$\|\cdot\| : \quad X \longrightarrow \mathbb{R}, \quad x \longmapsto \|x\|$$

heißt eine Norm *auf X, wenn folgendes gilt:*

(N1) $\forall\, x \in X : \qquad \|x\| \geq 0\,; \quad \|x\| = 0 \iff x = 0$ (Positivität)

(N2) $\forall\, x \in X,\ \forall\, \alpha \in \mathbb{K} : \quad \|\alpha\, x\| \ = \ |\alpha|\, \|x\|$ (Homogenität)

(N3) $\forall\, x_1, x_2 \in X : \quad \|x_1 + x_2\| \ \leq \ \|x_1\| + \|x_2\|$ (Dreiecksungleichung)

Der Vektorraum X, zusammen mit einer Norm $\|\cdot\|$ auf X, heißt dann ein normierter Raum*, und wir schreiben gelegentlich $(X, \|\cdot\|)$, aber meist kürzer X für einen solchen normierten Raum. Sind auf einem Vektorraum X* zwei *Normen $\|\cdot\|_1, \|\cdot\|_2$ gegeben, so nennen wir diese* äquivalent*, wenn es Konstanten $C, K \in \mathbb{R}_+$ gibt, für die*

$$\forall\, x \in X : \qquad C\,\|x\|_1 \ \leq \ \|x\|_2 \ \leq \ K\,\|x\|_1\,.$$

Dies ist eine Äquivalenzrelation auf der Menge aller Normen auf X.

Aufgabe 1.3.2 *Zeige, dass die oben definierte Äquivalenz von Normen tatsächlich die drei Eigenschaften einer Äquivalenzrelation besitzt.*

Aufgabe 1.3.3 (Kartesische Produkte normierter Räume) *Gegeben seien zwei normierte Räume $(X_1, \|\cdot\|_1)$ und $(X_2, \|\cdot\|_2)$ über dem gleichen Körper $\mathbb{K}$. Zeige dass die Abbildung*

$$(x_1, x_2) \ \mapsto \ \|x_1\|_1 \ + \ \|x_2\|_2$$

eine Norm auf dem kartesischen Produkt $X_1 \times X_2$ ist.

Aufgabe 1.3.4 (Dreiecksungleichung nach unten) *Zeige: In einem normierten Raum X gilt*

$$\Big|\, \|x_1\| \ - \ \|x_2\| \,\Big| \ \leq \|x_1 \ - \ x_2\| \qquad \forall\, x_1, x_2 \in X\,.$$

Beispiel 1.3.5 *Für $x = (x_1, \ldots, x_n)^T \in \mathbb{K}^n$ und $1 \leq p \leq \infty$ sei*

$$\|x\|_p \ = \ \begin{cases} \Big(|x_1|^p + \ldots + |x_n|^p\Big)^{1/p} & (1 \leq p < \infty), \\[2ex] \sup\{|x_1|, \ldots, |x_n|\} & (p = \infty). \end{cases}$$

Dadurch ist für jedes feste p eine Norm auf $\mathbb{K}^n$ definiert; für $p < \infty$ ist die Dreiecksungleichung äquivalent zur Minkowskischen Ungleichung. Wir nennen $\|\cdot\|_p$ die p-Norm auf $\mathbb{K}^n$ *und sprechen für $p = 2$ auch von der* Euklidischen Norm*. Offenbar gilt $|x_k| \ \leq \|x\|_p$ für alle $k = 1, \ldots, n$, und daraus ergeben sich folgende Abschätzungen:*

$$\|x\|_\infty \ \leq \|x\|_p \ \leq n^{1/p}\, \|x\|_\infty \qquad \forall\ p \in [1, \infty), \quad \forall x \in \mathbb{K}^n.$$

Also sind die Normen $\|\cdot\|_p$ und $\|\cdot\|_\infty$ äquivalent. Daraus ergibt sich aber leicht, dass zwei beliebige p-Normen auf $\mathbb{K}^n$ immer äquivalent sind.

Aufgabe 1.3.6 *Zeige: Für alle $x \in \mathbb{K}^n$ gilt*

$$\lim_{p\to\infty} \|x\|_p \;=\; \|x\|_\infty \,.$$

Beispiel 1.3.7 *Sei $[a, b]$ ein abgeschlossenes Intervall, und sei $C[a, b]$ die Menge aller dort stetigen Funktionen. Für $f \in C[a, b]$ sei*

$$\|f\|_p \;=\; \begin{cases} \left(\int_a^b |f(x)|^p dx \right)^{1/p} & (1 \le p < \infty), \\ \sup\{|f(x)| : \; a \le x \le b\} & (p = \infty). \end{cases}$$

Dies sind ebenfalls Normen auf $C[a, b]$. Diese Normen sind allerdings nicht *äquivalent!*

Definition 1.3.8 *Die Norm $\|f\|_\infty = \sup\{|f(x)| : \; a \le x \le b\}$ heißt* die Supremumsnorm der Funktion *f. Manchmal nennen wir $\|x\| = \sup\{|x_1|, \ldots, |x_n|\}$ auch* die Supremumsnorm des Vektors *$x \in \mathbb{K}^n$.*

Aufgabe 1.3.9 *Finde eine Funktionenfolge (f_n) in $C[a, b]$, für die gilt*

$$\|f_n\|_\infty = 1 \;\forall\; n \ge 1, \qquad \lim_{n\to\infty} \|f_n\|_p = 0 \;\forall\; p \in [1, \infty).$$

Schließe daraus, dass die p-Norm und die Supremumsnorm auf $C[a, b]$ nicht äquivalent sind.

Aufgabe 1.3.10 *Sei $R[a, b]$ die Menge aller auf dem abgeschlossenen Intervall $[a, b]$, $a < b$, Riemann-integrierbaren Funktionen, und sei $\|f\|_p$ für $1 \le p < \infty$ und $f \in R[a, b]$ wie in obigem Beispiel definiert. Zeige:*

1. Auf $R[a, b]$ erfüllt $\|\cdot\|_p$ nicht das Axiom (N1), ist also *keine* Norm.
2. Definiert man eine Äquivalenzrelation auf $R[a, b]$ wie in Aufgabe 1.2.4, so gilt

   $$f \;\sim\; g \quad\Longrightarrow\quad \|f\|_p \;=\; \|g\|_p \qquad \forall\; f, g \in R[a, b] \,.$$

   Das bedeutet, dass die p-Norm von Funktionen innerhalb einer Äquivalenzklasse immer gleich ist.

Also ist es berechtigt, von der Norm einer Äquivalenzklasse zu sprechen. Da es nach Aufgabe 1.2.4 *auch Sinn macht, Äquivalenzklassen zu addieren und zu multiplizieren, folgt: Die Menge der Äquivalenzklassen in $R[a, b]$ ist ein normierter Raum.*

Kapitel 2

Topologie in normierten Räumen

2.1 Innere Punkte und offene Mengen

Definition 2.1.1 *Sei X ein normierter Raum. Für ein $r \in \mathbb{R}_+$ und $x_0 \in X$ heißt*

$$\mathcal{U}_r(x_0) \;=\; K(x_0, r) \;=\; \{x \in X : \quad \|x - x_0\| < r\}$$

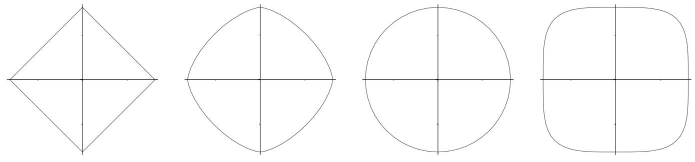
die r-Umgebung von x_0 *oder* die Kugel um x_0 mit Radius r. *Ein x_0 heißt* innerer Punkt *einer Teilmenge $A \subset X$, wenn es ein $r > 0$ gibt mit $\mathcal{U}_r(x_0) \subset A$. Die Teilmenge A heißt* offen, *wenn jeder ihrer Punkte ein innerer Punkt ist.*

Aufgabe 2.1.2 *Skizziere in $\mathbb{R}^2$ die Kugel vom Radius $r = 1$ um den Nullpunkt für die p-Normen mit $p = 1$, $p = 3/2$, $p = 2$ und $p = 4$.*

Lösung: Eine bequeme Lösung der Aufgabe mit MAPLE geschieht mit dem Kommando

```
> plot([signum(cos(t))*(abs(cos(t)))^(2/p),
  signum(sin(t))*(abs(sin(t)))^(2/p),t=0..2*Pi],
  scaling=constrained,tickmarks=[1,1]);
```

und vorheriger Zuweisung der einzelnen Werte für p. Tut man dies, ergeben sich die folgenden vier Bilder:



In Aufgabe 1.3.6 wurde gezeigt, dass der Rand der Kugel für wachsendes p gegen die entsprechende Figur für den Wert $p = \infty$ strebt, und diese ist ein achsenparalleles Quadrat der Seitenlänge 2 mit dem Ursprung als Mittelpunkt. $\square$

Aufgabe 2.1.3 *Begründe, warum in $\mathbb{R}^2$ mit der euklidischen Norm die Menge $O = \{(x,y)^T : \quad x < 2\,y^2\}$ offen ist.*

Lösung: Sei $(x_0, y_0)^T \in O$, also $2\,y_0^2 - x_0 > 0$. Für $r_0 > 0$ ist ein Punkt $(x,y)^T \in K((x_0,y_0)^T, r_0)$ immer von der Form $(x,y)^T = (x_0 + r\,\cos\phi, y_0 + r\,\sin\phi)^T$ mit $r < r_0$ und beliebigem $\phi \in \mathbb{R}$. Die Ungleichung $x < 2\,y^2$ ist also äquivalent zu

$$r(\cos\phi \;-\; 4\,y_0\,\sin\phi) \;-\; 2\,r^2\,\sin^2\phi \;<\; 2\,y_0^2 \;-\; x_0\,.$$

Da die linke Seite höchstens gleich $r\,(1 + 4\,|y_0|)$ ist, gilt diese Ungleichung sicher für $r < r_0$ mit

$$r_0 \;=\; \frac{2\,y_0^2 - x_0}{1 + 4\,|y_0|}\,.$$

Das war zu zeigen. □

Behauptung 2.1.4 *Kugeln um beliebige Punkte mit beliebigen Radien sind immer offen.*

Beweis: Seien $x_0 \in X$, $r \in \mathbb{R}_+$, und sei $x \in K(x_0, r)$, also $\|x - x_0\| < r$. Sei $\rho = r - \|x - x_0\|$. Dann ist $\rho > 0$ und für $\tilde{x} \in \mathcal{U}_\rho(x)$ folgt mit der Dreiecksungleichung

$$\|\tilde{x} - x_0\| \;\le\; \|\tilde{x} - x\| \;+\; \|x - x_0\| \;<\; \rho \;+\; \|x - x_0\| \;=\; r.$$

Also ist $\mathcal{U}_\rho(x) \subset K(x_0, r)$, und somit ist x innerer Punkt von $K(x_0, r)$. Daher folgt die Behauptung. □

Proposition 2.1.5 (Rechenregeln für offene Mengen) *In jedem normierten Raum X gilt:*

(a) *Der Gesamtraum X und die leere Menge sind offen.*

(b) *Der Durchschnitt endlich vieler offener Mengen ist offen.*

(c) *Die Vereinigung beliebig vieler offener Mengen ist offen.*

Beweis: Aussage (a) folgt sofort mit der Definition offener Mengen. **Zu (b):** Seien $O_j \subset X$ alle offen, für $j = 1, \ldots, n$. Sei O ihr Durchschnitt, und sei $x \in O$. Dann gibt es Radien $r_j > 0$ mit $\mathcal{U}_{r_j}(x) \subset O_j$ für alle $j = 1, \ldots, n$. Sei $r = \min\{r_1, \ldots, r_n\}$. Dann folgt $\mathcal{U}_r(x) \subset \mathcal{U}_{r_j}(x) \subset O_j$ für alle $j = 1, \ldots, n$, und deshalb ist $\mathcal{U}_r(x) \subset O$. Daher ist x innerer Punkt von O, und da $x \in O$ beliebig war, folgt dass O offen ist. **Zu (c):** Seien $O_j \subset X$ offen, für alle $j \in J$. Sei O die Vereinigung aller O_j, und sei $x \in O$. Dann gibt es ein $j = j_x \in J$ mit $x \in O_j$, und somit existiert ein $r = r_j > 0$ mit $\mathcal{U}_r(x) \subset O_j$. Dann gilt aber erst recht $\mathcal{U}_r(x) \subset O$, und deshalb ist O offen. □

Aufgabe 2.1.6 *Sei X ein normierter Raum, und seien $O_n = K(0, 1/n) \subset X$ für $n \in \mathbb{N}$. Zeige: Alle O_n sind offen, aber ihr Durchschnitt ist es nicht.*

Definition 2.1.7 *Sei X ein normierter Raum. Die Familie aller offenen Teilmengen von X heißt auch* eine Topologie *auf X.*

Proposition 2.1.8 *Äquivalente Normen auf einem Vektorraum X über $\mathbb{K}$ definieren die gleiche Topologie auf X.*

Beweis: Seien $\|\cdot\|_1$ und $\|\cdot\|_2$ äquivalente Normen auf X. Sei $O \subset X$ offen bzgl. $\|\cdot\|_1$. Dann gibt es zu jedem $x_0 \in O$ ein $r > 0$ so, dass alle $x \in X$ mit $\|x - x_0\|_1 < r$ in O liegen. Nach Definition der Äquivalenz von Normen gibt es eine Konstante $c > 0$ mit $\|x\|_1 \le c\,\|x\|_2$. Wenn wir also $\rho = r/c$ setzen, so folgt für alle $x \in X$ mit $\|x - x_0\|_2 < \rho$, dass

$$\|x - x_0\|_1 \le c\,\|x - x_0\|_2 < c\,\rho = r.$$

Demnach ist jedes $x \in X$ mit $\|x - x_0\|_2 < \rho$ in O, und deshalb ist O offen bzgl. $\|\cdot\|_2$. Da wir die Numerierung der beiden Normen vertauschen können, folgt auch die Umkehrung. □

Nach obiger Proposition ist es im Folgenden unerheblich, welche der p-Normen auf $\mathbb{K}^n$ wir betrachten. Wir wollen aber vereinbaren, dass wir immer die euklidische Norm zugrundelegen, wenn nicht ausdrücklich etwas anderes gesagt ist!

Aufgabe 2.1.9 *Zeige: Jedes offene Intervall ist eine offene Teilmenge von $\mathbb{R}$, und jede offene Teilmenge von $\mathbb{R}$ ist die Vereinigung von höchstens abzählbar vielen offenen Intervallen.*

2.2 Häufungspunkte, isolierte Punkte, abgeschlossene Mengen

Definition 2.2.1 *Sei X ein normierter Raum, und sei $B \subset X$ beliebig. Ein Punkt $x_0 \in X$ heißt* Häufungspunkt *von B, wenn*

$$\forall\, \varepsilon > 0 \quad \exists\, x \in B \setminus \{x_0\} : \quad \|x - x_0\| < \varepsilon\,.$$

Die Menge aller Häufungspunkte von B sei mit B' bezeichnet. Ein Punkt $x_0 \in B$ heißt ein isolierter Punkt von B, *wenn er kein Häufungspunkt von B ist. Eine Teilmenge $A \subset X$ heißt* abgeschlossen, *falls ihr Komplement $O = X \setminus A$ offen ist.*

Lemma 2.2.2 *Sei X ein normierter Raum, und sei $B \subset X$ beliebig. Genau dann ist B abgeschlossen, wenn es alle seine Häufungspunkte enthält.*

Beweis: Sei $x_0 \notin B$ ein Häufungspunkt von B. Dann gehört x_0 zum Komplement von B, kann aber kein innerer Punkt des Komplementes sein. Also kann B nicht abgeschlossen sein. Sei jetzt B nicht abgeschlossen. Dann gibt es ein $x_0 \notin B$, welches kein innerer Punkt des Komplementes von B ist. Also muss es in jeder Kugel um x_0 einen Punkt von B geben, und deshalb ist x_0 ein Häufungspunkt von B. □

Proposition 2.2.3 (Rechenregeln für abgeschlossene Mengen) *Gegeben sei ein normierter Raum X. Dann gilt immer:*

(a) *Der Gesamtraum X und die leere Menge sind abgeschlossen.*

(b) *Die Vereinigung endlich vieler abgeschlossener Mengen ist abgeschlossen.*

(c) *Der Durchschnitt beliebig vieler abgeschlossener Mengen ist abgeschlossen.*

Beweis: Folgt mit Hilfe der de Morganschen Regeln aus den ensprechenden Regeln für offene Mengen. □

Satz 2.2.4 *Sei X ein normierter Raum, und sei eine Teilmenge $A \subset X$ offen und abgeschlossen zugleich. Dann ist $A = X$ oder $A = \emptyset$.*

Beweis: Sei $B = X \setminus A$, und seien A und B beide nicht leer. Für $x_0 \in A$ und $x_1 \in B$ bilden die Vektoren

$$x(t) \; = \; t\,x_1 \; + \; (1-t)\,x_0\,, \qquad 0 \;\le t \;\le 1,$$

die *Verbindungsstrecke* von x_0 nach x_1. Sei $t_0 = \sup\{t \in [0,1] : \; x(t) \in A\}$. Dann kann $x = x(t_0)$ kein innerer Punkt von A, aber auch keiner von B sein. Falls x zu A gehört, ist deshalb A nicht offen, und sonst muss $x \in B$ sein, so dass dann B nicht offen sein kann. □

Aufgabe 2.2.5 *Finde abgeschlossene Mengen in einem normierten Raum X, deren Vereinigung nicht abgeschlossen ist.*

2.3 Abgeschlossene Hülle, offener Kern, Rand

Definition 2.3.1 *Sei X ein normierter Raum. Für ein $B \subset X$ sei*

$$\mathring{B} \; = \; \bigcup_{\substack{O \subset B \\ O \text{ offen}}} O.$$

Nach den Rechenregeln für offene Mengen ist $\mathring{B}$ offen, und wenn O eine offene Teilmenge von B ist, dann folgt $O \subset \mathring{B}$. Also ist $\mathring{B}$ die größte offene Teilmenge von *B und wird als* der offene Kern von *B bezeichnet. Weiter sei*

$$\overline{B} \; = \; \bigcap_{\substack{A \supset B \\ A \text{ abgeschl.}}} A.$$

Nach den Rechenregeln für abgeschlossene Mengen ist $\overline{B}$ abgeschlossen, und wenn A eine abgeschlossene Obermenge von B ist, dann folgt $A \supset \overline{B}$. Also ist $\overline{B}$ die kleinste abgeschlossene Obermenge von *B und wird als* die abgeschlossene Hülle von *B bezeichnet. Die Menge $\partial B = \overline{B} \setminus \mathring{B}$ heißt* der Rand *von B, und jedes $x \in \partial B$ heißt* ein Randpunkt *von B.*

Aufgabe 2.3.2 *Sei X ein normierter Raum, und sei $B \subset X$. Zeige:*

$$\overline{B} \; = \; B \iff B \text{ abgeschlossen}, \qquad \mathring{B} \; = \; B \iff B \text{ offen}.$$

Proposition 2.3.3 *Sei X normierter Raum, und seien $B \subset X$, $C = X \setminus B$. Dann gilt:*

(a) $\mathring{B}$ *ist die Menge aller inneren Punkte von B.*

(b) $\overline{C} \; = \; X \setminus \mathring{B}, \qquad \mathring{C} \; = \; X \setminus \overline{B}.$

(c) $\partial C \; = \; \partial B \; = \; X \setminus (\mathring{B} \cup \mathring{C}).$

(d) $\overline{B} \; = \; B \, \cup \, B' \; = \; B \, \cup \, \partial B.$

Beweis: **Zu (a):** Ein innerer Punkt von B ist in einer kleinen Kugel enthalten, die ihrerseits ganz zu B gehört. Da diese Kugel offen ist, gehört sie auch zum offenen Kern von B. Umgekehrt ist $\mathring{B}$ offen, und somit ist jedes $x \in \mathring{B}$ ein innerer Punkt von B. **Zu (b):** $X \setminus \mathring{B}$ ist eine abgeschlossene Obermenge von C, und daher gilt $\overline{C} \subset X \setminus \mathring{B}$. Umgekehrt ist $X \setminus \overline{C}$ eine offene Teilmenge von B, also in $\mathring{B}$ enthalten. Daraus folgt die erste Aussage. Die zweite ergibt sich durch Vertauschen von B und C und Anwenden der de

Morganschen Regeln. **Zu (c):** Ein $x \in \partial B$ kann nicht zu $\mathring{B}$, aber auch nicht zu $\mathring{C}$ gehören. Umgekehrt, ist $x \notin \mathring{B} \cup \mathring{C}$, so folgt wegen (b), dass $x \in \overline{B}$ sein muss, und dann ist $x \in \partial B$ nach Definition des Randes. Also gilt die rechte Gleichung, und durch Vertauschen von B und C folgt auch die linke. **Zu (d):** Ein Häufungspunkt x von B ist kein innerer Punkt von C, also folgt $x \in \overline{B}$ wegen (b) und (c). Umgekehrt, ist $x \in \overline{B} \setminus B$, so ist x kein innerer Punkt von C, und daraus folgt, dass x ein Häufungspunkt von B sein muss. Also gilt die linke Gleichung. Die rechte folgt, weil einerseits $\partial B \subset \overline{B}$ ist, und andererseits ein Punkt $x \in \overline{B} \setminus B$ auch in $\overline{B} \setminus \mathring{B} = \partial B$ sein muss. □

Aufgabe 2.3.4 *Sei X ein normierter Raum. Zeige: Ist $B \subset X$ mit $\partial B = \emptyset$, so folgt $B = X$ oder $B = \emptyset$. Untersuche, ob auch die Umkehrung gilt.*

Aufgabe 2.3.5 *Sei X normierter Raum, und seien $x_0 \in X$ und $r \in \mathbb{R}_+$. Zeige:*

$$\overline{K(x_0, r)} = \{x \in X : \ \|x - x_0\| \le r\}.$$

Wir nennen diese Menge auch die abgeschlossene Kugel um x_0 vom Radius r.

2.4 Beschränkte und kompakte Mengen

Definition 2.4.1 *Sei X ein normierter Raum, und sei $B \subset X$. Wir nennen B* beschränkt*, falls ein $r \in \mathbb{R}_+$ existiert, für welches $B \subset K(0, r)$ ist. Wir sagen, dass die Mengen $\{O_\alpha : \alpha \in I\}$ eine* offene Überdeckung *von B bilden, wenn alle O_α offen sind, und wenn*

$$B \ \subset \ \bigcup_{\alpha \in I} O_\alpha .$$

Wir sagen weiter, dass eine solche offene Überdeckung von B eine endliche Teilüberdeckung besitzt*, wenn es $\alpha_1, \ldots, \alpha_n \in I$ gibt mit*

$$B \ \subset \ \bigcup_{j=1}^{n} O_{\alpha_j} .$$

Wir nennen die Menge B kompakt*, wenn* jede *offene Überdeckung von B eine endliche Teilüberdeckung besitzt.*

Aufgabe 2.4.2 (Durchmesser einer Menge) *Sei X ein normierter Raum, und sei $B \subset X$. Wir nennen $d = \sup\{\|x_1 - x_2\| : x_1, x_2 \in B\}$* den Durchmesser *von B. Zeige: Genau dann ist B beschränkt, wenn sein Durchmesser endlich ist.*

Aufgabe 2.4.3 *Sei $X \neq \{0\}$ ein normierter Raum, und seien $x_0 \in X$ und $r \in \mathbb{R}_+$. Berechne den Durchmesser d von $K(x_0, r)$. Untersuche ob dieser Durchmesser angenommen wird, d. h., ob es $x_1, x_2 \in K(x_0, r)$ gibt mit $d = \|x_1 - x_2\|$.*

Satz 2.4.4 *Sei X normierter Raum, und sei $B \subset X$ kompakt. Dann ist B abgeschlossen und beschränkt.*

Beweis: Die Menge aller Kugeln $K(0, n)$, mit beliebigem $n \in \mathbb{N}$, bilden sicher eine offene Überdeckung von B. Da B kompakt ist, muss es eine endliche Teilüberdeckung geben. Da $K(0, n) \subset K(0, n+1)$ ist, bedeutet das die Existenz eines $n_0 \in \mathbb{N}$ mit $B \subset K(0, n_0)$. Also ist B beschränkt. Sei $x_0 \notin B$. Die Mengen $O_n = \{x \in X : \ \|x - x_0\| > 1/n\}$, $n \in \mathbb{N}$, sind offen wegen Aufgabe 2.3.5 und bilden offenbar eine

Überdeckung von B. Da B kompakt ist, und da wiederum $O_n \subset O_{n+1}$ ist, gibt es ein $n_0 \in \mathbb{N}$ mit $B \subset O_{n_0}$. Daraus folgt, dass $K(x_0, 1/n_0)$ keinen Punkt von B enthält, und deshalb ist x_0 kein Häufungspunkt von B. Also ist B abgeschlossen. □

Aufgabe 2.4.5 *Zeige: Ist X normierter Raum und $K \subset X$ kompakt, so ist jede abgeschlossene Teilmenge von K ebenfalls kompakt. Insbesondere ist die leere Menge kompakt!*

Aufgabe 2.4.6 *Zeige: In jedem normierten Raum ist der Durchschnitt beliebig vieler kompakter Mengen wieder kompakt.*

2.5 Konvexe und sternförmige Mengen

Definition 2.5.1 *Sei X ein beliebiger Vektorraum über $\mathbb{K}$. Eine Teilmenge $B \subset X$ heißt* sternförmig *bzgl. eines $x_0 \in B$, wenn für jedes $x \in B$ die ganze Verbindungsstrecke von x_0 nach x zu B gehört. Wir nennen B* konvex, *wenn es sternförmig bzgl. jedes Punktes $x_0 \in B$ ist. Also ist B genau dann konvex, wenn mit je zwei Punkten auch deren Verbindungsstrecke zu B gehört.*

Aufgabe 2.5.2 *Sei X normierter Raum. Zeige: Jede Kugel in X ist konvex.*

Behauptung 2.5.3 *Durchschnitte beliebig vieler konvexer Mengen sind wieder konvex.*

Beweis: Folgt direkt aus der Definition (auch, falls der Durchschnitt leer sein sollte). □

Definition 2.5.4 *Sei X ein beliebiger Vektorraum über $\mathbb{K}$, und sei $B \subset X$. Die Menge*

$$B_h \;=\; \bigcap_{\substack{C \supset B \\ C \ \text{konvex}}} C$$

ist nach obiger Behauptung konvex und enthält B. Also ist B_h die kleinste konvexe Obermenge *von B und wird* die konvexe Hülle *von B genannt.*

Aufgabe 2.5.5 *Sei X ein beliebiger Vektorraum über $\mathbb{K}$, und sei $x_0 \in X$. Zeige: Der Durchschnitt beliebig vieler bzgl. x_0 sternförmiger Mengen ist wieder sternförmig bzgl. x_0. Benutze dies, um eine sternförmige Hülle bzgl. x_0 zu definieren.*

Aufgabe 2.5.6 *Sei X ein beliebiger Vektorraum über $\mathbb{K}$, sei $A \subset X$ konvex, und seien $x_1, \ldots, x_n \in A$ sowie $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ mit $\alpha_j \geq 0$ für alle j und $\sum_{j=1}^n \alpha_j = 1$. Zeige*

$$c(x_j; \alpha_j) \;:=\; \sum_{j=1}^n \alpha_j \, x_j \;\in\; A\,.$$

Man nennt jede solche Summe auch eine Konvexkombination *der Vektoren $x_1, \ldots, x_n$.*

Aufgabe 2.5.7 *Sei X ein beliebiger Vektorraum über $\mathbb{K}$, und seien $x_1, \ldots, x_n \in X$ gegeben. Zeige: Die Menge aller Konvexkombinationen der $x_1, \ldots, x_n$ ist genau gleich der konvexen Hülle der endlichen Menge $\{x_1, \ldots, x_n\}$. Interpretiere diese Menge anschaulich für $n \leq 3$.*

2.6 Zusammenhang

Definition 2.6.1 *Sei X ein normierter Raum, und sei $B \subset X$. Wir nennen B* unzusammenhängend*, falls offene Mengen $O_1, O_2 \subset X$ existieren mit*

$$B \subset O_1 \cup O_2 \,, \quad B \cap O_1 \cap O_2 = \emptyset \,, \quad B \cap O_j \neq \emptyset \,, \quad 1 \leq j \leq 2.$$

Falls dies nicht so ist, heißt B zusammenhängend*. Seien $x_0, \ldots, x_m \in X$. Das Tupel aller Verbindungsstrecken von x_{j-1} nach x_j, für $j = 1, \ldots, m$, heißt* der Streckenzug *oder* das Polygon *zu den Eckpunkten x_j. Wir sagen auch: Der Streckenzug* verbindet x_0 mit x_m in der Menge B*, wenn alle Verbindungsstrecken in B liegen. Wir nennen B* polygonzusammenhängend*, wenn sich je zwei Punkte von B durch ein Polygon in B verbinden lassen.*

Aufgabe 2.6.2 *Gib ein Beispiel für eine Teilmenge von $\mathbb{R}^2$, welche zusammenhängend, aber nicht polygonzusammenhängend ist.*

Aufgabe 2.6.3 *Zeige: Eine offene Teilmenge eines normierten Raumes ist genau dann unzusammenhängend, wenn sie Vereinigung zweier offener nichtleerer und disjunkter Mengen ist.*

Aufgabe 2.6.4 *Zeige: Konvexe und sternförmige Mengen sind immer polygonzusammenhängend.*

Lemma 2.6.5 *Sei X ein normierter Raum, und sei $B \subset X$ polygonzusammenhängend. Dann ist B zusammenhängend.*

Beweis: Seien offene Mengen O_1, O_2 in X mit $B \subset O_1 \cup O_2$, $B \cap O_j \neq \emptyset$, für $1 \leq j \leq 2$, gegeben. Zu zeigen ist dann $B \cap O_1 \cap O_2 \neq \emptyset$. Seien $x \in B \cap O_1$, $\tilde{x} \in B \cap O_2$ gewählt. Dann gibt es einen Streckenzug, der beide Punkte in B verbindet. Seien $x_0 (= x), x_1, \ldots, x_m (= \tilde{x})$ seine Eckpunkte, und sei $k \in \{0, \ldots, m\}$ maximal gewählt, so dass $x_k \in B \cap O_1$ gilt. Falls $k = m$ ist, ist nichts mehr zu zeigen. Andernfalls sei $t_0 = \sup \{t \in [0,1] : \ x(t) = (1-t)\, x_k + t\, x_{k+1} \in O_1\}$. Da O_1 offen ist, folgt $x(t_0) \notin O_1$, also $x(t_0) \in O_2$. Da O_2 ebenfalls offen ist, gilt $x(t) \in O_2$ für $t < t_0$, mit $t_0 - t$ genügend klein. Dann ist aber $x(t) \in B \cap O_1 \cap O_2$ für wenigstens ein solches t, was zu zeigen war. □

Definition 2.6.6 *Sei X ein normierter Raum. Eine offene und zusammenhängende Teilmenge von X heißt* ein Gebiet*.*

Satz 2.6.7 (Polygonzusammenhang von Gebieten) *In einem normierten Raum ist jedes Gebiet G polygonzusammenhängend.*

Beweis: Wir nennen zwei Punkte in G äquivalent, wenn sie sich durch ein Polygon in G verbinden lassen. Man sieht sofort, dass dies tatsächlich eine Äquivalenzrelation auf G ist, und dass alle Äquivalenzklassen offen sind. Wenn man annimmt, dass es mehrere Äquivalenzklassen gibt, folgt mit Aufgabe 2.6.3 ein Widerspruch zum Zusammenhang von G. □

Aufgabe 2.6.8 *Sei B eine zusammenhängende Teilmenge eines normierten Raumes X, und sei $B \subset C \subset \overline{B}$. Zeige: Dann ist auch C zusammenhängend.*

Aufgabe 2.6.9 *Zeige: Eine Teilmenge von $\mathbb{R}$ ist genau dann zusammenhängend, wenn sie ein Intervall ist.*

Aufgabe 2.6.10 *Zeige: Eine offene Teilmenge O eines normierten Raumes X ist Vereinigung von paarweise disjunkten Gebieten. Jedes solche Gebiet heißt auch eine* Zusammenhangskomponente *von O.*

Aufgabe 2.6.11 *Zeige: In $\mathbb{R}^n$ hat eine offene Menge höchstens abzählbar viele Zusammenhangskomponenten.*

Kapitel 3

Konvergenz in normierten Räumen

3.1 Konvergenz und Vollständigkeit

Definition 3.1.1 *Sei X ein normierter Raum, und sei $(x_m)_{m=p}^{\infty}$ eine Folge in X. Wir nennen (x_m)* beschränkt*, wenn ein $r \in \mathbb{R}_+$ existiert mit $\|x_m\| \leq r$ für alle $m \geq p$. Wir nennen (x_m)* Cauchy-Folge*, wenn*

$$\forall\, \varepsilon > 0 \quad \exists\, N \in \mathbb{R}_+ \quad \forall\, m, \mu \in \mathbb{N}: \quad m, \mu \geq N \implies \|x_m - x_\mu\| < \varepsilon\,.$$

Weiter heißt (x_m) konvergent*, wenn ein $x \in X$ existiert, für welches*

$$\forall\, \varepsilon > 0 \quad \exists\, N \in \mathbb{R}_+ \quad \forall\, m \in \mathbb{N}: \quad m \geq N \implies \|x_m - x\| < \varepsilon\,.$$

Wenn dem so ist, heißt x Grenzwert *der Folge. Wir schreiben dann $\lim\limits_{m\to\infty} x_m = x$ oder $x_m \longrightarrow x$ $(m \to \infty)$. Schließlich heißt die Folge (x_m)* divergent*, wenn sie nicht konvergiert.*

Die folgenden Ergebnisse beweist man genauso wie die entsprechenden Aussagen aus *Analysis I*, wobei man für Vektoren statt $|\cdot|$ jeweils $\|\cdot\|$ schreiben muss. Beachte allerdings, dass nicht behauptet wird, dass jede Cauchy-Folge konvergiert; dies trifft im Allgemeinen nicht zu!

Proposition 3.1.2 *In jedem normierten Raum X gilt:*

(a) *Eine konvergente Folge hat nur einen Grenzwert.*

(b) *Eine konvergente Folge ist Cauchy-Folge.*

(c) *Eine Cauchy-Folge ist beschränkt.*

(d) *Wenn eine Cauchy-Folge eine konvergente Teilfolge besitzt, so ist sie selber konvergent.*

Satz 3.1.3 (Rechenregeln für Grenzwerte) *Sei X normierter Raum, und seien $x_m, y_m \in X$, $\alpha_m \in \mathbb{K}$, für alle $m \in \mathbb{N}$.*

(a) *Falls die Folgen (x_m) und (y_m) beide konvergieren, dann ist auch $(x_m + y_m)$ konvergent, und es gilt*

$$\lim_{m\to\infty} (x_m + y_m) \;=\; \lim_{m\to\infty} x_m \;+\; \lim_{m\to\infty} y_m\,.$$

(b) *Falls die Folgen (x_m) und (α_m) beide konvergieren, dann ist auch $(\alpha_m\, x_m)$ konvergent, und es gilt*

$$\lim_{m\to\infty} \alpha_m\, x_m \;=\; \lim_{m\to\infty} \alpha_m \, \lim_{m\to\infty} x_m\,.$$

Aufgabe 3.1.4 *Beweise* Proposition 3.1.2 *und die anschließenden Rechenregeln.*

Aufgabe 3.1.5 *Sei X ein normierter Raum, und sei $B \subset X$. Zeige: Jeder Häufungspunkt von B ist Grenzwert einer Folge mit Gliedern in B.*

Definition 3.1.6 *Ein normierter Raum X heißt* vollständig, *falls jede Cauchy-Folge konvergiert. Ein vollständiger normierter Raum heißt auch* Banach-Raum. *Allgemeiner heißt eine Teilmenge $B \subset X$* vollständig, *wenn jede Cauchyfolge mit Gliedern in B auch einen Grenzwert in B hat.*

Nicht jeder normierte Raum hat diese wichtige Eigenschaft der Vollständigkeit. Wir werden aber in Abschnitt 3.3 zeigen, dass die Räume $\mathbb{K}^n$ vollständig sind.

Aufgabe 3.1.7 *Zeige: In $C[a,b]$ mit der Supremumsnorm bedeutet Konvergenz einer Folge dasselbe wie die gleichmäßige Konvergenz. Schließe aus einem Satz der* Analysis I, *dass $C[a,b]$ vollständig ist. Zeige weiter, dass der Teilraum der Polynome nicht vollständig ist.*

Aufgabe 3.1.8 *Sei X ein normierter Raum, und sei $B \subset X$. Zeige: Genau dann ist B abgeschlossen, wenn für jede konvergente Folge (x_n), deren Glieder alle zu B gehören, auch der Grenzwert in B liegt.*

Aufgabe 3.1.9 (Intervallschachtelungsprinzip) *Sei X vollständiger normierter Raum. Seien A_n nicht-leere, abgeschlossene und beschränkte Teilmengen von X mit $A_n \supset A_{n+1}$, für alle $n \in \mathbb{N}$, und so dass der Durchmesser der A_n für $n \to \infty$ gegen 0 geht. Zeige: Dann gibt es genau ein $x \in X$ mit $x \in A_n$ für alle $n \in \mathbb{N}$.*

3.2 Folgenkompaktheit, totale Beschränktheit

Definition 3.2.1 *Sei X ein normierter Raum, und sei $B \subset X$. Wir nennen B* folgenkompakt, *wenn jede Folge in B eine konvergente Teilfolge besitzt, deren Grenzwert ebenfalls zu B behört. Wir sagen, dass B* total beschränkt *ist, wenn gilt*

$$\forall\, \varepsilon > 0 \quad \exists\, x_1, \ldots, x_n \in B: \qquad B \subset \bigcup_{j=1}^{n} K(x_j, \varepsilon)\,.$$

Wir nennen B vollständig, *wenn jede Cauchy-Folge mit Gliedern in B einen Grenzwert besitzt, der ebenfalls zu B gehört.*

Aufgabe 3.2.2 *Zeige: Ist X ein normierter Raum, und ist A eine vollständige Teilmenge von X, so ist jede abgeschlossene Teilmenge von A ebenfalls vollständig.*

Aufgabe 3.2.3 *Sei c_0 die Menge aller Nullfolgen, versehen mit der Supremumsnorm*

$$\|(x_k)\| \;=\; \sup\{|x_k|: \quad k \in \mathbb{N}\}\,.$$

Sei $B = \{x = (x_k) \in c_0: \quad \|x\| \le 1\}$, also ist B sicherlich beschränkt. Zeige: In B gibt es unendlich viele Elemente x_m mit $\|x_m - x_\mu\| = 2$ für $m \neq \mu$. Schließe daraus, dass B nicht total beschränkt ist.

Satz 3.2.4 (Charakterisierung der Kompaktheit) *Für jeden normierten Raum X sind die folgenden Aussagen für eine Teilmenge $B \subset X$ äquivalent:*

(a) B *ist kompakt.*

(b) B *ist folgenkompakt.*

(c) B *ist vollständig und total beschränkt.*

Beweis: Zu (a) $\Longrightarrow$ (b): Sei B nicht folgenkompakt. Falls B nicht abgeschlossen ist, kann es wegen Satz 2.4.4 auch nicht kompakt sein, und deshalb sei B als abgeschlossen vorausgesetzt. Dann muss es eine Folge (x_m) mit Gliedern aus B geben, welche keine konvergente Teilfolge besitzt. Die Glieder der Folge bilden dann eine unendliche Teilmenge A von B, und aus Aufgabe 3.1.5 folgt, dass A keinen Häufungspunkt hat. Also ist insbesondere A abgeschlossen, und somit $O_0 = X \setminus A$ offen. Ferner muss es zu jedem $m \in \mathbb{N}$ ein $r_m > 0$ geben, für welches $O_m = K(x_m, r_m)$ keinen weiteren Punkt von A (außer x_m) mehr enthält, denn andernfalls wäre ja x_m ein Häufungspunkt von A. Die Mengen $\{O_m : M \in \mathbb{N}_0\}$ bilden dann eine offene Überdeckung von B ohne endliche Teilüberdeckung, und somit kann B nicht kompakt sein. **Zu (b) $\Longrightarrow$ (c):** Jede Cauchy-Folge aus B muss wegen der Folgenkompaktheit von B eine konvergente Teilfolge besitzen und ist deshalb selber konvergent. Daher muss B vollständig sein. Sei jetzt ein $\varepsilon > 0$ und ein $x_1 \in B$ betrachtet. Falls $B \in K(x_1, \varepsilon)$ liegt, brechen wir ab; im anderen Fall wählen wir ein $x_2 \in B \setminus K(x_1, \varepsilon)$. Allgemein: Sind $x_1, \ldots, x_m \in B$ schon gewählt, ist entweder $B \subset \cup_{k=1}^{m} K(x_k, \varepsilon)$, oder wir können $x_{m+1} \in B \setminus \cup_{k=1}^{m} K(x_k, \varepsilon)$ wählen. Würde diese Konstruktion niemals abbrechen, so ergäbe sich eine Folge (x_m) aus B, die keine konvergente Teilfolge enthalten könnte. Das kann aber nicht sein, und somit muss B total beschränkt sein. **Zu (c) $\Longrightarrow$ (a):** Sei angenommen, dass B vollständig und total beschränkt, aber nicht kompakt ist. Dann gibt es eine offene Überdeckung $\{O_\alpha : \alpha \in I\}$ ohne endliche Teilüberdeckung. Zu $\varepsilon = 1/2$ gibt es dann $x_{11}, \ldots, x_{1m_1} \in B$ so, dass $B \subset \cup_{j=1}^{m_1} K(x_{1j}, 1/2)$ ist. Für mindestens ein j wird $B \cap K(x_{1j}, 1/2)$ nicht von endlich vielen O_α überdeckt; o. B. d. A. sei dies für $j = 1$ der Fall, und wir setzen $B_1 = B \cap K(x_{11}, 1/2)$. Allgemein: Zu $\mu \geq 2$ und $\varepsilon = 2^{-\mu}$ gibt es $x_{\mu 1}, \ldots, x_{\mu m_\mu} \in B$ so, dass $B \subset \cup_{j=1}^{m_\mu} K(x_{\mu j}, 2^{-\mu})$ ist, und für $B_\mu = B_{\mu-1} \cap K(x_{\mu 1}, 2^{-\mu})$ reichen endlich viel O_α nicht zur Überdeckung aus. Sicherlich kann B_μ nach Konstruktion nicht die leere Menge sein. Also folgt $K(x_{\mu 1}, 2^{-\mu}) \cap K(x_{\mu-1,1}, 2^{-\mu+1}) \neq \emptyset$, und deshalb muss gelten:

$$\|x_{1\mu} - x_{1,\mu-1}\| < 2^{-\mu} + 2^{-\mu+1} = \frac{3}{2^\mu} \qquad \forall\, \mu \in \mathbb{N}\,.$$

Daher gilt für beliebiges $\mu, p \in \mathbb{N}$:

$$\|x_{1\mu} - x_{1,\mu+p}\| \leq \sum_{j=\mu+1}^{p+\mu} \|x_{1j} - x_{1,j-1}\| \leq \sum_{j=\mu+1}^{p+\mu} \frac{3}{2^j} < \frac{3}{2^\mu}\,.$$

Daher ist $(x_{1\mu})$ eine Cauchy-Folge aus B und muss deshalb gegen ein $x \in B$ konvergieren. Zu diesem x gibt es ein $\alpha_0 \in I$ und ein $\varepsilon > 0$ mit $x \in K(x, \varepsilon) \subset O_{\alpha_0}$. Dann müssen aber alle B_μ für $\mu \geq \mu_0$ ebenfalls in O_{α_0} liegen, was ein Widerspruch zur Konstruktion der Folge ist. □

Aufgabe 3.2.5 *Zeige folgende andere Form des Intervallschachtelungsprinzips: Sind A_n nicht-leere kompakte Teilmengen eines normierten Raumes X mit $A_n \supset A_{n+1}$, für alle $n \in \mathbb{N}$, so ist $\cap_n A_n$ nicht leer.*

3.3 Konvergenz und Kompaktheit in mehreren Dimensionen

Die für diese Vorlesung wichtigsten Räume sind $\mathbb{R}^n$ oder $\mathbb{C}^n$. Dabei sei $n \in \mathbb{N}$ immer fest gewählt, und wir stellen uns $n \geq 2$ vor, obwohl die gemachten Aussagen auch alle für $n = 1$ richtig sind. Wenn wir einen Vektor $x \in \mathbb{K}^n$ betrachten, soll immer $x = (x_1, \ldots, x_n)^T$ gelten; für Folgen (x_m) in $\mathbb{K}^n$ sei entsprechend $x_m = (x_1^{(m)}, \ldots, x_n^{(m)})^T$ gesetzt. Als Norm auf $\mathbb{K}^n$ sei stets die euklidische Norm gewählt, obwohl die folgenden Aussagen auch für jede andere p-Norm gelten. Der Einfachheit halber schreiben wir nur $\|\cdot\|$ an Stelle von $\|\cdot\|_2$.

Satz 3.3.1 (Vollständigkeit von $\mathbb{K}^n$)

(a) *Eine Folge (x_m) aus $\mathbb{K}^n$ konvergiert genau dann gegen ein $x \in \mathbb{K}^n$, wenn gilt*

$$\lim_{m\to\infty} x_k^{(m)} \;=\; x_k \qquad \forall\; k \;=\; 1,\dots,n.$$

(b) *Der Raum $\mathbb{K}^n$ ist vollständig.*

Beweis: **Zu (a):** Gelte $\lim_{m\to\infty} x_m = x$, also per Definition $\lim_{m\to\infty} \|x_m - x\| = 0$. Wegen $|x_k^{(m)} - x_k| \le \|x_m - x\|$ folgt dann also $\lim_{m\to\infty} x_k^{(m)} = x_k$, für alle $k = 1,\dots,n$. Umgekehrt folgt aus $\lim_{m\to\infty} x_k^{(m)} = x_k$ und den Rechenregeln für Grenzwerte, dass $\lim_{m\to\infty} \sum_{k=1}^n |x_k^{(m)} - x_k|^2 = 0$ ist, und aus der Stetigkeit der Quadratwurzel folgt dann $\lim_{m\to\infty} \|x_m - x\| = 0$. **Zu (b):** Wenn (x_m) eine Cauchy-Folge in $\mathbb{K}^n$ ist, so ist für jedes $k = 1,\dots,n$ die Folge $(x_k^{(m)})_{m=1}^\infty$ eine Cauchy-Folge in $\mathbb{K}$. Nach dem Cauchy-Kriterium sind diese Folgen alle konvergent. Wenn wir die Grenzwerte mit x_k bezeichnen und daraus einen Vektor $x \in \mathbb{K}^n$ bilden, folgt mit Teil (a) des Satzes, dass (x_m) gegen x konvergiert. $\square$

In einem allgemeinen normierten Raum ist es nicht leicht zu erkennen, ob eine Teilmenge kompakt ist; in $\mathbb{K}^n$ ist dies anders:

Satz 3.3.2 (Satz von Heine-Borel) *Eine Teilmenge von $\mathbb{K}^n$ ist genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.*

Beweis: Wegen Satz 2.4.4 ist die eine Richtung der Aussage in jedem normierten Raum richtig. Sei deshalb jetzt $B \subset \mathbb{K}^n$ abgeschlossen und beschränkt. Mit Satz 3.2.4 ist nur die Folgenkompaktheit von B zu zeigen. Dazu sei (x_m) eine Folge aus B. Dann ist (x_m) beschränkt, und wegen $|x_k^{(m)}| \le \|x_m\|$ ist auch jede der Folgen $(x_k^{(m)})_{m=1}^\infty$ beschränkt. Nach dem Satz von Bolzano und Weierstraß besitzt $(x_1^{(m)})_{m=1}^\infty$ eine konvergente Teilfolge. Die entsprechende Teilfolge von $(x_2^{(m)})_{m=1}^\infty$ ist ebenfalls beschränkt und besitzt wieder eine konvergent Teilfolge, u. s. w. Insgesamt sieht man, dass die Ausgangsfolge (x_m) eine Teilfolge besitzt, für die *alle* Koordinatenfolgen konvergieren. Nach dem vorausgegangenen Satz ist dies aber gleichbedeutend mit der Konvergenz der Teilfolge selber. Da B abgeschlossen ist, muss der Grenzwert dieser Teilfolge zu B gehören, und das zeigt die Folgenkompaktheit. $\square$

Korollar zu Satz 3.3.2 (Satz von Bolzano und Weierstraß in $\mathbb{K}^n$) *Eine beschränkte Folge in $\mathbb{K}^n$ besitzt eine konvergente Teilfolge.*

Beweis: Sei (x_m) beschränkt, und sei B die abgeschlossene Hülle von $\{x_m :\; m \in \mathbb{N}\}$. Dann ist B abgeschlossen und beschränkt, also kompakt. Mit Satz 3.2.4 ist B auch folgenkompakt, und das ist die Behauptung. $\square$

Aufgabe 3.3.3 *Zeige, ohne den Satz von Heine-Borel zu benutzen, dass in $\mathbb{K}^n$ jede beschränkte Menge auch total beschränkt ist.*

Aufgabe 3.3.4 *Sei c_0 die Menge aller Nullfolgen mit Gliedern in $\mathbb{K}$, zusammen mit der Supremumsnorm. Finde in c_0 eine beschränkte Folge ohne konvergente Teilfolge. Schließe daraus, dass der Satz von Heine-Borel im Allgemeinen für unendlich-dimensionale normierte Räume falsch ist.*

Kapitel 4

Stetigkeit

4.1 Stetigkeit in normierten Räumen

Im Folgenden seien X_1, X_2 zwei normierte Räume. Für diese Vorlesung ist es ausreichend, sich $X_1 = \mathbb{K}^n$ und $X_2 = \mathbb{K}^m$ vorzustellen (mit $n, m \in \mathbb{N}$), aber die folgenden Begriffe sind auch allgemein sinnvoll.

Definition 4.1.1 *Sei $D \subset X_1$ nicht leer, und sei $f : D \longrightarrow X_2$. Wir nennen f* stetig in einem Punkt *$x_0 \in D$, falls gilt*

$$\forall\, \varepsilon > 0 \quad \exists\, \delta > 0 \quad \forall\, x \in D : \quad \|x - x_0\| < \delta \implies \|f(x) - f(x_0)\| < \varepsilon. \tag{4.1.1}$$

Falls f in jedem Punkt von D stetig ist, sagen wir kurz: f ist auf D stetig*. Wir sagen, dass f* auf D einer Lipschitzbedingung genügt*, oder* Lipschitzstetig ist*, falls eine Konstante $L \in \mathbb{R}_+$ existiert, so dass*

$$\|f(x_0) \; - \; f(x_1)\| \;\le\; L\,\|x_0 \; - \; x_1\| \qquad \forall\, x_0, x_1 \in D\,.$$

Jedes solche L heißt auch Lipschitzkonstante für f (auf D).

Beispiel 4.1.2 *Überraschenderweise ist in unendlich-dimensionalen normierten Räumen nicht jede lineare Abbildung von X_1 nach X_2 stetig; dies wird in der Vorlesung* Funktionalanalysis *näher besprochen. Auf jedem normierten Raum X ist aber die Abbildung $x \mapsto \|x\|$ stetig auf dem ganzen Raum X; dies folgt aus der Dreiecksungleichung nach unten. Ausserdem ist die identische Abbildung stetig auf X.*

Bemerkung 4.1.3 *Die folgenden Aussagen werden genau wie in* Analysis I *bewiesen:*

(a) *Falls x_0 ein isolierter Punkt von D ist, ist jede auf D definierte Funktion in x_0 stetig.*

(b) *Folgenstetigkeit ist äquivalent zur Stetigkeit; d. h., f ist genau dann stetig in x_0, wenn für jede Folge (x_n) aus D gilt*
$$\lim_{n\to\infty} x_n \;=\; x_0 \in D \qquad \Longrightarrow \qquad \lim_{n\to\infty} f(x_n) \;=\; f(x_0)\,.$$

(c) *Erfüllt f eine Lipschitzbedingung auf D, so ist f dort stetig.*

(d) *Sind $f : D \longrightarrow X_2$ und $g : D \longrightarrow X_2$ stetig in $x_0 \in D$, so ist auch $g + f$ dort stetig.*

(e) *Sind $f : D \longrightarrow X_2$ und $g : D \longrightarrow \mathbb{K}$ stetig in $x_0 \in D$, so ist auch $g\,f$ dort stetig.*

(f) *Die Hintereinanderausführung stetiger Funktionen ist stetig. Genauer: Ist auch X_3 ein normierter Raum, ist $D_2 \subset X_2$, und sind $f : D \longrightarrow D_2$ stetig in $x_0 \in D$ sowie $g : D_2 \longrightarrow X_3$ stetig in $f(x_0)$, so ist $g \circ f$ stetig in x_0.*

(g) *Im Fall $X_2 = \mathbb{K}$ ist der Kehrwert einer Funktion f dort stetig, wo f selber stetig und von 0 verschieden ist.*

Proposition 4.1.4 *Seien X_1, X_2 normierte Räume, sei $D \subset X_1$ nicht leer, und sei $f : D \longrightarrow X_2$. Genau dann ist f stetig auf D, wenn zu jeder offenen Teilmenge $O_2 \subset X_2$ eine offene Teilmenge $O_1 \subset X_1$ existiert, für welche $f^{-1}(O_2) = D \cap O_1$ ist.*

Beweis: Sei f stetig auf D, und sei $O_2 \subset X_2$ offen. Zu jedem $x_0 \in f^{-1}(O_2)$ existiert ein $\varepsilon = \varepsilon_{x_0} > 0$ mit $K(f(x_0), \varepsilon) \subset O_2$. Wegen der Stetigkeit von f gibt es ein $\delta = \delta_{x_0} > 0$ derart, dass für alle $x \in D \cap K(x_0, \delta)$ gilt $f(x) \in K(f(x_0), \varepsilon)$. Sei

$$O_1 = \bigcup_{x_0 \in f^{-1}(O_2)} K(x_0, \delta_{x_0}).$$

Dann ist O_1 offen, und für $x \in O_1 \cap D$ gibt es ein $x_0 \in f^{-1}(O_2)$ mit $x \in K(x_0, \delta_{x_0})$, woraus $f(x) \in O_2$ folgt. Umgekehrt sei für ein $\varepsilon > 0$ und ein $x_0 \in D$ die Menge $O_2 = K(f(x_0), \varepsilon)$ betrachtet. Dazu gibt es nach Voraussetzung eine offene Menge $O_1 \subset X_1$ mit $f^{-1}(O_2) = D \cap O_1$. Also ist $x_0 \in O_1$, und es gibt ein $\delta > 0$ mit $K(x_0, \delta) \subset O_1$. Deshalb gilt für alle $x \in D$: Wenn $\|x - x_0\| < \delta$ ist, dann ist $x \in D \cap O_1$, und deshalb folgt $f(x) \in O_2$, oder anders ausgedrückt: $\|f(x) - f(x_0)\| < \varepsilon$, und daher ist f stetig in x_0. □

Aufgabe 4.1.5 *Sei $A = [\alpha_{jk}]$ eine (n, m)-Matrix mit Elementen aus $\mathbb{K}$, und sei*

$$\|A\| = \Big(\sum_{j=1}^{n} \sum_{k=1}^{m} |\alpha_{jk}|^2 \Big)^{1/2}.$$

Zeige: $\|A\,x\| \leq \|A\|\,\|x\|$ für alle $x \in \mathbb{K}^m$, wobei links die euklidische Norm in $\mathbb{K}^n$, und rechts für x die euklidische Norm in $\mathbb{K}^m$ zu nehmen ist. Leite daraus die Stetigkeit von $x \mapsto A\,x$ auf $\mathbb{K}^m$ ab.

Aufgabe 4.1.6 *Zeige: Ist $[a, b]$ eine abgeschlossenes Intervall in $\mathbb{R}$, und ist $X = C[a, b]$ der Raum der auf $[a, b]$ stetigen Funktionen mit der Supremumsnorm, so erfüllt die Abbildung $f \longmapsto \int_a^b f(x)\,dx$ auf $C[a, b]$ eine Lipschitzbedingung und ist deshalb dort stetig.*

4.2 Stetigkeit und Zusammenhang

Der folgende Satz entspricht genau dem Zwischenwertsatz aus *Analysis I*:

Satz 4.2.1 *Seien X_1, X_2 normierte Räume, sei $D \subset X_1$ nicht leer, und sei $f : D \longrightarrow X_2$ stetig auf D. Ist B eine zusammenhängende Teilmenge von D, so ist $f(B)$ ebenfalls zusammenhängend.*

Beweis: Sei $f(B)$ unzusammenhängend. Dann gibt es per Definition zwei offene Mengen $O_1, O_2 \subset X_2$ mit

$$f(B) \;\subset\; O_1 \cup O_2\,, \quad f(B) \cap O_1 \cap O_2 \;=\; \emptyset, \quad f(B) \cap O_j \;\neq\; \emptyset\,, \quad (1 \leq j \leq 2).$$

Nach Proposition 4.1.4 gibt es dann zwei offene Mengen $\tilde{O}_1, \tilde{O}_2 \subset X_1$ mit $D \cap \tilde{O}_j = f^{-1}(O_j)$, für $1 \leq j \leq 2$. Damit folgt

$$B \;\subset\; \tilde{O}_1 \cup \tilde{O}_2\,, \quad B \cap \tilde{O}_1 \cap \tilde{O}_2 \;=\; \emptyset\,, \quad B \cap \tilde{O}_j \;\neq\; \emptyset\,, \quad (1 \leq j \leq 2).$$

Daher ist auch B unzusammenhängend. □

Aufgabe 4.2.2 (Zwischenwertsatz) *Sei X ein normierter Raum, sei $D \subset X$ nicht leer, und sei $f : D \longrightarrow \mathbb{R}$ stetig auf D. Zeige mit Hilfe von* Satz 4.2.1*: Ist B eine zusammenhängende Teilmenge von D, und sind $x_1, x_2 \in B$ mit $f(x_1) = y_1 < f(x_2) = y_2$, so gibt es zu jedem $y \in (y_1, y_2)$ ein $x \in B$ mit $f(x) = y$.*

4.3 Stetigkeit und Kompaktheit

Definition 4.3.1 *Seien X_1, X_2 normierte Räume, sei $D \subset X_1$ nicht leer, und sei $f : D \longrightarrow X_2$. Wir nennen f* auf D gleichmäßig stetig, *falls folgendes gilt:*

$$\forall\, \varepsilon > 0 \quad \exists\, \delta > 0 \quad \forall\, x_1, x_2 \in D : \quad \|x_1 - x_2\| \;<\; \delta \quad \Longrightarrow \quad \|f(x_1) - f(x_2)\| \;<\; \varepsilon\,.$$

Falls $X_2 = \mathbb{R}$ ist, definieren wir die Begriffe Maximum *und* Minimum *genau wie in* Analysis I.

Satz 4.3.2 *Seien X_1, X_2 normierte Räume, sei $D \subset X_1$ nicht leer und kompakt, und sei $f : D \longrightarrow X_2$ stetig auf D. Dann ist f sogar gleichmäßig stetig auf D, und $f(D)$ ist kompakt. Ist $X_2 = \mathbb{R}$, so nimmt f auf D ein Maximum und ein Minimum an.*

Beweis: Sei $\varepsilon > 0$. Wegen der Stetigkeit von f gibt es zu jedem $x \in D$ ein $\delta_x > 0$ derart, dass für alle $\tilde{x} \in D \cap K(x, \delta_x)$ gilt $f(\tilde{x}) \in K(f(x), \varepsilon/2)$. Die Menge aller $K(x, \delta_x/2)$ ist eine offene Überdeckung von D und besitzt wegen der Kompaktheit eine endliche Teilüberdeckung. Also existieren $x_1, \ldots, x_m \in D$ so, dass $D \subset \cup_{k=1}^m K(x_k, \delta_{x_k}/2)$ ist. Mit $2\,\delta = \min\{\delta_{x_1}, \ldots, \delta_{x_m}\}$ gilt dann: Für zwei beliebige Punkte $x, \tilde{x} \in D$ ist $x \in K(x_k, \delta_{x_k}/2)$ für wenigstens ein $k \in \{1, \ldots, m\}$. Wenn $\|x - \tilde{x}\| < \delta$ ist, dann folgt $\tilde{x} \in K(x_k, \delta_{x_k})$. Daraus folgt mit der Dreiecksungleichung, dass $\|f(x) - f(\tilde{x})\| \le \|f(x) - f(x_k)\| + \|f(x_k) - f(\tilde{x})\| < \varepsilon$ ist. Das zeigt die gleichmäßige Stetigkeit von f auf D. Sei jetzt (y_m) eine Folge aus $f(D)$. Dann gibt es für jedes $m \in \mathbb{N}$ mindestens ein $x_m \in D$ mit $f(x_m) = y_m$. Da D kompakt, also auch folgenkompakt ist, muss die Folge (x_m) eine konvergente Teilfolge mit einem Grenzwert $x \in D$ haben. Wegen der Stetigkeit von f konvergiert die entsprechende Teilfolge von (y_n) gegen $f(x) \in f(D)$. Das zeigt die Folgenkompaktheit und damit die Kompaktheit von $f(D)$. Für $X_2 = \mathbb{R}$ folgt: Da $f(D)$ kompakt ist, ist die Menge insbesondere beschränkt und abgeschlossen und muss deshalb ein endliches Supremum haben, welches sogar zur Menge gehört und deshalb ein Maximum ist. Dieses Maximum wird dann natürlich an einer Stelle $x \in D$ angenommen. Genauso schließt man für das Minimum. □

Satz 4.3.3 (Stetigkeit der Umkehrabbildung) *Seien X_1, X_2 normierte Räume, sei $D \subset X_1$ nicht leer und kompakt, und sei $f : D \longrightarrow X_2$ stetig und injektiv auf D. Dann ist die Umkehrfunktion auf $\tilde{D} = f(D)$ stetig.*

Beweis: Sei (y_m) eine Folge aus $\tilde{D}$, welche gegen ein $y \in \tilde{D}$ konvergiert. Seien x_m und x die (eindeutig bestimmten) Urbilder dieser Punkte. Zu zeigen ist $\lim_{m\to\infty} x_m = x$. Wenn wir das Gegenteil annehmen, dann existiert ein $\varepsilon > 0$, für welches $\|x_m - x\| \ge \varepsilon$ für unendlich viele m, d. h. für eine Teilfolge von (x_m), gilt. Wegen der Folgenkompaktheit von D enthält diese Teilfolge selbst wieder eine Teilfolge, welche gegen ein $\tilde{x} \in D$ konvergiert. Wegen der Stetigkeit von f konvergieren die Bilder dieser Teilfolge, also die entsprechende Teilfolge von (y_m), gegen $f(\tilde{x})$, und daher gilt $f(\tilde{x}) = y$, also $\tilde{x} = x$. Das widerspricht aber der Wahl der Teilfolge. □

Aufgabe 4.3.4 *Finde ein Beispiel für eine stetige Funktion mit nicht-kompaktem Definitionsbereich aber kompakter Wertemenge.*

Aufgabe 4.3.5 *Finde ein Beispiel für eine stetige Funktion mit beschränktem Definitionsbereich aber unbeschränkter Wertemenge.*

4.4 Stetigkeit und gleichmäßige Konvergenz

Definition 4.4.1 *Seien X_1, X_2 normierte Räume, sei $D \subset X_1$ nicht leer, und seien $f_m, g_k : D \longrightarrow X_2$. Dann nennen wir (f_m) eine* Funktionenfolge *auf D, und $\sum_{k=1}^{\infty} g_k$ eine* Funktionenreihe *auf D. Die Funktionenfolge (f_m) heißt* auf D punktweise konvergent, *falls für jedes $x \in D$ die Folge $(f_m(x))$ konvergent ist. Ist dies der Fall, so heißt $f : D \longrightarrow X_2$, mit $f(x) = \lim_{m\to\infty} f_m(x)$ für alle $x \in D$,* die Grenzfunktion *der Folge. Analog heißt die Funktionenreihe $\sum_{k=1}^{\infty} g_k$* auf D punktweise konvergent, *falls für jedes $x \in D$ die Reihe $\sum_{k=1}^{\infty} g_k(x)$ konvergent ist, und die Grenzfunktion f ist dann durch $f(x) = \sum_{k=1}^{\infty} g_k(x)$ für alle $x \in D$ gegeben. Die Funktionenfolge (f_m) heißt* auf D gleichmäßig konvergent, *falls sie punktweise gegen die Grenzfunktion f konvergiert, und falls weiter gilt:*

$$\forall\, \varepsilon > 0 \quad \exists\, N \in \mathbb{R}_+ \quad \forall\, m \in \mathbb{N} \quad \forall\, x \in D : \qquad m \geq N \quad \Longrightarrow \quad \|f_m(x) - f(x)\| \;<\; \varepsilon\,.$$

Analog heißt die Funktionenreihe $\sum_{k=1}^{\infty} g_k$ auf D gleichmäßig konvergent, *falls sie punktweise gegen die Grenzfunktion f konvergiert, und falls weiter gilt:*

$$\forall\, \varepsilon > 0 \quad \exists\, N \in \mathbb{R}_+ \quad \forall\, m \in \mathbb{N} \quad \forall\, x \in D : \qquad m \geq N \quad \Longrightarrow \quad \Big\| \sum_{k=1}^{m} g_k(x) - f(x) \Big\| \;<\; \varepsilon\,.$$

Die gleichmäßige Konvergenz der Reihe $\sum_{k=1}^{\infty} g_k$ ist also äquivalent mit der gleichmäßigen Konvergenz der Folge ihrer Partialsummen.

Genau wie in *Analysis I* zeigt man die folgenden Resultate in jedem vollständigen normierten Raum X:

Satz 4.4.2 (Cauchy-Kriterium für gleichmäßige Konvergenz) *Mit den Bezeichnungen wie in der obigen Definition gilt: Die Funktionenfolge (f_n) ist genau dann auf D gleichmäßig konvergent, wenn*

$$\forall\, \varepsilon > 0 \quad \exists\, N \in \mathbb{R}_+ \quad \forall\, m, \mu \in \mathbb{N} \quad \forall\, x \in D : \qquad m, \mu \geq N \quad \Longrightarrow \quad \|f_m(x) - f_\mu(x)\| \;<\; \varepsilon\,.$$

Die Funktionenreihe $\sum_{k=1}^{\infty} g_k$ ist genau dann auf D gleichmäßig konvergent, wenn

$$\forall\, \varepsilon > 0 \quad \exists\, N \in \mathbb{R}_+ \quad \forall\, m, p \in \mathbb{N} \quad \forall\, x \in D : \qquad m \geq N \quad \Longrightarrow \quad \Big\| \sum_{k=m+1}^{m+p} g_k(x) \Big\| \;<\; \varepsilon\,.$$

Satz 4.4.3 (Majorantenkriterium für gleichmäßige Konvergenz) *Mit den Bezeichnungen wie in der obigen Definition gilt: Falls Zahlen $a_k \in \mathbb{R}_+$ existieren, für welche gilt*

$$\|g_k(x)\| \;\leq\; a_k \quad \forall\, x \in D, \qquad \sum_{k=1}^{\infty} a_k \;<\; \infty,$$

dann ist die Funktionenreihe $\sum_{k=1}^{\infty} g_k$ auf D gleichmäßig konvergent.

Satz 4.4.4 *Mit den Bezeichnungen wie in der obigen Definition gilt:*

(a) *Sind alle f_n in einem Punkt $x_0 \in D$ stetig, und ist die Funktionenfolge (f_n) auf D gleichmäßig konvergent, so ist die Grenzfunktion f ebenfalls stetig in x_0.*

(b) *Sind alle g_k in einem Punkt $x_0 \in D$ stetig, und ist die Funktionenreihe $\sum_{k=1}^{\infty} g_k$ auf D gleichmäßig konvergent, so ist die Grenzfunktion f ebenfalls stetig in x_0.*

Ein nützliches Ergebnis, welches wir in *Analysis I* ausgelassen haben, ist der folgende Satz:

Satz 4.4.5 (Satz von Dini) *Sei X ein normierter Raum, sei $D \subset X$ kompakt, und seien $f_m : D \longrightarrow \mathbb{R}$ alle stetig auf D. Weiter seien für alle $x \in D$ die Zahlenfolgen $(f_m(x))$ monoton fallende Nullfolgen. Dann ist (f_m) auf D gleichmäßig konvergent.*

Beweis: Für $\varepsilon > 0$ sei $D_m = \{x \in D : \ f_m(x) < \varepsilon\}$. Zu $x \in D_m$ sei $\tilde{\varepsilon}_x = \varepsilon - f_m(x)$. Wegen der Stetigkeit von f_m gibt es ein $\delta_x > 0$ so, dass für alle $\tilde{x} \in D$ gilt:

$$\|x - \tilde{x}\| < \delta_x \quad \Longrightarrow \quad |f_m(x) - f_m(\tilde{x})| < \tilde{\varepsilon}_x .$$

Sei $O_m = \cup_{x \in D_m} K(x, \delta_x)$. Dann ist O_m offen und enthält D_m. Umgekehrt, ist $\tilde{x} \in D \cap O_m$, so ist $\tilde{x} \in K(x, \delta_x)$ für ein $x \in D_m$, und dann folgt $|f_m(\tilde{x})| \leq |f_m(x)| + |f_m(\tilde{x}) - f_m(x)| < |f_m(x)| + \tilde{\varepsilon}_x = \varepsilon$, und deshalb ist $\tilde{x} \in D_m$. Also gilt $D_m = D \cap O_m$. Da jedes $x \in D$ in einem der D_m, also erst recht in O_m, liegen muss, bilden die O_m eine offene Überdeckung von D. Wegen der Kompaktheit von D gibt es ein $N \in \mathbb{N}$ mit $D \subset \cup_{m=1}^{N} O_m$. Wegen $D_m = D \cap O_m$ folgt daraus $D \subset \cup_{m=1}^{N} D_m$, und da aus der Monotonievoraussetzung folgt dass $D_m \subset D_{m+1}$ ist, folgt sogar $D \subset D_N$, und dann sogar $D \subset D_m$ für alle $m \geq N$. Mit der Definition von D_m folgt dann die Behauptung. $\square$

Aufgabe 4.4.6 *Seien $g_k(x_1, x_2) = (x_1^k + x_2)/k^2$, für $k \in \mathbb{N}$ und $x_1, x_2 \in \mathbb{R}$. Finde einen möglichst großen Bereich D, auf dem die Funktionenreihe $\sum_{k=1}^{\infty} g_k$ gleichmäßig konvergiert und untersuche dort die Stetigkeit der Grenzfunktion.*

Aufgabe 4.4.7 *Leite die gleichmäßige Konvergenz der Funktionenfolge (f_m), mit $f_m(x) = x^m(1-x)$ für $x \in [0,1] \subset \mathbb{R}$, aus dem Satz von Dini ab.*

4.5 Stetigkeit von vektorwertigen Abbildungen

Definition 4.5.1 *Sei X ein normierter Raum, sei $D \subset X$ nicht leer, und sei $f : D \longrightarrow \mathbb{K}^m$. Dann ist jeder Funktionswert $f(x)$ ein Vektor*

$$f(x) = (f_1(x), \dots, f_m(x))^T \qquad \forall\, x \in D.$$

Wir nennen dann die Funktionen $f_k : D \longrightarrow \mathbb{K}$ auch die Koordinatenfunktionen *der vektorwertigen Funktion f und schreiben kurz $f = (f_1, \dots, f_n)^T$. Ist $X = \mathbb{K}^n$, so nennen wir f, aber auch die f_k,* Funktionen der n Variablen $x_1, \dots, x_n$. *Wenn $n \geq 2$ ist, sprechen wir auch von Funktionen* mehrerer Variabler *und schreiben statt $f(x)$ auch $f(x_1, \dots, x_n)$.*

Bemerkung 4.5.2 *Wenn eine Funktion f wie oben gegeben ist, zeigt man leicht, dass die Stetigkeit von f an einer Stelle x_0 äquivalent zur Stetigkeit jeder der Koordinatenfunktionen an dieser Stelle ist. Beachte allerdings, dass bei Funktionen mehrerer Variabler die Stetigkeit nicht schon dann gesichert ist, wenn die Funktion stetig in jeder einzelnen Variablen ist! Dies ist gerade Inhalt der folgenden Aufgaben.*

Aufgabe 4.5.3 *Benutze die Äquivalenz von Stetigkeit und Folgenstetigkeit zusammen mit* Satz 1*, um zu zeigen, dass die Stetigkeit von f äquivalent zur Stetigkeit der Koordinatenfunktionen ist.*

Aufgabe 4.5.4 *Für $x_1, x_2 \in \mathbb{R}$ sei*

$$f(x_1, x_2) = \begin{cases} \dfrac{x_1^2 x_2}{x_1^4 + x_2^2} & \text{für } (x_1, x_2) \neq (0,0), \\[2ex] 0 & \text{für } (x_1, x_2) = (0,0). \end{cases}$$

Zeige: Wenn man für x_1 oder x_2 den Wert 0 einsetzt, wird dieses f eine stetige Funktion der anderen Variablen auf ganz $\mathbb{R}$. Als Funktion von zwei Variablen ist f aber nicht stetig im Nullpunkt.

4.6 Polynome, rationale Funktionen und Potenzreihen

Definition 4.6.1 *Für $x = (x_1, \dots, x_n)^T \in \mathbb{K}^n$ und $p = (p_1, \dots, p_n) \in \mathbb{N}_0^n$ sei*

$$x^p \;=\; x_1^{p_1} \cdot \ldots \cdot x_n^{p_n}\,.$$

Jedes solche p heißt ein Multi-Index*, und die Zahl $|p| = p_1 + \ldots + p_n$ heißt* Betrag *oder* Länge *von p. Eine Funktion $P : \mathbb{K}^n \longrightarrow \mathbb{K}$ der Form*

$$P(x) \;=\; P(x_1, \dots, x_n) \;=\; \sum_{0 \le |p| \le m} a_p\, x^p \qquad \forall\, x \in \mathbb{K}^n$$

mit $a_p \in \mathbb{K}$, heißt ein Polynom *in n Variablen. Die Zahlen a_p heißen* die Koeffizienten *des Polynoms. Für die Menge aller Polynome mit Koeffizienten in $\mathbb{K}$ schreiben wir wieder $\mathbb{K}[x]$; beachte aber, dass hier $x \in \mathbb{K}^n$ ist! Falls alle Koeffizienten $a_p = 0$ sind, ist $P(x) = 0$ für alle $x \in \mathbb{K}^n$, und wir nennen dieses Polynom wieder* das Nullpolynom *oder* die Nullfunktion*. Im anderen Fall können wir annehmen, dass $a_p \neq 0$ ist für mindestens ein p mit $|p| = m$, und dann nennen wir $m = \deg P$ auch* den Grad *des Polynoms. Den Grad des Nullpolynoms definieren wir wieder als $-\infty$. Beachte aber, dass für $n \ge 2$ mehrere Terme in P mit $|p| = \deg P$ und $a_p \neq 0$ vorkommen können, so dass dann nicht klar ist, was man unter dem höchsten Koeffizienten von P verstehen soll. Für zwei Polynome $p, q \in \mathbb{K}[x]$, wobei q nicht das Nullpolynom ist, heißt $r = p/q$ eine* rationale Funktion *in n Variablen. Sie ist überall da definiert, wo das Nennerpolynom nicht verschwindet. Die Menge aller rationalen Funktionen wird wieder mit $\mathbb{K}(x)$ bezeichnet.*

Beispiel 4.6.2 *Sei A eine quadratische n-reihige Matrix. Dann ist die Abbildung $p(x) = x^T A\, x$, $x \in \mathbb{K}^n$, ein Polynom. Der Grad dieses Polynoms ist gleich 2, außer wenn A die Nullmatrix ist. Man spricht auch von der* quadratischen Form mit der Koeffizientenmatrix *A. Für $a \in \mathbb{K}^n$ ist $p(x) = a^T x$ ein Polynom vom Grad 1, außer wenn a der Nullvektor ist.*

Definition 4.6.3 *Ein Ausdruck der Form*

$$\sum_p a_p\ (z - z_0)^p\,, \qquad a_p \in \mathbb{C}, \quad z_0 \;=\; (z_1^{(0)}, \dots, z_n^{(0)})^T \in \mathbb{C}^n \tag{4.6.1}$$

mit Summation über alle Multi-Indizes p, heißt eine Potenzreihe in n Variablen $z = (z_1, \dots, z_n)^T$*. Die Zahlen a_p heißen* die Koeffizienten*, und z_0 heißt* der Entwicklungspunkt *der Potenzreihe.*

Beispiel 4.6.4 *Für $z_0 = 0$ und $a_p = 1$ für alle p erhalten wir die* geometrische Reihe in n Variablen*. Es gilt*

$$\sum_p z^p \;=\; \sum_{p_1=0}^{\infty} \cdots \sum_{p_1=0}^{\infty} z_1^{p_1} \cdot \ldots \cdot z_n^{p_n} \;=\; \frac{1}{1 - z_1} \cdot \ldots \cdot \frac{1}{1 - z_n}\,,$$

wobei die Reihe für alle z mit $\|z\|_\infty = \max\{|z_1|, \dots, |z_n|\} < 1$ absolut konvergiert.

Bemerkung 4.6.5 *Die Menge der Multi-Indizes ist abzählbar unendlich, also kann man durch Anordnung der Terme eine Potenzreihe in mehreren Variablen als normale Reihe schreiben. Allerdings ist klar, dass bei Konvergenzuntersuchungen der richtige Konvergenzbegriff der der* absoluten Konvergenz *sein muss, da sonst der Wert der Reihe von der gewählten Anordnung der Terme abhängen kann.*

Aufgabe 4.6.6 *Zeige die Abzählbarkeit der Menge aller Multi-Indizes.*

Für Potenzreihen in mehreren Variablen gibt es nichts, was dem Konvergenzradius bei bewöhnlichen Potenzreihen entspricht. Das Konvergenzverhalten einer solchen Reihe ist nicht leicht zu beschreiben. Es gilt aber folgendes Ergebnis:

Proposition 4.6.7 (Konvergenz einer Potenzreihe in mehreren Var.)

(a) *Sei die Reihe* (4.6.1) *für ein* $z = z_1 = (z_1^{(1)}, \ldots, z_n^{(1)})^T$ *mit* $z_k^{(1)} \neq z_k^{(0)}$, $1 \leq k \leq n$, *absolut konvergent. Dann konvergiert sie absolut für alle* $z = (z_1, \ldots, z_n)^T$ *mit* $|z_k - z_k^{(0)}| < |z_k^{(1)} - z_k^{(0)}|$ *für alle* $k = 1, \ldots, n$. *Die Konvergenz ist gleichmäßig für*

$$|z_k - z_k^{(0)}| \leq r_k \, |z_k^{(1)} - z_k^{(0)}| \qquad \forall k = 1, \ldots, n\,, \tag{4.6.2}$$

für beliebige Zahlen $0 < r_k < 1$, *für* $k = 1, \ldots, n$.

(b) *Ist* $\alpha_k = \max\{|a_p| : \ |p| = k\}$, *für* $k \in \mathbb{N}_0$, *und hat die Potenzreihe* $\sum_{k=0}^{\infty} \alpha_k \, t^k$ *einen Konvergenzradius* $R > 0$, *so konvergiert* (4.6.1) *absolut für*

$$\|z - z_0\|_\infty = \max\{|z_1 - z_1^{(0)}|, \ldots, |z_n - z_n^{(0)}|\} \ < \ R.$$

Beweis: Zu (a): Aus der Konvergenz für $z = z_1$ folgt die Existenz von $K > 0$ mit $|a_p\,(z_1 - z_0)^p| \leq K$ für alle Multi-Indizes p. Daher gilt für alle z wie in (4.6.2)

$$|a_p\,(z - z_0)^p| \ \leq \ K\, r_1^{p_1} \cdot \ldots \cdot r_n^{p_n} \qquad \forall \ p.$$

Mit dem Majorantenkriterium folgt die Behauptung. **Zu (b):** Sei $r = \|z - z_0\|_\infty < R$. Dann folgt $|(z - z_o)^p| \leq r^{|p|}$, also

$$\sum_p |a_p| \ |(z - z_0)^p| \ \leq \ \sum_{k=0}^{\infty} r^k \ \sum_{|p|=k} |a_p|.$$

Die Anzahl der Multi-Indizes vom Betrag k ist nicht größer als $(k+1)^n$, und deshalb ist die Reihe $\sum_k (k+1)^n \alpha_k \, r^k$ eine konvergente Majorante. □

Aufgabe 4.6.8 *Zeige: Ist* $p = e_k = (0, \ldots, 0, 1, 0, \ldots, 0)^T$, *mit der* 1 *an der* k*-ten Stelle, der* k*-te Basisvektor der kanonischen Basis von* $\mathbb{K}^n$, *so ist* $x^p = x_k$.

Aufgabe 4.6.9 *Zeige: Sind* p *und* q *zwei Multi-Indizes, so ist* $x^{p+q} = x^p \, x^q$.

Aufgabe 4.6.10 *Zeige: Jedes Polynom* $p \in \mathbb{K}[x]$ *vom Grade höchstens gleich* 2 *kann geschrieben werden als* $p(x) = x^T A\, x + a^T x + b$, *mit einer quadratischen* n*-reihigen Matrix* A, *einem Vektor* $a \in \mathbb{K}^n$ *und einer Zahl* $b \in \mathbb{K}$.

Kapitel 5

Differenzialrechnung mehrerer Variabler

Im Folgenden sei immer ein festes nicht leeres Gebiet $G \subset \mathbb{R}^n$ und eine Funktion $f : G \longrightarrow \mathbb{R}$ betrachtet. Wir schreiben dann auch $f(x) = f(x_1, \ldots, x_n)$.

5.1 Richtungsableitungen, partielle Ableitungen, Gradient

Definition 5.1.1 *Sei e ein beliebiger* Einheitsvektor*, d. h., $e \in \mathbb{R}^n$ und $\|e\| = 1$. Falls für ein $x \in G$ der Grenzwert*

$$f_e(x) \;=\; \lim_{t\to 0} \frac{f(x+t\,e) \;-\; f(x)}{t}$$

existiert, heißt $f_e(x)$ die Ableitung von f an der Stelle x in Richtung von e*. Jede solche Ableitung heißt auch kurz* eine Richtungsableitung von f im Punkt x*. Ist e der k-te Basisvektor e_k der kanonischen Basis von $\mathbb{R}^n$, so heißt diese Richtungsableitung auch* k-te partielle Ableitung *von f im Punkt x, oder* die Ableitung von f nach der Variablen x_k im Punkt x*, und wir schreiben auch*

$$f_{x_k}(x) \;=\; \frac{\partial f(x)}{\partial x_k} \;=\; \lim_{t\to 0} \frac{f(x_1, \ldots, x_{k-1}, x_k + t, x_{k+1}, \ldots, x_n) - f(x_1, \ldots, x_n)}{t}$$

für diese partielle Ableitung. Wenn an einer Stelle $x \in G$ alle partiellen Ableitungen existieren, so heißt f im Punkt x partiell differenzierbar*, und der Zeilenvektor*

$$\operatorname{grad} f(x) \;=\; (f_{x_1}(x), \ldots, f_{x_n}(x))$$

heißt der Gradient *von f im Punkt x.*

Bemerkung 5.1.2 *Die Berechnung einer partiellen Ableitung geschieht wie folgt: Alle Variablen werden wie Konstante behandelt mit Ausnahme einer, und nach dieser einen Variablen wird nach den Regeln aus* Analysis I *differenziert. Wie man andere Richtungsableitungen, außer über die Definition, berechnet, wird später behandelt. Wichtig ist,* dass aus der partiellen Differenzierbarkeit, oder sogar aus der Existenz aller Richtungsableitungen, i. A. nicht die Stetigkeit folgt*. Dazu siehe das folgende Beispiel:*

Beispiel 5.1.3 *Wir definieren f wie in* Aufgabe 4.5.4*. Sei $e = (\cos\phi, \sin\phi)^T$, für irgend ein $\phi \in \mathbb{R}$. Dann ist $\|e\| = 1$, und jeder Einheitsvektor in $\mathbb{R}^2$ kann so geschrieben werden. Es gilt offenbar*

$$\lim_{t\to 0} \frac{f(t\,\cos\phi, t\,\sin\phi) \;-\; f(0,0)}{t} \;=\; \cos^2\phi \,/\, \sin\phi \qquad \text{falls} \quad \sin\phi \neq 0\,,$$

und der Grenzwert ist gleich 0 *falls* $\sin\phi = 0$ *ist. Somit existiert die Ableitung in Richtung von* e *und hat den angegebenen Wert. Also existieren alle Richtungsableitungen im Punkt* $x = 0$*. Insbesondere ist* f *im Nullpunkt partiell differenzierbar. In* Aufgabe 4.5.4 *wurde aber gezeigt, dass* f *im Nullpunkt nicht stetig ist!*

Aufgabe 5.1.4 *Sei* $p = (p_1, \ldots, p_n)$ *ein Multi-Index, und sei* $f(x) = x^p$*. Zeige, dass* f *in jedem Punkt* $x \in \mathbb{R}^n$ *partiell differenzierbar ist, und dass* $f_{x_k}(x) = p_k\, x^{p-e_k}$ *ist – auch für* $p_k = 0$*, wenn wir davon absehen, dass dann* $p - e_k$ *kein Multi-Index mehr ist.*

Aufgabe 5.1.5 *Sei* $A = [\alpha_{jk}]$ *eine reelle, nicht notwendigerweise symmetrische* (n,n)*-Matrix, und sei* $f(x) = x^T A\, x$*. Zeige, dass* f *in jedem Punkt* $x \in \mathbb{R}^n$ *partiell differenzierbar ist, und dass* $f_{x_k}(x) = \sum_{j=1}^n (\alpha_{jk} + \alpha_{kj})\, x_j$ *ist. Überlege, wie man auf möglichst einfache Weise den Gradienten von* f *schreiben kann.*

Definition 5.1.6 *Falls* f *in jedem Punkt von* G *partiell differenzierbar ist, und falls alle partiellen Ableitungen auf* G *stetig sind, sagen wir kurz, dass* f auf G stetig partiell differenzierbar ist. *Falls es zu jedem* $x_0 \in G$ *einen Radius* $r > 0$ *mit* $K(x_0, r) \subset G$ *so gibt, dass* f *auf* $K(x_0, r)$ *einer Lipschitzbedingung genügt, so sagen wir dass* f *auf* G lokal eine Lipschitzbedingung *erfüllt.*

Lemma 5.1.7 *Falls* f *auf* G *stetig partiell differenzierbar ist, dann erfüllt* f *auf* G *lokal eine Lipschitzbedingung und ist somit dort stetig.*

Beweis: Sei $x_0 \in G$, und sei $r > 0$ so, dass $K(x_0, 2\,r) \subset G$ ist. Dann gibt es ein $C_{x_0} > 0$ mit

$$|f_{x_k}(x)| \;\le\; C_{x_0} \qquad \forall\; k = 1, \ldots, n, \quad \forall\; x \in \overline{K(x_0, r)}.$$

Für $k = 0, \ldots, n$ sei $y_k = (x_1, \ldots, x_k, x_{k+1}^{(0)}, \ldots, x_n^{(0)})^T$, also $y_0 = x_0$, $y_n = x$, und $y_k = y_{k-1} + (x_k - x_k^{(0)})\, e_k$ für $k = 1, \ldots, n$. Dann gibt es nach dem ersten Mittelwertsatz der Differenzialrechnung zu jedem $x \in K(x_0, r)$ ein θ_k zwischen $x_k^{(0)}$ und x_k mit

$$|f(y_k) - f(y_{k-1})| \;=\; |x_k - x_k^{(0)}|\, |f_{x_k}(y_k + \theta_k\, e_k)| \;\le\; |x_k - x_k^{(0)}|\, C_{x_0}\,.$$

Daraus folgt

$$|f(x) - f(x_0)| \;\le\; \sum_{k=1}^n |f(y_k) - f(y_{k-1})| \;\le\; C_{x_0}\, \|x - x_0\|_1 \;\le\; C_{x_0}\, \sqrt{n}\, \|x - x_0\|_2\,,$$

wobei die letzte Abschätzung aus Aufgabe 1.1.3 folgt. Dies ist die Lipschitzbedingung auf $K(x_0, r)$. □

Lemma 5.1.8 (Vertauschen von Integral und Ableitung) *Sei* $n = 2$*, sei* f *auf* G *nach* x_2 *partiell differenzierbar, und seien* f *und* f_{x_2} *auf* G *stetig. Seien weiter* $a < b$ *und* $c < d$ *so, dass* $[a,b] \times [c,d] \subset G$ *ist, und sei*

$$g(x) \;=\; \int_a^b f(t, x)\, dt \qquad \forall\; x \in [c,d].$$

Dann ist g *auf* $[c,d]$ *stetig differenzierbar, und es gilt*

$$g'(x) \;=\; \int_a^b f_{x_2}(t, x)\, dt \qquad \forall\; x \in [c,d].$$

Beweis: Sei $x \in [c,d]$, und sei $\tau \neq 0$ so, dass $x+\tau \in [c,d]$ ist. Nach dem ersten Mittelwertsatz der Differenzialrechnung gibt es zu jedem $t \in [a,b]$ ein θ zwischen x und $x+\tau$, so dass

$$\frac{f(t,x+\tau) \;-\; f(t,x)}{\tau} \;=\; f_{x_2}(t,\theta)\,.$$

Da aus der Stetigkeit von f_{x_2} auf $[a,b]\times[c,d]$ die gleichmäßige Stetigkeit folgt, gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$ so, dass für $|\tau| < \delta$ und beliebiges $t \in [a,b]$ folgt

$$\Big|\, \frac{f(t,x+\tau) - f(t,x)}{\tau} \;-\; f_{x_2}(t,x) \,\Big| \;=\; |f_{x_2}(t,\theta) \;-\; f_{x_2}(t,x)| < \varepsilon\,.$$

Daraus folgt durch Integration über t, dass für alle diese x und τ gilt

$$\Big|\, \frac{g(x+\tau) - g(x)}{\tau} \;-\; \int_a^b f_{x_2}(t,x)\,dt \,\Big| \;<\; \varepsilon(b-a).$$

Dies ergibt die Behauptung. □

Bemerkung 5.1.9 *Beachte, dass obiges Lemma sinngemäß auch dann gilt, wenn die Variable x ein Vektor ist; in diesem Fall kann man ja alle x_j bis auf eines festhalten und nach der verbleibenden Variablen partiell differenzieren.*

Definition 5.1.10 *Sei $\alpha \in \mathbb{N}_0$. Die durch das Integral*

$$J_\alpha(x) \;=\; \frac{1}{\pi} \int_0^\pi \cos(x\,\sin t - \alpha\, t)\,dt \qquad \forall\; x \in \mathbb{R}$$

definierte Funktion heißt Besselsche Funktion *vom Index α. Mit dem obigen Lemma folgt, dass diese Funktion beliebig oft differenzierbar ist.*

Aufgabe 5.1.11 *Zeige, dass die Besselfunktion vom Index $\alpha \in \mathbb{N}_0$ eine Lösung der Differenzialgleichung*

$$x^2\,y'' \;+\; x\,y' \;+\; (x^2-\alpha^2)\,y \;=\; 0$$

ist. Diese Gleichung heißt auch Besselsche Differenzialgleichung.

5.2 Vertauschen der Differenziationsreihenfolge

Definition 5.2.1 *Wenn f auf G nach x_k partiell differenzierbar ist, und wenn f_{x_k} in einem Punkt x nach x_j partiell differenzierbar ist, dann heißt diese Ableitung auch* die zweite partielle Ableitung von f nach x_k und x_j im Punkt x, *und wir schreiben auch hierfür*

$$f_{x_k x_j}(x) \;=\; \frac{\partial^2 f(x)}{\partial x_j\,\partial x_k} \;=\; \frac{\partial^2}{\partial x_j\,\partial x_k}\, f(x)\,.$$

Entsprechend werden höhere als zweite partielle Ableitungen definiert. Im Allgemeinen muss man hier auf die Reihenfolge der Differenziationen achten!

Beispiel 5.2.2 *Für $x_1, x_2 \in \mathbb{R}$ sei*

$$f(x_1,x_2) \;=\; \begin{cases} \dfrac{x_1\,x_2^3}{x_1^2+x_2^2} & \text{für } (x_1,x_2) \neq (0,0), \\[2ex] 0 & \text{für } (x_1,x_2) = (0,0). \end{cases}$$

Dann sind

$$f_{x_1}(x_1,x_2) \;=\; \begin{cases} \dfrac{(x_2^2-x_1^2)\,x_2^3}{(x_1^2+x_2^2)^2} & \textit{für } (x_1,x_2)\neq(0,0),\\[2mm] 0 & \textit{für } (x_1,x_2)=(0,0),\end{cases}$$

$$f_{x_2}(x_1,x_2) \;=\; \begin{cases} \dfrac{(3\,x_1^2+x_2^2)\,x_1\,x_2^2}{(x_1^2+x_2^2)^2} & \textit{für } (x_1,x_2)\neq(0,0),\\[2mm] 0 & \textit{für } (x_1,x_2)=(0,0).\end{cases}$$

Daraus ergibt sich $f_{x_1x_2}(0,0)=1$, $f_{x_2x_1}(0,0)=0$.

Lemma 5.2.3 *Sei* $n=2$, $J=[a_1,a_2]\times[b_1,b_2]\subset G$, *mit* $a_1<a_2$, $b_1<b_2$. *Sei* f *auf* G *partiell differenzierbar, und seien* f *und seine beiden Ableitungen auf* J *stetig. Sei weiter*

$$\square\, f \;=\; f(a_2,b_2) \;+\; f(a_1,b_1) \;-\; f(a_1,b_2) \;-\; f(a_2,b_1)\,.$$

Falls $f_{x_1x_2}$ *auf* $\mathring{J}$ *existiert, dann gibt es ein* $(\xi,\eta)^T\in\mathring{J}$ *mit*

$$\square\, f \;=\; (a_2\;-\;a_1)\,(b_2\;-\;b_1)\,f_{x_1x_2}(\xi,\eta)\,.$$

Eine entsprechende Aussage, aber im Allgemeinen mit anderem (ξ,η), *gilt auch für* $f_{x_2x_1}$.

Beweis: Wir beweisen nur die erste Aussage; die zweite folgt analog oder durch Vertauschen der Variablen x_1 und x_2. Sei $g(x_1)=f(x_1,b_2)-f(x_1,b_1)$, also $\square\, f=g(a_2)-g(a_1)$. Eine Anwendung des ersten Mittelwertsatzes der Differenzialrechnung ergibt die Existenz von $\xi\in(a_1,a_2)$ mit $\square\, f=(a_2-a_1)\,g'(\xi)=(a_2-a_1)\,(f_{x_1}(\xi,b_2)-f_{x_1}(\xi,b_1))$. Erneute Anwendung auf $h(x_2)=f_{x_1}(\xi,x_2)$ liefert die Existenz von $\eta\in(b_2,b_1)$ mit $(f_{x_1}(\xi,b_2)-f_{x_1}(\xi,b_1))=(b_2-b_1)\,f_{x_1x_2}(\xi,\eta)$, und daraus folgt die Behauptung. □

Satz 5.2.4 (Satz von Schwarz) *Seien* $j,k\in\{1,\ldots,n\}$, *mit* $j\neq k$, *gegeben. Sei* f *in* G *nach* x_j *und* x_k *partiell differenzierbar, und seien* f *und diese beiden partiellen Ableitungen in* G *stetig. Ferner existiere* $f_{x_jx_k}$ *auf ganz* G *und sei stetig in einem Punkt* $x_0\in G$. *Dann existiert auch* $f_{x_kx_j}$ *im Punkt* x_0, *und es gilt*

$$f_{x_jx_k}(x_0) \;=\; f_{x_kx_j}(x_0)\,.$$

Beweis: O. B. d. A. sei $n=2$ und $j=1$, $k=2$. Falls $f_{x_1x_2}(x_0)\neq0$ sein sollte, können wir zu $f(x_1,x_2)-c\,x_1\,x_2$ übergehen. Deshalb sei jetzt $f_{x_1x_2}(x_0)=0$ angenommen. Wegen der Stetigkeitsvoraussetzung gibt es zu jedem $\varepsilon>0$ ein $\delta>0$ mit $|f_{x_1x_2}(x)|\le\varepsilon$ für $x=(x_1,x_2)^T\in K(x_0,\delta)$. Für $J\subset K(x_0,\delta)$ folgt nach obigem Lemma, dass $|\square f|\le\varepsilon\,(a_2-a_1)\,(b_2-b_1)$ ist. Wegen $\square f=(b_2-b_1)\,(f_{x_2}(a_2,\zeta)-f_{x_2}(a_1,\zeta))$ folgt hieraus

$$\left|\frac{f_{x_2}(a_2,\zeta)\;-\;f_{x_2}(a_1,\zeta)}{a_2\;-\;a_1}\right| \quad<\quad \varepsilon$$

für ein $\zeta\in(b_1,b_2)$. Für $x_0=(x_1^{(0)},x_2^{(0)})^T$ wählen wir jetzt $b_1=x_2^{(0)}$, $b_2=x_2^{(0)}+1/n$, mit genügend großem $n\in\mathbb{N}$, so dass $J\subset K(x_0,\delta)$ ist. Wenn $n\to\infty$ geht, folgt daraus, dass

$$\left|\frac{f_{x_2}(a_2,x_2^{(0)})\;-\;f_{x_2}(a_1,x_2^{(0)})}{a_2\;-\;a_1}\right| \quad\le\quad \varepsilon.$$

Setzt man nun $a_1=x_1^{(0)}$ und $a_2=x_1^{(0)}+h$ oder umgekehrt, je nachdem, ob $h>0$ ist oder nicht, so folgt, dass $|f_{x_2}(x_1^{(0)}+h,x_2^{(0)})\;-\;f_{x_2}(x_1^{(0)},x_2^{(0)})|/|h|\le\varepsilon$, wenn nur h genügend klein ist. Daraus folgt aber $f_{x_2x_1}(x_0)=0$. □

Aufgabe 5.2.5 *Überprüfe die Behauptungen in* Beispiel 5.2.2.

Aufgabe 5.2.6 *Berechne alle zweiten partiellen Ableitungen von*

$$f(x_1, x_2) = x_1^3 \cos(x_1^2 + x_2).$$

Aufgabe 5.2.7 *Berechne alle zweiten partiellen Ableitungen von f aus* Aufgabe 5.1.5.

5.3 Totale Differenzierbarkeit

Definition 5.3.1 *Wir nennen f* im Punkt $x_0 \in G$ total differenzierbar, *wenn es eine lineare Abbildung $\ell : \mathbb{R}^n \longrightarrow \mathbb{R}$ gibt, sodass für alle genügend kleinen $h \in \mathbb{R}^n$, für welche dann insbesondere $x_0 + h \in G$ liegt, gilt*

$$f(x_0 + h) \;=\; f(x_0) \;+\; \ell(h) \;+\; \|h\|\, r(h), \qquad r(h) \longrightarrow 0 \quad (h \to 0).$$

Falls dies für alle $x_0 \in G$ gilt, heißt f kurz auf G total differenzierbar, *oder einfach* auf G differenzierbar.

Jede lineare Abbildung $\ell : \mathbb{R}^n \longrightarrow \mathbb{R}$ ist nach Resultaten aus der linearen Algebra von der Form $\ell(h) = a\, h$ mit einem Zeilenvektor $a = (a_1, \ldots, a_n)$, d. h., einer $(1, n)$-Matrix. Diese Matrix heißt dann die Ableitung *oder* das totale Differenzial *von f im Punkt x_0, und wir schreiben $f'(x_0)$ an Stelle von a. Der Graph von $g(x) = f'(x_0)\,(x - x_0) + f(x_0)$ ist eine lineare Mannigfaltigkeit in $\mathbb{R}^{n+1}$ und heißt, etwas ungenau,* die Tangentialhyperebene *zu f im Punkt x_0.*

Bemerkung 5.3.2 *Man sieht leicht mit obiger Definition, dass die Summe zweier total differenzierbarer Funktionen selbst wieder total differenzierbar ist.*

Beispiel 5.3.3 *Ein Polynom $p \in \mathbb{R}[x]$ höchstens ersten Grades kann immer geschrieben werden als $p(x) = a\,x + b$, mit einem Zeilenvektor a und einer reellen Zahl b. Beachtet man, dass $p(x_0 + h) = p(x_0) + a\,h$ gilt, so folgt direkt aus der Definition, mit $r(h)$ gleich der Nullfunktion, dass p an jeder Stelle von $\mathbb{R}^n$ total differenzierbar ist, und dass $f'(x_0) = a$ ist, unabhängig von x_0.*

Aufgabe 5.3.4 *Für eine n-reihige quadratische Matrix A sei $f(x) = x^T A\, x$ die zugehörige quadratische Form. Zeige, dass dann für $x_0, h \in \mathbb{R}^n$ gilt*

$$f(x_0 + h) \;=\; f(x_0) \;+\; x_0^T (A + A^T)\, h \;+\; \|h\|\, e^T(h)\, A\, h,$$

mit einem Einheitsvektor $e(h)$, nämlich $e(h) = \|h\|^{-1} h$. Leite daraus her, dass f im Punkt x_0 total differenzierbar ist, und dass $f'(x_0) = x_0^T (A + A^T)$ ist.

Satz 5.3.5 (Berechnung der Ableitung) *Sei f in $x_0 \in G$ total differenzierbar. Dann ist f dort sowohl stetig als auch partiell differenzierbar, und es gilt*

$$f'(x_0) \;=\; \operatorname{grad}\, f(x_0)\,.$$

Ferner existiert auch für jeden Einheitsvektor e die Ableitung von f in Richtung von e, und es ist

$$f_e(x_0) \;=\; \operatorname{grad}\, f(x_0)\, e.$$

Beweis: Die Stetigkeit folgt direkt aus der Definition der totalen Differenzierbarkeit. Setzt man dort $h = t\,e$ mit kleinem $t \in \mathbb{R}$ ein, so folgt wegen $\|h\| = |t|$

$$\frac{f(x_0 + t\,e) \;-\; f(x_0)}{t} \;=\; f'(x_0)\,e \pm r(t\,e) \;\longrightarrow\; f'(x_0)\,e \qquad (t \to 0),$$

und deshalb existiert die Ableitung von f in Richtung von e und ist gleich $f'(x_0)\,e$. Für $e = e_k$ folgt, dass f im Punkt x_0 nach x_k partiell differenzierbar ist, und $f_{x_k}(x_0) = f'(x_0)\,e_k$. Also ist $\operatorname{grad} f(x_0) = f'(x_0)$. □

Bemerkung 5.3.6 *Wir haben gezeigt, dass die Richtungsableitung einer total differenzierbaren Funktion das innere Produkt aus dem Gradienten und dem Einheitsvektor e ist. Dieses innere Produkt ist am größten, wenn e die gleiche Richtung wie der Gradient von f hat. Daraus ergibt sich folgende Interpretation des Gradienten:*

Der Gradient einer total differenzierbaren Funktion f zeigt in Richtung des stärksten Anstieges der Funktionswerte von f.

Die Definition der totalen Differenzierbarkeit ist nicht leicht zu überprüfen. Der folgende Satz gibt aber ein bequemes Kriterium für die Differenzierbarkeit auf dem Gebiet G:

Satz 5.3.7 *Sei f auf G stetig partiell differenzierbar. Dann ist f dort auch total differenzierbar, und die Ableitung f' ist dann stetig auf G.*

Beweis: Die Stetigkeit von f' ist klar auf Grund der Definition der Ableitung. Sei $g(x) = f(x) - \operatorname{grad} f(x_0)\,(x - x_0)$. Dann ist g ebenfalls auf G stetig partiell differenzierbar, und die totale Differenzierbarkeit von f ist wegen Beispiel 5.3.3 und Bemerkung 5.3.2 äquivalent zu der von g. Es gilt aber $\operatorname{grad} g(x_0) = 0$, und deshalb folgt aus der Stetigkeit der partiellen Ableitungen, dass diese in der Nähe von x_0 alle klein sind. Also ergibt sich wie im Beweis von Lemma 5.1.7, dass es zu jedem $\varepsilon > 0$ einen Radius $r > 0$ gibt mit $|g(x) - g(x_0)| \le \varepsilon\,\|x - x_0\|$ für alle $x \in K(x_0, r)$. Deshalb ist für $h = x - x_0 \neq 0$ und $r(h) = \|h\|^{-1}(g(x_0 + h) - g(x_0))$ klar, dass $r(h) \to 0$ für $h \to 0$ gilt. Das aber ist die Behauptung. □

Definition 5.3.8 *Sei jetzt $f = (f_1, \ldots, f_m) : G \longrightarrow \mathbb{R}^m$ eine vektorwertige Funktion auf G. Dann nennen wir f in einem Punkt $x_0 \in G$ oder auf G total oder nach x_k partiell differenzierbar, wenn dasselbe für jede der Koordinatenfunktionen f_j zutrifft. Wir setzen dann*

$$f_{x_k}(x_0) \;=\; \frac{\partial f}{\partial x_k}(x_0) \;=\; \left(\frac{\partial f_1}{\partial x_k}(x_0), \ldots, \frac{\partial f_m}{\partial x_k}(x_0)\right)^T.$$

Eine partielle Ableitung ist also jetzt ein Spaltenvektor, und entsprechend ist die Ableitung von f gleich der Matrix

$$f'(x_0) \;=\; \frac{\partial f}{\partial x}(x_0) \;=\; \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \frac{\partial f_1}{\partial x_2}(x_0) & \ldots & \frac{\partial f_1}{\partial x_n}(x_0) \\ \frac{\partial f_2}{\partial x_1}(x_0) & \frac{\partial f_2}{\partial x_2}(x_0) & \ldots & \frac{\partial f_2}{\partial x_n}(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(x_0) & \frac{\partial f_m}{\partial x_2}(x_0) & \ldots & \frac{\partial f_m}{\partial x_n}(x_0) \end{bmatrix}$$

Man nennt diese Matrix auch Funktionalmatrix *oder* Jacobi-Matrix *von f. Für $m = n$ ist die Funktionalmatrix quadratisch, und ihre Determinante heißt* Funktionaldeterminante. *Außer für $m = 1$ ist es nicht üblich, vom Gradienten von f zu sprechen.*

Aufgabe 5.3.9 *Zeige, dass Polynome und rationale Funktionen auf ihrem natürlichen Definitionsbereich überall total differenzierbar sind.*

Aufgabe 5.3.10 *Zeige, dass* affine Abbildungen $f(x) = A\,x + b$, *mit einer reellen* (m,n)*-Matrix* A *und einem* $b \in \mathbb{R}^m$, *überall total differenzierbar sind, und dass* $f'(x) = A$ *ist, für alle* $x \in \mathbb{R}^n$.

5.4 Die Kettenregel

Satz 5.4.1 (Kettenregel) *Seien* $f : G \longrightarrow G_1 \subset \mathbb{R}^m$ *und* $g : G_1 \longrightarrow \mathbb{R}^p$ *gegeben, und seien* f *in* $x_0 \in G$ *und* g *in* $y_0 = f(x_0) \in G_1$ *total differenzierbar. Dann ist* $\ell = g \circ f$ *in* x_0 *total differenzierbar, und es gilt*

$$\ell'(x_0) \;=\; g'(f(x_0))\; f'(x_0)\,.$$

Beweis: Aus der Definition der totalen Differenzierbarkeit folgt

$$f(x_0 + h) \;=\; f(x_0) \;+\; f'(x_0)\,h \;+\; \|h\|\,r_f(h), \qquad r_f(h) \longrightarrow 0 \quad (h \to 0),$$

$$g(y_0 + h) \;=\; g(y_0) \;+\; g'(y_0)\,h \;+\; \|h\|\,r_g(h), \qquad r_g(h) \longrightarrow 0 \quad (h \to 0),$$

wobei in der ersten (zweiten) Gleichung $h \in \mathbb{R}^n$ ($h \in \mathbb{R}^m$) zu nehmen ist. Daraus folgt

$$\ell(x_0 + h) \;=\; g(y_0 + \tilde h) \;=\; g(y_0) \;+\; g'(y_0)\,\tilde h \;+\; \|\tilde h\|\,r_g(\tilde h),$$

mit $\tilde h = f'(x_0)\,h + \|h\|\,r_f(h)$. Dies ist aber äquivalent zu

$$\ell(x_0 + h) \;=\; \ell(x_0) \;+ g'(f(x_0))\,f'(x_0)\,h \;+\; \|h\|\,r_\ell(h),$$

$$r_\ell(h) \;=\; g'(y_0)\,r_f(h) \;+\; \frac{1}{\|h\|}\,\|\tilde h\|\,r_g(\tilde h).$$

Da $\tilde h \longrightarrow 0$ für $h \to 0$, während $\|h\|^{-1}\tilde h$ beschränkt bleibt, folgt insgesamt $r_\ell(h) \to 0$ für $h \to 0$. Das ergibt die Behauptung. □

Korollar zu Satz 5.4.1 *Sei* (a,b) *ein offenes Intervall in* $\mathbb{R}$ *sowie* $G \subset \mathbb{R}^n$ *ein Gebiet, seien* $f_k : (a,b) \to \mathbb{R}$ *für* $k = 1, \ldots, n$, *mit* $f(t) = (f_1(t), \ldots, f_n(t))^T \in G$ *für alle* $t \in (a,b)$, *und sei* $g : G \longrightarrow \mathbb{R}$. *Falls dann alle* f_k *an einer Stelle* $t_0 \in (a,b)$ *differenzierbar sind, und falls* g *im Punkt* $x_0 = (f_1(t_0), \ldots, f_n(t_0))^T$ *total differenzierbar ist, dann ist* $g \circ f$ *in* t_0 *differenzierbar, und es gilt*

$$(g \circ f)'(t_0) \;=\; \operatorname{grad} g(f(t_0))\; f'(t_0) \;=\; \sum_{k=1}^{n} g_{x_k}(f_1(t_0), \ldots, f_n(t_0))\, f_k'(t_0)\,.$$

Beweis: Folgt direkt aus der Kettenregel! □

Als eine Anwendung zeigen wir:

Behauptung 5.4.2 *Seien* $a < b$ *und* $c < d$, *sei* $f(x_1, x_2)$ *auf* $(a,b) \times (c,d)$ *nach* x_2 *partiell differenzierbar, und seien* f *und* f_{x_2} *auf* $(a,b) \times (c,d)$ *stetig. Seien weiter* $g_1, g_2 : (c,d) \longrightarrow (a,b)$ *stetig differenzierbar, und sei schließlich*

$$g(x) \;=\; \int_{g_1(x)}^{g_2(x)} f(t,x)\,dt \qquad \forall\; x \in (c,d).$$

Dann ist g auf (c,d) differenzierbar, und

$$g'(x) \;=\; f(g_2(x),x)\,g_2'(x) \;-\; f(g_1(x),x)\,g_1'(x) \;+\; \int_{g_1(x)}^{g_2(x)} f_{x_2}(t,x)\,dt$$

für alle $x \in (c,d)$.

Beweis: Sei $F(\xi,\eta,\zeta) = \int_\xi^\eta f(t,\zeta)\,dt$, für $\xi,\eta \in (a,b)$, $\zeta \in (c,d)$. Dann ist F nach allen Variablen partiell differenzierbar, und

$$F_\xi(\xi,\eta,\zeta) \;=\; -f(\xi,\zeta), \quad F_\eta(\xi,\eta,\zeta) \;=\; f(\eta,\zeta),$$

bzw. (wegen Lemma 5.1.8)

$$F_\zeta(\xi,\eta,\zeta) \;=\; \int_\xi^\eta f_{x_2}(t,\zeta)\,dt.$$

Da alle partiellen Ableitungen stetig sind, ist F auch total differenzierbar. Durch Anwendung der Kettenregel auf $g(x) = F(g_1(x),g_2(x),x)$ folgt dann die Behauptung. □

Aufgabe 5.4.3 *Sei $G \subset \mathbb{R}^2$ ein Gebiet, und sei $F = F(x,y)$ auf G total differenzierbar. Sei weiter $f(x)$ eine auf (a,b) differenzierbare Funktion, für die $(x,f(x))^T \in G$ und $F(x,f(x))$ konstant ist für alle $x \in (a,b)$. Zeige: Ist $F_y(x,y) \neq 0$ auf G, so ist*

$$f'(x) \;=\; -\,\frac{F_x(x,f(x))}{F_y(x,f(x))} \quad \forall\, x \in (a,b)\,.$$

Aufgabe 5.4.4 *Verallgemeinere obige Aufgabe auf den Fall, dass $x \in \mathbb{R}^n$ ist.*

5.5 Der Mittelwertsatz und der Satz von Taylor

Satz 5.5.1 (Mittelwertsatz) *Sei $f : G \longrightarrow \mathbb{R}$ total differenzierbar auf G, und seien $a,b \in G$ so, dass die Verbindungsstrecke von a nach b ganz zu G gehört. Dann gibt es ein ξ auf dieser Verbindungsstrecke mit*

$$f(b) \;-\; f(a) \;=\; f'(\xi)\;(b \;-\; a)\,.$$

Beweis: Setze $g(t) = f(t\,b + (1-t)\,a)$, $0 \le t \le 1$, und wende den ersten Mittelwertsatz der Differentialrechnung sowie die Kettenregel an. □

Bei vektorwertigen Funktionen gilt im Allgemeinen kein Mittelwertsatz. Als Ersatz zeigen wir aber, dass stetig differenzierbare Funktionen mit beschränkter Ableitung eine Lipschitzbedingung erfüllen:

Satz 5.5.2 (Lipschitzbedingung für stetig differenzierbare Funktionen) *Sei G ein konvexes Gebiet, sei $f : G \longrightarrow \mathbb{R}^m$ stetig partiell differenzierbar auf G, und sei $L = \sup\{\|f'(x)\| : \; x \in G\} < \infty$. Dann gilt*

$$\|f(x_1) \;-\; f(x_2)\| \;\le\; L\,\|x_1 \;-\; x_2\| \qquad \forall\, x_1,x_2 \in G\,.$$

Beweis: Seien $x_1, x_2 \in G$, und sei $g(t) = \|f(t\,x_2 + (1-t)\,x_1) - f(x_1)\|$, für $0 \le t \le 1$. Dieses g ist stetig auf $[0,1]$ und differenzierbar in jedem Punkt t mit $g(t) > 0$. Falls $g(1) = 0$ ist, ist nichts zu zeigen. Im anderen Fall gibt es ein $t_0 \in [0,1)$ mit $g(t_0) = 0$ und $g(t) > 0$ für alle $t \in (t_0, 1]$. Dann ist der erste Mittelwertsatz der Differentialrechnung auf g und das Intervall $[t_0, 1]$ anwendbar, und wir erhalten die Existenz eines $\xi \in (t_0, 1)$ mit

$$g(1) \;=\; g(1) \;-\; g(t_0) \;=\; g'(\xi)\,(1 \;-\; t_0) \;\le g'(\xi)\,.$$

Man zeigt aber mit der Kettenregel, dass

$$g'(t) \;=\; \frac{1}{g(t)}\,(f^T(x(t)) \;-\; f^T(x_1))\,f'(x(t))\,(x_2 - x_1)\,,$$

mit $x(t) = t\,x_2 + (1-t)\,x_1$, falls nur $g(t) > 0$ ist. Daraus folgt aber

$$\|g'(t)\| \;\le\; \|f'(x(t))\|\,\|x_2 - x_1\| \;\le\; L\,\|x_2 - x_1\|\,,$$

und hieraus ergibt sich die Behauptung. □

Im nächsten Satz benutzen wir folgende Bezeichnungen:

$$\nabla \;=\; \left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}\right)$$

bezeichnet *den Gradientenoperator*, d. h. $\nabla f = \operatorname{grad} f$. Für $h = (h_1, \ldots, h_n)^T \in \mathbb{R}^n$ sei

$$\nabla\,h \;=\; \sum_{k=1}^{n} h_k\,\frac{\partial}{\partial x_k},$$

d. h. $\nabla\,h$ ist ein *Differentialoperator*, den man auf jede partiell differenzierbare Funktion f anwenden kann, und $(\nabla\,h)\,f = \sum_{k=1}^{n} h_k\,\frac{\partial f}{\partial x_k}$. Weiter nennen wir eine Funktion $(m+1)$*-mal stetig partiell differenzierbar* auf G, wenn auf G alle ihre partiellen Ableitungen bis zur $(m+1)$-ten Ordnung existieren und stetig sind. Insbesondere kann man dann den Operator $\nabla\,h$ auf jede solche Funktion $(m+1)$-mal anwenden, und wir schreiben dann auch $(\nabla\,h)^k\,f$ für die Funktion, die man durch k-maliges Anwenden dieses Operators auf f erhält. Wir definieren noch $(\nabla\,h)^0\,f = f$.

Satz 5.5.3 (Der Satz von Taylor) *Sei $f : G \longrightarrow \mathbb{R}$ $(m+1)$-mal stetig partiell differenzierbar, und seien x und h so, dass die Verbindungsstrecke von x nach $x+h$ ganz in G liegt. Dann gilt*

$$f(x+h) \;=\; \sum_{k=0}^{m} \frac{(\nabla\,h)^k f(x)}{k!} \;+\; \int_0^1 \frac{(1-t)^m}{m!}\;(\nabla\,h)^{m+1} f(x + t\,h)\,dt.$$

Beweis: Sei $g(t) = f(x + t\,h)$, $0 \le t \le 1$. Mit dem Taylorschen Satz aus Analysis 1 folgt dann

$$g(1) \;=\; \sum_{k=0}^{m} \frac{g^{(k)}(0)}{k!} \;+\; \int_0^1 \frac{(1-t)^m}{m!}\;g^{(m+1)}(t)\,dt.$$

Mit der Kettenregel und vollständiger Induktion zeigt man, dass $g^{(k)}(t) \;=\; (\nabla\,h)^k f(x + t\,h)$ gilt, für $k = 0, \ldots, m+1$, und daraus folgt die Behauptung. □

Definition 5.5.4 *Wir nennen, analog zu den Bezeichnungen aus* Analysis 1*, die Summe*

$$P_m(h,x) \;=\; \sum_{k=0}^{m} \frac{(\nabla\,h)^k f(x)}{k!}$$

das Taylorpolynom *der Ordnung m; dieser Ausdruck ist ein Polynom vom Grade $\le m$ in den Variablen $h_1, \ldots, h_n$. Die Differenz $r_m(x,h) = f(x+h) - P_m(h,x)$ heißt dann auch* das Restglied *der Ordnung m.*

Aufgabe 5.5.5 *Sei, zusätzlich zu den übrigen Voraussetzungen des Mittelwertsatzes, noch angenommen, dass* grad f *auf* G *stetig ist. Zeige, dass dann* f *auf* G *lokal einer Lipschitzbedingung genügt.*

Aufgabe 5.5.6 *Sei* $p = (p_1, \ldots, p_n)$ *ein Multi-Index, und sei*

$$D^p \;=\; \frac{\partial^{p_1}}{\partial x_1^{p_1}} \cdots \frac{\partial^{p_n}}{\partial x_n^{p_n}} \,,$$

d. h., $D^p f$ *erhält man aus* f *durch* p_1*-maliges partielles Differenzieren nach* x_1*,* p_2*-maliges Differenzieren nach* x_2*, u. s. w., wobei stillschweigend vorausgesetzt ist, dass die Reihenfolge des Differenzierens ohne Bedeutung ist, was nach dem Satz von Schwarz sicher gilt, wenn* f *hinreichend oft stetig partiell differenzierbar ist. Zeige:*

$$\frac{(\nabla\, h)^k f(x)}{k!} \;=\; \sum_{|p|=k} \frac{h^p\, D^p f(x)}{p!} \;,$$

wenn man $p! = (p_1!) \cdots (p_n!)$ *setzt.*

Kapitel 6

Implizite Funktionen, lokale Extrema

6.1 Der Banachsche Fixpunktsatz

Definition 6.1.1 *Sei X ein normierter Raum, sei $D \subset X$ nicht leer, und sei $\phi : D \longrightarrow X$. Wir nennen ϕ* eine Kontraktion *auf D, falls ϕ einer Lipschitzbedingung mit einer Lipschitzkonstanten $L < 1$ genügt. Ein solches L heißt dann auch* Kontraktionsparameter*. Ein $x \in D$ heißt* ein Fixpunkt von ϕ, *falls $\phi(x) = x$ ist.*

Satz 6.1.2 (Banachscher Fixpunktsatz) *Sei X ein normierter Raum, sei $D \subset X$ vollständig, und sei $\phi : D \longrightarrow D$ eine Kontraktion auf D. Dann hat ϕ genau einen Fixpunkt $\xi \in D$. Weiter gilt: Ist $L\,(< 1)$ ein Kontraktionsparameter von ϕ, ist $x_0 \in D$, und setzt man $x_{m+1} = \phi(x_m)$ für $m \in \mathbb{N}_0$, so folgt*

$$\|x_m - \xi\| \leq \frac{1}{1-L}\,\|x_{m+1} - x_m\| \leq \frac{L^m}{1-L}\,\|x_1 - x_0\|\,,$$

und daher gilt $\lim x_m = \xi$.

Beweis: Es ist auf Grund der Definition der (x_m)

$$\|x_{m+1} - x_m\| = \|\phi(x_m) - \phi(x_{m-1})\| \leq L\,\|x_m - x_{m-1}\|\,,$$

und daher folgt mit Induktion über m:

$$\|x_{m+1} - x_m\| \leq L^m\,\|x_1 - x_0\| \qquad \forall\, m \in \mathbb{N}_0\,.$$

Für $p \in \mathbb{N}_0$ ist $x_{m+p+1} - x_m = \sum_{k=m}^{m+p} (x_{k+1} - x_k)$, und daraus folgt

$$\|x_{m+p+1} - x_m\| \leq \sum_{k=m}^{m+p} \|x_{k+1} - x_k\| \leq \|x_1 - x_0\| \sum_{k=m}^{\infty} L^k = \frac{\|x_1 - x_0\|\,L^m}{1-L}\,.$$

Da $L < 1$ ist, folgt hieraus, dass (x_m) eine Cauchy-Folge ist, und wegen der Vollständigkeit von D existiert $\xi = \lim_{m\to\infty} x_m \in D$. Da jede Kontraktion auch stetig ist, folgt hieraus $\phi(\xi) = \lim_{m\to\infty} \phi(x_m) = \lim_{m\to\infty} x_{m+1} = \xi$. Also ist ξ ein Fixpunkt von ϕ. Ist $\tilde{\xi}$ ebenfalls ein Fixpunkt, so gilt $\|\tilde{\xi} - \xi\| = \|\phi(\tilde{\xi}) - \phi(\xi)\| \leq L\,\|\tilde{\xi} - \xi\|$, was nur für $\|\tilde{\xi} - \xi\| = 0$ richtig ist. Daher ist also ξ der einzige Fixpunkt von f. Weiter folgt

$$\begin{aligned}
\|x_m - \xi\| &\leq \|x_m - x_{m+1}\| + \|x_{m+1} - \xi\| \\
&= \|x_m - x_{m+1}\| + \|\phi(x_m) - \phi(\xi)\| \\
&\leq \|x_m - x_{m+1}\| + L\,\|x_m - \xi\|\,,
\end{aligned}$$

woraus folgt

$$(1-L)\,\|x_m-\xi\| \;\le\; \|x_m-x_{m+1}\| \;\le\; L^m\,\|x_1 \;-\; x_0\| \qquad \forall\, m\in\mathbb{N}_0\,.$$

Das ist alles, was noch zu zeigen war. □

Bemerkung 6.1.3 *Der für uns interessanteste Fall des Banachschen Fixpunktsatzes betrifft den Raum $X=\mathbb{R}^n$.* Satz 5.5.2 *zeigt, dass eine Funktion ϕ dann eine Kontraktion auf einer konvexen offenen Menge G ist, wenn sie dort stetig partiell differenzierbar und das Supremum der Norm ihrer Ableitung echt kleiner als* 1 *ist. Ist $D=\overline{G}$, und ist ϕ auf D stetig, so sieht man leicht, dass ϕ auch auf D eine Kontraktion mit dem gleichen Kontraktionsparameter wie auf G ist. Ausserdem ist D als abgeschlossene Teilmenge eines vollständigen Raumes selber vollständig. Es ist aber oft nicht leicht zu zeigen, dass ϕ die Menge D in sich abbildet. Deshalb halten wir fest:*

Ist $D=\overline{K(a,r)}$, mit einem $a\in\mathbb{R}^n$ und $r>0$, so ist D abgeschlossen und somit auch vollständig. Ist $\phi:D\longrightarrow X$ eine Kontraktion mit Kontraktionsparameter L, und gilt $\|\phi(a)-a\|\le(1-L)\,r$, so folgt für $x\in D$ (also $\|x-a\|\le r$), dass $\|\phi(x)-a\|\le\|\phi(x)-\phi(a)\|+\|\phi(a)-a\|\le L\,\|x-a\|+(1-L)\,r\le r$. Deshalb folgt $\phi:D\longrightarrow D$, so dass der Banachsche Fixpunktsatz anwendbar ist.

In dieser Form lassen sich die Voraussetzungen des Banachschen Fixpunktsatzes meist am leichtesten nachprüfen.

Aufgabe 6.1.4 *In* Analysis 1 *wurde kurz das Newton-Verfahren zur Nullstellenberechnung vorgestellt. Zeige: Ist $f:(a,b)\longrightarrow\mathbb{R}$ zweimal stetig differenzierbar auf dem offenen Intervall (a,b), und ist $f(x_0)=0$ für ein $x_0\in(a,b)$, aber $f'(x_0)\ne0$, so ist die Funktion $\phi(x)=x-f(x)/f'(x)$ auf einem kleinen Intervall um x_0 herum kontrahierend.*

Aufgabe 6.1.5 *Seien f, x_0 und ϕ wie oben. Zeige die Existenz eines $\varepsilon>0$ derart, dass ϕ das Intervall $(x_0-\varepsilon,x_0+\varepsilon)$ in sich abbildet.*

6.2 Das Newton-Verfahren

Wir wollen nun die mehrdimensionale Version des Newton-Verfahrens zur Berechnung von Lösungen nichtlinearer Gleichungssysteme kennenlernen: Dazu sei $D\subset\mathbb{R}^n$ nicht leer und $f=(f_1,\dots,f_n)^T:D\longrightarrow\mathbb{R}^n$. Die Gleichung $f(x)=0$ ist offenbar eine bequeme Schreibweise für das System von n (im Allgemeinen nicht linearen) Gleichungen

$$f_j(x_1,\dots,x_n) \;=\; 0\,, \qquad 1\le j\le n, \tag{6.2.1}$$

in den *Unbekannten* $x_1,,\dots,x_n$. Jede Lösung der Gleichung $f(x)=0$ heißt dann auch *eine Nullstelle von* f (in D). Um den Banachschen Fixpunktsatz anwenden zu können, muss man eine andere Funktion ϕ finden, deren Fixpunkte gerade die Nullstellen von f sind. Die beim Newton-Verfahren verwendete *Schrittfunktion* ist

$$\phi(x) \;=\; x \;-\; f'(x)^{-1}f(x) \qquad (x\in D). \tag{6.2.2}$$

Dabei muss vorausgesetzt werden, dass D eine offene Menge ist, auf welcher f stetig partiell differenzierbar ist, und dass die Ableitung $f'(x)$ (welche eine n-reihige quadratische Matrix ist) immer invertierbar ist. Wenn f sogar zweimal stetig partiell differenzierbar auf D ist, dann ist ϕ dort wenigstens einmal stetig partiell differenzierbar, und aus $f(\xi)=0$ folgt $\phi'(\xi)=0$. Daher ist ϕ auf einer genügend kleinen Kugel um ξ kontrahierend und bildet diese Kugel in sich ab. Aus dem Banachschen Fixpunktsatz folgt dann die *lokale Konvergenz* des Newton-Verfahrens, d. h. genauer: *Ist x_0 nahe genug bei ξ, und ist $x_{n+1}=\phi(x_n)=x_n-f'(x_n)^{-1}f(x_n)$ für $n\in\mathbb{N}_0$, so konvergiert (x_n) gegen ξ.* Allerdings ist die Berechnung von $f'(x_n)^{-1}$ in jedem Iterationsschritt recht aufwendig, und deshalb benutzt man oft das sogenannte *vereinfachte Newton-Verfahren* mit der Schrittfunktion

$$\phi(x) \;=\; x \;-\; A\, f(x) \qquad (x \in D),$$

wobei A eine feste quadratische Matrix ist. Natürlich wird man die Matrix A geeignet wählen müssen, damit ϕ kontrahierend ist. Im Idealfall wäre $A = f'(\xi)^{-1}$ zu setzen, denn dann ist $\phi'(\xi) = 0$. Allerdings ist uns ξ ja nicht bekannt, sondern soll ja gerade erst berechnet werden.

Für eine spätere Anwendung formulieren wir eine einfache Folgerung aus Bemerkung 6.1.3:

Lemma 6.2.1 *Sei G ein konvexes Gebiet in $\mathbb{R}^n$, sei $f : G \longrightarrow \mathbb{R}^n$ stetig partiell differenzierbar. Sei weiter die Schrittfunktion ϕ des vereinfachten Newton-Verfahrens auf einer abgeschlossenen Kugel $\overline{K(a,r)} \subset G$ eine Kontraktion mit Kontraktionsparameter L (< 1), und gelte $\|A\, f(a)\| \le (1-L)\, r$. Dann hat f in $\overline{K(a,r)}$ genau eine Nullstelle ξ.*

Aufgabe 6.2.2 *Sei D eine offene Menge, und sei $f : D \longrightarrow \mathbb{R}^n$ zweimal stetig partiell differenzierbar ist. Weiter sei $\xi \in D$ so, dass $f(\xi) = 0$, aber $\det f'(\xi) \neq 0$ ist. Sei schließlich ϕ wie in (6.2.2). Zeige, dass dann $\phi'(\xi) = 0$ ist, und folgere hieraus die Existenz eines $r > 0$, für welches ϕ die abgeschlossene Kugel um ξ mit Radius r in sich abbildet und dort kontrahierend ist.*

Aufgabe 6.2.3 *Sei D eine offene Menge, und sei $f : D \longrightarrow \mathbb{R}^n$ stetig partiell differenzierbar ist. Weiter sei $\xi \in D$ so, dass $f(\xi) = 0$, aber $\det f'(\xi) \neq 0$ ist, und sei A eine Matrix mit $\|I - A\, f'(\xi)\| < 1$. Sei schließlich ϕ die Schrittfunktion des vereinfachten Newtonverfahrens mit obiger Matrix A. Zeige, dass dann $\|\phi'(\xi)\| < 1$ ist, und folgere hieraus die Existenz eines $r > 0$, für welches ϕ die abgeschlossene Kugel um ξ mit Radius r in sich abbildet und dort kontrahierend ist.*

6.3 Implizite Funktionen und Umkehrfunktion

Wir wollen nun das Newton-Verfahren verwenden, um die Lösbarkeit eines allgemeinen nichtlinearen Gleichungssystems wie (6.2.1), aber mit weniger Gleichungen als Unbekannten, zu untersuchen. Dazu verwenden wir folgende Bezeichnungen: Für $n, m \in \mathbb{N}$ sei $G \subset \mathbb{R}^{n+m}$ ein Gebiet, und $f = (f_1, \ldots, f_m)^T : G \longrightarrow \mathbb{R}^m$ eine dort definierte Funktion. Dann ist eine Nullstelle von f genau die Lösung eines Systems von m Gleichungen in $n+m$ Unbekannten. Also liegt die Vermutung nahe, dass man für gewisse der Unbekannten (n an der Zahl) beliebige Werte einsetzen kann und die Gleichungen nach den übrigen m Unbekannten auflösen kann. Allerdings können die Gleichungen auch abhängig sein; um dies auszuschließen, definieren wir:

Definition 6.3.1 *Mit obigen Bezeichnungen nennen wir die Gleichung $f(x) = 0$* wohlgestellt*, wenn die Matrix $f'(x)$ für jedes $x \in G$ den Rang m hat.*

Nach welchen der Unbekannten man auflösen kann, ist im Allgemeinen nicht klar. Durch eine geeignete Umnumerierung der Unbekannten können wir uns aber auf den Fall beschränken, dass wir nach den letzten m Unbekannten auflösen wollen. Um diese besser von den anderen zu unterscheiden, fassen wir einen Vektor aus $\mathbb{R}^{n+m}$ hier als ein Paar (x, y) mit $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$, $y = (y_1, \ldots, y_m)^T \in \mathbb{R}^m$ auf. Die Aufgabe lautet also, die Gleichungen

$$f_j(x_1, \ldots, x_n, y_1, \ldots, y_m) \;=\; 0\,, \qquad 1 \le j \le m, \tag{6.3.1}$$

bei gegebenen Werten für $x_1, \ldots, x_n$, nach den Unbekannten $y_1, \ldots, y_m$ aufzulösen.

Falls f partiell differenzierbar ist, schreiben wir

$$\frac{\partial f}{\partial x}(x,y) \;=\; f_x(x,y) \;=\; \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x,y) & \dots & \frac{\partial f_1}{\partial x_n}(x,y) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x,y) & \dots & \frac{\partial f_m}{\partial x_n}(x,y) \end{bmatrix},$$

und analog

$$\frac{\partial f}{\partial y}(x,y) \;=\; f_y(x,y) \;=\; \begin{bmatrix} \frac{\partial f_1}{\partial y_1}(x,y) & \dots & \frac{\partial f_1}{\partial y_m}(x,y) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial y_1}(x,y) & \dots & \frac{\partial f_m}{\partial y_m}(x,y) \end{bmatrix}.$$

Beachte, dass f_y eine quadratische Matrix ist!

In Vektorschreibweise ist (6.3.1) äquivalent zur Gleichung $f(x,y) = 0$. Es kann vorkommen, dass diese Gleichung für jedes x unlösbar ist. Wir wollen aber zeigen: Wenn ein x_0 existiert, für welches die Gleichung lösbar ist, und wenn eine zusätzliche Bedingung an f erfüllt ist, dann ist die Gleichung für alle x in der Nähe von x_0 eindeutig nach y auflösbar. Die Auflösung ergibt dann eine Funktion $y = g(x)$, welche wir eine *implizite Funktion* nennen wollen.

Satz 6.3.2 (Hauptsatz über implizite Funktionen) *Für $n, m \in \mathbb{N}$ sei G ein Gebiet in $\mathbb{R}^{n+m}$, und $f : G \longrightarrow \mathbb{R}^m$ sei stetig partiell differenzierbar auf G. Für ein $(\xi,\eta)^T \in G$ gelte $f(\xi,\eta) = 0$, und die Matrix*

$$\frac{\partial f}{\partial y}(x,y) \;=\; \begin{bmatrix} \frac{\partial f_1}{\partial y_1}(x,y) & \dots & \frac{\partial f_1}{\partial y_m}(x,y) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial y_1}(x,y) & \dots & \frac{\partial f_m}{\partial y_m}(x,y) \end{bmatrix}$$

sei für $(x,y) = (\xi,\eta)$ invertierbar. Dann gibt es $r_1, r_2 > 0$ derart, dass eine eindeutig bestimmte Funktion $g : K(\xi, r_1) \longrightarrow K(\eta, r_2)$ existiert mit

$$\forall\, x \in K(\xi, r_1)\,,\ y \in K(\eta, r_2) : \qquad f(x,y) \;=\; 0 \quad\iff\quad y \;=\; g(x)\ .$$

Weiter gilt: Die Funktion g ist stetig partiell differenzierbar auf $K(\xi, r_1)$, und

$$g'(x) \;=\; -\left(\frac{\partial f}{\partial y}(x, g(x))\right)^{-1} \frac{\partial f}{\partial x}(x, g(x)) \quad \forall\, x \in K(\xi, r_1)\,. \tag{6.3.2}$$

Beweis: Wir bezeichnen die inverse Matrix von $f_y(\xi,\eta)$ mit A und setzen $\phi(x,y) = y - A\,f(x,y)$. Dann ist ϕ nach $y_1, \dots, y_m$ stetig partiell differenzierbar, und $\phi_y(\xi,\eta) = 0$. Da ϕ_y stetig von (x,y) abhängt, existieren $r_1, r_2 > 0$ so, dass $\|\phi_y(x,y)\| \le 1/2$ ist für $\|x-\xi\| < r_1$, $\|y-\eta\| < r_2$. Wenn wir evtl. r_1 verkleinern, gilt $\|A\,f(x,\eta)\| \le r_2/2$ für $\|x-\xi\| < r_1$ (denn f ist stetig und $f(\xi,\eta) = 0$). Wenn wir jetzt ein festes x mit $\|x-\xi\| < r_1$ betrachten, können wir Lemma 6.2.1 auf die Funktion $f(x,\cdot)$ (und $a = \eta$, $L = 1/2$) anwenden und erhalten die Existenz eines eindeutig bestimmten $y = g(x)$ mit $\|y-\eta\| \le r_2$, für welches $f(x,y) = 0$ ist. Um jetzt noch die Differenzierbarkeit von g zu zeigen, seien $(x,y) \in K(\xi,r_1) \times K(\eta,r_2)$ gegeben. Auf jede Komponente f_j des Vektors f kann man den Mittelwertsatz anwenden, allerdings wird die entsprechende Zwischenstelle dabei von j abhängen. Genauer heißt das: Zu (x,y) und j existiert ein Paar $(\tilde{x}_j, \tilde{y}_j)$ auf der Verbindungsstrecke von (ξ,η) und (x,y), für welches gilt

$$f_j(x,y) \;-\; f_j(\xi,\eta) \;=\; \frac{\partial f_j}{\partial x}(\tilde{x}_j, \tilde{y}_j)\,(x-\xi) \;+\; \frac{\partial f_j}{\partial y}(\tilde{x}_j, \tilde{y}_j)\,(y-\eta).$$

Wenn wir diese Gleichungen zu einer Vektorgleichung zusammenfassen und $y = g(x)$ setzen, folgt (wegen $f(x, g(x)) = f(\xi, \eta) = 0$)

$$0 \;=\; f(x, g(x)) \;-\; f(\xi, \eta) \;=\; B(x)\,(x - \xi) \;+\; C(x)\,(g(x) - \eta)\,,$$

mit Matrizen $B(x)$ und $C(x)$, wobei $B(x)$ beschränkt ist, während $C(x)$, jedenfalls für genügend kleine Werte von r_1, r_2, invertierbar ist und eine beschränkte Inverse hat. Deshalb können wir die Gleichung nach $g(x)$ auflösen und erhalten (wegen $g(\xi) = \eta$)

$$g(x) \;=\; g(\xi) \;-\; C(x)^{-1}B(x)\,(x \;-\; \xi)\,. \tag{6.3.3}$$

Wegen der Beschränktheit von $C(x)^{-1}B(x)$ folgt jetzt, dass g im Punkt ξ stetig ist. Daraus folgt wiederum, dass $C(x)$ und $B(x)$ beide im Punkt ξ stetig sind, und $B(\xi) = f_x(\xi, g(\xi))$, $C(\xi) = f_y(\xi, g(\xi))$. Dann ist aber auch $C(x)^{-1}$ im Punkt ξ stetig, und hiermit ergibt sich aus (6.3.3), dass g im Punkt ξ partiell differenzierbar ist, und dass $g'(\xi) = -f_y(\xi, g(\xi))^{-1} f_x(\xi, g(\xi))$ ist. Da wir jetzt aber das Paar (ξ, η) durch jedes andere Paar $(x, g(x))$ ersetzen können, folgt genauso die Gültigkeit von (6.3.2), und aus dieser Gleichung lesen wir die Stetigkeit von g' ab. □

Satz 6.3.3 (Existenz und Ableitung der Umkehrfunktion) *Sei $G \subset \mathbb{R}^n$ ein Gebiet, und sei $f = (f_1, \ldots, f_n)^T : G \longrightarrow \mathbb{R}^n$ stetig partiell differenzierbar auf G. Sei weiter für ein $\eta \in G$*

$$\det f'(\eta) \;=\; \det \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\eta) & \ldots & \frac{\partial f_1}{\partial x_n}(\eta) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(\eta) & \ldots & \frac{\partial f_n}{\partial x_n}(\eta) \end{bmatrix} \;\neq\; 0\,.$$

Dann gibt es Gebiete G_1, G_2 mit $\eta \in G_1 \subset G$, $f(\eta) \in G_2 \subset \mathbb{R}^n$ derart, dass $f : G_1 \longrightarrow G_2$ bijektiv und $\det f'(x) \neq 0$ für alle $x \in G_1$ ist. Die Umkehrfunktion ist dann auf G_2 differenzierbar, und es gilt

$$\frac{\partial f^{-1}}{\partial x}(x) \;=\; \bigl(f'(f^{-1}(x))\bigr)^{-1} \quad \forall\, x \in G_2\,.$$

Beweis: Eine Anwendung des Hauptsatzes über implizite Funktionen auf $f(x, y) = f(y) - x$ und $\xi = f(\eta)$ liefert die Existenz von $r_1, r_2 > 0$, so dass

$$\forall\, x \in K(\xi, r_1) \quad \exists_1\, y \in K(\eta, r_2) \;:\quad x = f(y)\,.$$

Aus Proposition 4.1.4 folgt, dass die Urbildmenge von $K(\xi, r_1)$ eine offene Teilmenge $O \subset G$ ist, und wir setzen $G_1 = O \cap K(\eta, r_2)$. Dann ist $f : G_1 \longrightarrow G_2 = K(\xi, r_1)$ bijektiv. Aus dem oben schon angewandten Hauptsatz folgt aber auch die stetige partielle Differenzierbarkeit der Umkehrfunktion (also auch deren Stetigkeit), und da G_2 zusammenhängend ist, folgt mit Satz 4.2.1 der Zusammenhang von G_1. Die behauptete Aussage über die Ableitung von f^{-1} folgt dann aus (6.3.2). □

Definition 6.3.4 *Sei $G \subset \mathbb{R}^n$ ein Gebiet, und sei $f : G \longrightarrow \mathbb{R}^n$. Wir nennen f* lokal umkehrbar *oder* lokal injektiv, *wenn es zu jedem $x \in G$ ein $r > 0$ gibt, für welches die Restriktion von f auf $K(x, r)$ injektiv ist. Falls f auf G injektiv ist, nennen wir f manchmal auch* global umkehrbar.

Der obige Satz zeigt, dass eine stetig partiell differenzierbare Funktion, deren Funktionaldeterminante nie verschwindet, lokal umkehrbar ist. In Aufgabe 6.3.6 wird ein Beispiel dafür gegeben, dass ein solches f nicht global umkehrbar sein muss.

Aufgabe 6.3.5 *Finde eine einfache Funktion* $f : \mathbb{R}^{n+m} \longrightarrow \mathbb{R}^m$, *für welche die Gleichung* $f(x, y) = 0$ *für jedes* $x \in \mathbb{R}^n$ *unlösbar ist.*

Aufgabe 6.3.6 *Sei* $(r, \phi)^T \mapsto f(r, \phi) = (r\, \cos\phi, r\, \sin\phi)^T$, *für* $r \in \mathbb{R}_+$, $\phi \in \mathbb{R}$. *Zeige: Es ist* $\det f'(r, \phi) \neq 0$ *für alle diese* r, ϕ. *Also ist die Abbildung* f *lokal umkehrbar. Untersuche* f *auf globale Umkehrbarkeit!*

6.4 Lokale Extrema

Definition 6.4.1 *Sei* $G \subset \mathbb{R}^n$ *ein Gebiet, und sei* $f : G \longrightarrow \mathbb{R}$. *Ein* $\xi \in G$ *heißt* lokales Maximum (Minimum) *von* f, *falls ein* $r > 0$ *existiert, für welches*

$$\forall\, x \in G \cap K(\xi, r) \;:\qquad f(x) \;\le\; f(\xi) \quad \bigl(\, f(x) \;\ge\; f(\xi)\bigr)\,.$$

In beiden Fällen nennen wir ξ *dann auch* lokales Extremum *von* f.

Satz 6.4.2 *Sei* $G \subset \mathbb{R}^n$ *ein Gebiet, und habe* $f : G \longrightarrow \mathbb{R}$ *in einem Punkt* $\xi \in G$ *ein lokales Extremum. Falls dann für einen Einheitsvektor* e *die Richtungsableitung* $f_e(\xi)$ *existiert, folgt* $f_e(\xi) = 0$. *Insbesondere folgt* $\operatorname{grad} f(\xi) = 0$, *falls* f *im Punkt* ξ *partiell differenzierbar ist.*

Beweis: Die Funktion $g(t) = f(x + t\, e)$ ist für kleine $|t|$ definiert und hat im Nullpunkt ein lokales Extremum. Also folgt $g'(0) = f_e(x) = 0$. □

Definition 6.4.3 *Sei* $G \subset \mathbb{R}^n$ *ein Gebiet, und sei* $f : G \longrightarrow \mathbb{R}$. *Falls* f *in* $\xi \in G$ *partiell differenzierbar ist, und falls* $\operatorname{grad} f(\xi) = 0$ *ist, heißt* ξ *ein* kritischer *oder* stationärer *Punkt von* f. *Falls* f *in* ξ *zweimal partiell differenzierbar ist, heißt die quadratische Matrix*

$$H_f(\xi) \;=\; \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(\xi) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\xi) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\xi) \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\xi) & \frac{\partial^2 f}{\partial x_n \partial x_2}(\xi) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(\xi) \end{bmatrix}$$

die Hesse-Matrix *von* f *im Punkt* ξ. *Falls* f *auf* G *zweimal stetig partiell differenzierbar ist, ist* H_f *nach dem Satz von Schwarz symmetrisch.*

Satz 6.4.4 *Sei* $G \subset \mathbb{R}^n$ *ein Gebiet, sei* $f : G \longrightarrow \mathbb{R}$ *zweimal stetig partiell differenzierbar auf* G, *und sei* $\xi \in G$ *ein kritischer Punkt von* f. *Dann gilt:*

(a) *Falls* $H_f(\xi)$ *positiv definit ist, ist* ξ *ein lokales Minimum von* f.

(b) *Falls* $H_f(\xi)$ *negativ definit ist, ist* ξ *ein lokales Maximum von* f.

(c) *Falls* $H_f(\xi)$ *indefinit ist, ist* ξ *kein lokales Extremum von* f.

Beweis: Aus dem Satz von Taylor (mit $m = 1$) folgt für genügend kleine $h \in \mathbb{R}^n \setminus \{0\}$

$$f(\xi + h) \;=\; f(\xi) \;+\; h^T \left(\int_0^1 (1 - t)\, H_f(\xi + t\, h)\; dt \right) h\,. \tag{6.4.1}$$

Falls $H_f(\xi)$ positiv definit ist, ist $x^T H_f(\xi)\, x > 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$. Die Funktion $g(t,e) = e^T H_f(\xi + t\,h)\, e$ ist stetig auf der Menge $[0,1] \times \{e \in \mathbb{R}^n : \|e\| = 1\}$. Diese Menge ist kompakt, und deshalb nimmt g dort ein Minimum an, welches positiv sein muss, wenn nur $\|h\|$ klein genug ist. Wenn wir jetzt $e = \|h\|^{-1} h$ setzen, folgt $h^T H_f(\xi + t\,h)\, h > 0$ für alle $t \in [0,1]$, und daher ergibt sich aus (6.4.1), dass $f(\xi + t\,h) > f(\xi)$ ist. Also gilt (a). Der Beweis von (b) ergibt sich aus (a), angewandt auf $-f$. Zu (c): Aus der Indefinitheit folgt die Existenz von Einheitsvektoren $e_\pm$ mit $e_+^T H_f(\xi)\, e_+ > 0$, $e_-^T H_f(\xi)\, e_- < 0$. Analog wie im Beweis von (a) schließt man daraus, dass für $h_\pm = \lambda\, e_\pm$ und alle $t \in [0,1]$ gilt $h_+^T H_f(\xi + t\,h_+)\, h_+ > 0$, $h_-^T H_f(\xi + t\,h_-)\, h_- < 0$, jedenfalls wenn $\|h_\pm\|$ klein genug ist. Daraus folgt die Behauptung. □

Aufgabe 6.4.5 *Sei $G \subset \mathbb{R}^n$ ein Gebiet, und seien $g, h : G \longrightarrow \mathbb{R}$ partiell differenzierbar sowie $h(x) \neq 0$ für alle $x \in G$. Zeige: Genau dann ist $x \in G$ kritischer Punkt von $f = g/h$, wenn* grad $g(x) = f(x)$ grad $h(x)$ *gilt.*

Aufgabe 6.4.6 *Sei A eine n-reihige symmetrische Matrix, und sei $f(x) = \|x\|^{-2}\, x^T A\, x$. Zeige: Die stationären Punkte von f sind genau die Eigenvektoren von A. Man nennt dieses f auch den* Raleigh-Quotienten.

6.5 Extrema unter Nebenbedingungen

Definition 6.5.1 *Sei $G \subset \mathbb{R}^n$ ein Gebiet, und seien $f : G \longrightarrow \mathbb{R}$, $g : G \longrightarrow \mathbb{R}^m$, mit $1 \le m \le n-1$, und sei $\xi \in G$ mit $g(\xi) = 0$ gegeben. Dann heißt ξ* ein lokales Maximum (Minimum) *von f* unter der Nebenbedingung $g(x) = 0$, *falls ein $r > 0$ existiert mit*

$$\forall\ x \in G \cap K(\xi, r) \cap \{x \in G : g(x) = 0\} \ :\quad f(x) \ \le\ f(\xi) \quad \big(\, f(x) \ \ge\ f(\xi) \,\big)\,.$$

Wieder sprechen wir in beiden Fällen auch von einem lokalen Extremum unter der Nebenbedingung $g(x) = 0$.

Satz 6.5.2 (Lagrangesche Multiplikatorregel) *Sei G ein beliebiges Gebiet in $\mathbb{R}^n$, und seien $f : G \longrightarrow \mathbb{R}$, $g = (g_1, \ldots, g_m)^T : G \longrightarrow \mathbb{R}^m$, mit $1 \le m \le n-1$, beide stetig partiell differenzierbar auf G. Sei weiter ξ ein lokales Extremum von f unter der Nebenbedingung $g(x) = 0$, und habe $g'(\xi)$ den Rang m. Dann gibt es ein $\lambda_0 = (\lambda_1^{(0)}, \ldots, \lambda_m^{(0)})^T \in \mathbb{R}^m$ derart, dass die Funktion h, definiert durch*

$$h(x, \lambda) \ =\ f(x) \ +\ \lambda^T g(x) \qquad \forall\ \lambda \in \mathbb{R}^n,\ x \in G,$$

einen kritischen Punkt bei (ξ, λ_0) hat; d. h. also, dass die Gleichungen

$$\frac{\partial f}{\partial x_j}(\xi) \ + \sum_{k=1}^{m} \lambda_k^{(0)}\, \frac{\partial g_k}{\partial x_j}(\xi) \ =\ 0, \quad 1 \le j \le n,$$

$$g_k(\xi) \ =\ 0, \qquad 1 \ \le\ k \le\ n,$$

gelten müssen.

Beweis: Da der Rang von $g'(\xi)$ gleich m ist, existiert mindestens eine m-reihige Untermatrix mit nichtverschwindender Determinante. Für den Beweis können wir o. B. d. A. annehmen, dass diese aus den letzten m Spalten von $g'(\xi)$ besteht. Dann folgt aus dem Hauptsatz über implizite Funktionen, dass die Gleichung $g(x) = 0$ in der Nähe von ξ nach den letzten m Variablen eindeutig aufgelöst werden kann. Das heißt genauer: Ist $p = n - m$, und ist $\tilde{x} = (x_1, \ldots, x_p)^T$, $\tilde{\xi} = (\xi_1, \ldots, \xi_p)^T$, so gibt es ein $r > 0$ und eine auf $K(\tilde{\xi}, r)$ stetig partiell differenzierbare Funktion $k(\tilde{x})$ derart, dass $g(x) = 0$ (für x nahe bei ξ)

genau dann gilt, wenn $x = (\tilde{x}, k(\tilde{x}))^T$ ist. Offenbar hat dann $f(\tilde{x}, k(\tilde{x}))$ als Funktion von $\tilde{x}$ im Punkt $\tilde{\xi}$ ein lokales Extremum im üblichen Sinne, und daher muss gelten

$$\frac{\partial f}{\partial x_j}(\xi) \; + \sum_{\nu=p+1}^{n} \frac{\partial f}{\partial x_\nu}(\xi) \, \frac{\partial k_\nu}{\partial x_j}(\tilde{\xi}) \; = \; 0, \quad 1 \le j \le p.$$

Durch Differenzieren von $g(x) = g(\tilde{x}, k(\tilde{x})) = 0$ erhält man weiter die Gleichungen

$$\frac{\partial g_\mu}{\partial x_j}(\xi) \; + \sum_{\nu=p+1}^{n} \frac{\partial g_\mu}{\partial x_\nu}(\xi) \, \frac{\partial k_\nu}{\partial x_j}(\tilde{\xi}) \; = \; 0, \quad 1 \le \mu \le m, \; 1 \le j \le p.$$

Wenn man $\hat{x} = (x_{p+1}, \ldots, x_n)^T$ setzt, kann man diese Beziehungen wie folgt schreiben:

$$f_{\tilde{x}}(\xi) \; + \; f_{\hat{x}}(\xi) \, k_{\tilde{x}}(\tilde{\xi}) \; = \; 0, \quad g_{\tilde{x}}(\xi) \; + \; g_{\hat{x}}(\xi) \, k_{\tilde{x}}(\tilde{\xi}) \; = \; 0.$$

Dabei ist $g_{\hat{x}}(\xi)$ eine invertierbare Matrix, so dass wir die zweite Gleichung nach $k_{\tilde{x}}(\tilde{\xi})$ auflösen und in die erste einsetzen können. Dies ergibt

$$f_{\tilde{x}}(\xi) \; - \; f_{\hat{x}}(\xi) \, (g_{\hat{x}}(\xi))^{-1} \, g_{\tilde{x}}(\xi) \; = \; 0\,.$$

Mit $\lambda_0 = -f_{\hat{x}}(\xi) \, (g_{\hat{x}}(\xi))^{-1}$ erhält man jetzt $f_{\tilde{x}}(\xi) + \lambda_0 \, g_{\tilde{x}}(\xi) = 0$, und auch $f_{\hat{x}}(\xi) + \lambda_0 \, g_{\hat{x}}(\xi) = 0$. Das ist aber die Behauptung. □

Aufgabe 6.5.3 *Zeige: Die Funktion $f(x, y) = x + y$ nimmt auf der Menge $\{(x, y)^T \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ ein Maximum und ein Minimum an, und bestimme diese Punkte mit der Lagrangeschen Multiplikatorregel.*

Aufgabe 6.5.4 *Finde dasjenige Rechteck vom Umfang 1, dessen Flächeninhalt maximal ist. Gibt es auch eines mit minimalem Flächeninhalt?*

Kapitel 7

Der Jordan-Inhalt

7.1 Definition des Inhalts

Definition 7.1.1 *Seien* $a = (a_1, \ldots, a_n)^T, b = (b_1, \ldots, b_n)^T \in \mathbb{R}^n$ *so, dass* $a_k \leq b_k$ *für alle* $k = 1, \ldots, n$ *gilt. Dann heißt die Menge*

$$I \;=\; [a, b] \;=\; \{x = (x_1, \ldots, x_n)^T : \; a_k \leq x_k \leq b_k \;\forall\; k = 1, \ldots, n\}$$

ein n-dimensinales abgeschlossenes Intervall *oder einfach ein* Intervall*. Wenn $\ell = b_k - a_k$ von k unabhängig ist, nennen wir I auch* einen Würfel*, und ℓ heißt seine* Kantenlänge*. Offenbar ist $I = I_1 \times \ldots \times I_n$ für $I_k = [a_k, b_k]$. Analog kann man auch offene und halboffene Intervalle in $\mathbb{R}^n$ definieren, wir wollen das aber hier nicht tun. Für ein solches Intervall I bezeichnen wir die Zahl*

$$|I| \;=\; \prod_{k=1}^{n} (b_k - a_k)$$

als n-dimensionalen Inhalt *oder* Volumen *von I. Auch die leere Menge wollen wir als Intervall auffassen und ihr den Inhalt 0 zuordnen. Für ein $\alpha \in \mathbb{R}$ und $k \in \{1, \ldots, n\}$ heißt die Menge der Punkte $x \in \mathbb{R}^n$ mit $x_k = \alpha$ eine* achsenparallele Hyperebene*. Wir sagen: Eine solche achsenparallele Hyperebene* zerlegt *ein beliebiges $M \subset \mathbb{R}^n$ in zwei Teilmengen*

$$M_1 = \{x \in M : \; x_k \leq \alpha\}, \quad M_2 = \{x \in M : \; x_k \geq \alpha\},$$

welche im Allgemeinen nicht disjunkt sein werden, und von denen eine auch leer sein kann. Ist $M = I$ ein Intervall, so sind auch die beiden Teile Intervalle I_1, I_2, und man prüft mit der Definition des Inhaltes nach, dass $|I| = |I_1| + |I_2|$ gilt. Zwei Intervalle I_1 und I_2 heißen fremd*, wenn sie keine inneren Punkte gemeinsam haben. Eine Vereinigung von endlich vielen Intervallen heißt* eine Intervallsumme*. Jede solche Intervallsumme S ist auch Vereinigung von paarweise fremden Intervallen $I_1, \ldots, I_N$, und in diesem Fall nennen wir $I_1, \ldots, I_N$ auch* Zerlegung *von S und definieren den Inhalt von S als*

$$|S| \;=\; \sum_{k=1}^{N} |I_k| \,.$$

Eine Intervallsumme besitzt immer unendlich viele verschiedene Zerlegungen. Es ist nicht ganz leicht zu überprüfen, aber jedenfalls richtig, dass die Zahl $|S|$ hierbei nicht von der gewählten Zerlegung abhängt. Außerdem gilt: Sind S_1 und S_2 Intervallsummen mit $S_1 \subset S_2$, so ist $|S_1| \leq |S_2|$. Sei jetzt $M \subset \mathbb{R}^n$ beschränkt und nicht leer. Dann bezeichnen wir das Supremum der Inhalte aller Intervallsummen $S \subset M$ als den inneren Inhalt $|M|_i$ *von M. Analog heißt das Infimum der Inhalte aller Intervallsummen $S \supset M$*

der äußere Inhalt $|M|_a$ *von* M*. Falls der äußere und der innere Inhalt gleich sind, heißt* M Jordan-messbar*, und* $|M|_a = |M|_i$ *heißt* (Jordan-) Inhalt *von* M*. Aus der Definition folgt sofort, dass jede Intervallsumme Jordan-messbar ist, und dass ihr Jordan-Inhalt gleich dem vorher definierten Inhalt ist.*

Bemerkung 7.1.2 *In den Fällen* $n = 2$ *und* $n = 3$ *ist es üblich und anschaulich, statt vom Inhalt einer Menge* M *von deren* Flächeninhalt *bzw.* Volumen *zu sprechen. Aus der Definition ergeben sich folgende einfache Eigenschaften des Inhaltsbegriffes, deren Beweise wir hier auslassen:*

(a) *Für jede beschränkte Menge* $M \subset \mathbb{R}^n$ *ist* $0 \le |M|_i \le |M|_a < \infty$.

(b) *Für beschränkte Mengen* $M_1 \subset M_2 \subset \mathbb{R}^n$ *ist* $|M_1|_i \le |M_2|_i$, $|M_1|_a \le |M_2|_a$*, also im Falle der Jordan-Messbarkeit auch* $|M_1| \le |M_2|$*; man spricht deshalb auch von der* Monotonie *des Inhaltes.*

(c) *Für beschränkte Mengen* $M_1, M_2 \subset \mathbb{R}^n$ *gilt*

$$|M_1 \cup M_2|_a \ \le\ |M_1|_a \ +\ |M_2|_a\,;$$

man nennt diese Tatsache auch die Subadditivität *des äußeren Inhaltes.*

(d) *Für beschränkte Mengen* $M_1, M_2 \subset \mathbb{R}^n$ *mit* $\mathring{M}_1 \cap \mathring{M}_2 = \emptyset$ *gilt*

$$|M_1 \cup M_2|_i \ \ge\ |M_1|_i \ +\ |M_2|_i\,.$$

Aufgabe 7.1.3 *Berechne den äußeren und inneren Inhalt der Menge*

$$M = \{x = (x_1, \ldots, x_n)^T : x_k \in \mathbb{Q},\ 0 \le x_k \le 1 \quad \forall\ k = 1, \ldots, n\},$$

und zeige, dass sie nicht Jordan-messbar ist.

Aufgabe 7.1.4 *Man nennt eine Menge* $M \subset \mathbb{R}^n$ *eine* Lebesguesche Nullmenge*, falls es zu jedem* $\varepsilon > 0$ *abzählbar viele Intervalle* I_k *gibt, so dass gilt*

$$M \subset \bigcup_{k=1}^{\infty} I_k\,, \qquad \sum_{k=1}^{\infty} |I_k| \le \varepsilon.$$

Zeige: Die Menge M *aus der vorigen Aufgabe ist eine Lebesguesche Nullmenge.*

Aufgabe 7.1.5 *Zeige: Die Vereinigung abzählbar vieler Lebesguescher Nullmengen ist wieder eine Lebesguesche Nullmenge.*

Aufgabe 7.1.6 *Finde ein Beispiel dafür, dass der innere Inhalt nicht subadditiv ist.*

7.2 Charakterisierung der Messbarkeit

Definition 7.2.1 *Eine beschränkte Menge* $M \subset \mathbb{R}^n$ *heißt* (Jordan-) Nullmenge*, falls* $|M|_a = 0$ *ist. Es folgt dann, dass auch* $|M|_i = 0$ *ist, und deshalb ist jede Nullmenge auch Jordan-messbar und* $|M| = 0$.

Aufgabe 7.2.2 *Zeige: Die Menge der Glieder einer beliebigen konvergenten Folge in* $\mathbb{R}^n$ *ist eine Nullmenge.*

Satz 7.2.3 *Eine beschränkte Menge* $M \subset \mathbb{R}^n$ *ist genau dann Jordan-messbar, wenn ihr Rand* ∂M *eine Nullmenge ist.*

Beweis: Für $k \in \mathbb{N}$ zerlegt die Menge aller achsenparallelen Hyperebenen der Form $x_j = p_j/2^k$, $p_j \in \mathbb{Z}$, den Raum $\mathbb{R}^n$ in Würfel der Kantenlänge 2^{-k}. Wir setzen $M_{k,i}$ gleich der Vereinigung aller der Würfel, welche ganz zum Inneren von M gehören, und $M_{k,a}$ gleich der Vereinigung aller der Würfel, welche wenigstens einen Punkt von M enthalten. Dann ist die Vereinigung der Würfel, welche zu $M_{k,a}$ aber nicht zu $M_{k,i}$ gehören, eine Obermenge des Randes von M. Außerdem zeigt man mit der Definition von äußerem und innerem Inhalt, dass gilt

$$|M|_i \;=\; \lim_{k\to\infty} |M_{k,i}|\,, \quad |M|_a \;=\; \lim_{k\to\infty} |M_{k,a}|\,.$$

Damit folgt aber $|M|_i + |\partial M|_a = |M|_a$. Dies ergibt die Behauptung. □

Satz 7.2.4 (Eigenschaften des Inhalts) *Für Jordan-messbare $M_1, M_2 \subset \mathbb{R}^n$ gilt:*

(a) *$M_1 \cup M_2$, $M_1 \cap M_2$ und $M_1 \setminus M_2$ sind ebenfalls Jordan-messbar.*

(b) *$|M_1 \cup M_2| \le |M_1| + |M_2|$, und Gleichheit gilt genau dann, wenn die Mengen keine inneren Punkte gemeinsam haben.*

(c) *Falls $M_1 \subset M_2$ gilt, folgt $|M_2 \setminus M_1| = |M_2| - |M_1|$.*

(d) *Ist $B \subset \mathbb{R}^{n-1}$ beschränkt, und ist $f : B \longrightarrow \mathbb{R}$ gleichmäßig stetig auf B, so ist der Graph von f eine Nullmenge.*

Beweis: Zu (a): Ist M eine der Mengen $M_1 \cup M_2$, $M_1 \cap M_1$, $M_1 \setminus M_2$, so gilt in jedem Fall $\partial M \subset (\partial M_1) \cup (\partial M_2)$. Daher folgt die Behauptung mit dem vorigen Satz und der Subadditivität des äußeren Inhaltes. Zu (b): Die Ungleichung folgt aus der Subadditivität des äußeren Inhalts und der Definition der Jordan-Messbarkeit. Die Aussage über die Gleichheit ergibt sich aus Bemerkung 7.1.2 (d). Zu (c): Folgt aus (b) wegen $M_2 = M_1 \cup (M_2 \setminus M_1)$. Zu (d): Sei I ein Intervall in $\mathbb{R}^{n-1}$, welches B enthält. Sei $I = I_1 \cup \ldots \cup I_N$ eine Zerlegung von I, und seien m_j bzw. M_j das Infimum bzw. Supremum von f auf der Menge $I_j \cap B$, falls diese nicht leer ist, bzw. $m_j = M_j = 0$ sonst. Dann ist der Graph von f enthalten in der Vereinigung der n-dimensionalen Intervalle $I_j \times [m_j, M_j]$. Wegen der gleichmäßigen Stetigkeit kann man zu jedem $\varepsilon > 0$ eine Zerlegung von I so finden, dass $M_j - m_j < \varepsilon$ ist, für alle $j = 1, \ldots, N$, und dann folgt, dass der äußere Inhalt des Graphen von f kleiner als $\varepsilon\, c$ ist, wobei c der $(n-1)$-dimensionale Inhalt von I ist. Daraus folgt die Behauptung. □

Beispiel 7.2.5 *Definiere die Menge B durch die Ungleichungen*

$$\begin{array}{rcccl} a_1 & \le & x_1 & \le & b_1\,, \\ a_2(x_1) & \le & x_2 & \le & b_2(x_1)\,, \\ \vdots & & \vdots & & \vdots \\ a_n(x_1, \ldots, x_{n-1}) & \le & x_n & \le & b_n(x_1, \ldots, x_{n-1})\,. \end{array}$$

Dabei seien a_1, b_1 Konstante und die übrigen a_j, b_j stetige Funktionen in den angegebenen Variablen auf den Bereichen, welche durch die vorhergehenden Ungleichungen beschrieben werden. Weiter gelte noch $a_1 \le b_1$, sowie für alle $j = 2, \ldots, n$ und alle Werte der Variablen $a_j(x_1, \ldots, x_{j-1}) \le b_j(x_1, \ldots, x_{j-1})$. Randpunkte dieser Menge sind dann gerade diejenigen $x \in B$, für welche für mindestens eine der Koordinaten x_j ein Gleichheitszeichen gilt. Nach Teil (d) *des vorangegangenen Satzes ist der Rand von B eine Nullmenge, und deshalb ist B Jordan-messbar.*

Lemma 7.2.6 (Bereichsapproximation) *Seien alle C_k, D_k Jordan-messbare Teilmengen von $\mathbb{R}^n$, und sei M so, dass $C_k \subset M \subset D_k$ für $k \in \mathbb{N}$ ist. Falls gilt*

$$\lim_{k\to\infty} |D_k \setminus C_k| \;=\; 0,$$

dann ist auch M Jordan-messbar, und

$$|M| = \lim_{k\to\infty} |C_k| = \lim_{k\to\infty} |D_k| .$$

Beweis: Aus der Monotonie des äußeren und inneren Inhaltes und der Messbarkeit der C_k, D_k folgt $|C_k| \le |M|_i \le |M|_a \le |D_k|$ für alle $k \in \mathbb{N}$. Wegen $D_k = C_k \cup (D_k \setminus C_k)$ folgt $|D_k| = |C_k| + |D_k \setminus C_k|$, und daher gilt die Behauptung. □

Aufgabe 7.2.7 *Zeige: Jede Jordan-Nullmenge ist auch Lebesgue-Nullmenge, aber die Umkehrung gilt nicht.*

Aufgabe 7.2.8 *Zeige: Jede beschränkte Teilmenge einer achsenparallelen Hyperebene ist eine Nullmenge.*

Aufgabe 7.2.9 *Zeige: Jede offene Kugel in $\mathbb{R}^n$ ist Jordan-messbar.*

Aufgabe 7.2.10 *Zeige: Ist M Jordan-messbar, so sind es auch $\overline{M}$ und $\mathring{M}$, und alle drei haben denselben Inhalt. Schließe daraus, dass auch jede Menge C mit $\mathring{M} \subset C \subset \overline{M}$ Jordan-meßbar ist, und dass $|C| = |M|$ ist.*

Aufgabe 7.2.11 *Finde ein Beispiel dafür, dass die Vereinigung abzählbar vieler Jordan-Nullmengen nicht in jedem Fall eine Jordan-Nullmenge sein muss.*

7.3 Berechnung von Inhalten

Bisher können wir nur wenige Inhalte wirklich berechnen. Dies wollen wir jetzt ändern. Dazu zeigen wir zunächst:

Proposition 7.3.1 (Produktregel für den Jordan-Inhalt)
Seien $n, m \in \mathbb{N}$, und seien $M_1 \subset \mathbb{R}^n$, $M_2 \subset \mathbb{R}^m$ Jordan-messbar. Dann ist $M_1 \times M_2$ ebenfalls Jordan-messbar, und es gilt

$$|M_1 \times M_2| = |M_1|\,|M_2| ,$$

wobei das Symbol $|\cdot|$ jeweils für den Inhalt im entsprechenden Raum steht.

Beweis: Wenn M_1 und M_2 Intervalle sind, ist das kartesische Produkt ebenfalls ein Intervall, und die Behauptung gilt per Definition. Als nächstes kann man die Behauptung auch für Intervallsummen verifizieren, und damit erhält man die Gültigkeit im Allgemeinen mittels des Lemmas über die Bereichsapproximation. □

Definition 7.3.2 *Für $M \subset \mathbb{R}^n$ und $t \in \mathbb{R}$ sei*

$$M(t) = \{(x_2, \dots, x_n)^T : (t, x_2, \dots, x_n)^T \in M\} \subset \mathbb{R}^{n-1} .$$

Im wesentlichen ist $M(t)$ also der Durchschnitt von M mit der achsenparallelen Hyperebene $x_1 = t$, allerdings projiziert nach $\mathbb{R}^{n-1}$. Falls M beschränkt ist, ist $M(t) = \emptyset$ für alle t außerhalb eines kompakten Intervalls, und jede der Mengen $M(t)$ ist eine beschränkte Teilmenge von $\mathbb{R}^{n-1}$.

Wir wollen jetzt die Berechnung des Inhalts einer Menge in $\mathbb{R}^n$ auf die von Mengen in $\mathbb{R}^{n+1}$, gefolgt von der Berechnung eines Integrales zurückführen:

Satz 7.3.3 (Satz von Fubini für Inhalte) *Sei $M \subset \mathbb{R}^n$ Jordan-messbar, und seien $\alpha, \beta \in \mathbb{R}$ so, dass $M(t) = \emptyset$ ist für alle $t \notin [\alpha, \beta]$. Dann sind $|M(t)|_i$ und $|M(t)|_a$ beide über $[\alpha, \beta]$ integrierbar, und es gilt*

$$|M| \;=\; \int_\alpha^\beta |M(t)|_i \; dt \;=\; \int_\alpha^\beta |M(t)|_a \; dt\,.$$

Beweis: Wenn M gleich einem Intervall I ist, dann ist $M(t)$ ebenfalls ein Intervall (in $\mathbb{R}^{n-1}$) oder die leere Menge, und man überprüft leicht, dass die Behauptung richtig ist. Daraus folgt dann mit der Additivität von Inhalt und Integral die Gültigkeit für Intervallsummen. Sei jetzt S eine solche Intervallsumme. Wenn $S \subset M$ ist, folgt $S(t) \subset M(t)$, und demnach $|S(t)|_i \le |M(t)|_i \le |M(t)|_a$, für alle $t \in \mathbb{R}$. Wenn man der Einfachheit halber im Folgenden statt dem Ober- (Unter-) Integral nur ein Integralzeichen schreibt, folgt dann

$$|S| \;=\; \int_\alpha^\beta |S(t)|_i \, dt \;\le\; \int_\alpha^\beta |M(t)|_i \;\le\; \int_\alpha^\beta |M(t)|_a \, dt,$$

und deshalb folgt $|M|_i \le \int_\alpha^\beta |M(t)|_i \, dt \;\le \int_\alpha^\beta |M(t)|_a$. Analog folgt aber auch aus $M \subset S$, dass $M(t) \subset S(t)$, also $|M(t)|_i \le |M(t)|_a \le |S(t)|_a$ für alle $t \in \mathbb{R}$. Wenn $[c, d]$ ein Intervall ist, für welches $S(t) = \emptyset$ ist für $t \notin [c, d]$, so gilt (ebenfalls sowohl für das Ober- als auch das Unterintegral)

$$\int_\alpha^\beta |M(t)|_i \, dt \;\le\; \int_\alpha^\beta |M(t)|_a \, dt \;=\; \int_c^d |M(t)|_a \, dt \;\le\; \int_c^d |S(t)|_a \, dt \;= |S|\,,$$

und hieraus folgt $\int_\alpha^\beta |M(t)|_i \, dt \le \int_\alpha^\beta |M(t)|_a \, dt \le |M|_a$. Wegen $|M|_i = |M|_a = |M|$ gilt also die Behauptung. $\square$

Bemerkung 7.3.4 *In obigem Satz haben wir $M(t)$ als Schnitt von M mit der achsenparallelen Hyperebene $x_1 = t$ definiert. Es ist nicht schwer zu sehen, dass ein analoger Satz gilt, wenn man $M(t)$ mittels eines Schnittes mit der Hyperebene $x_j = t$, für irgend ein $j \in \{1, \ldots, n\}$, definiert. Weiter sei betont, dass die Mengen $M(t)$ im Allgemeinen keine Jordan-messbaren Teilmengen von $\mathbb{R}^{n-1}$ sein müssen.*

Aufgabe 7.3.5 *Sei $B \subset \mathbb{R}^2$ beschrieben durch die Ungleichungen*

$$a \;\le\; x_1 \;\le\; b, \quad f(x_1) \;\le\; x_2 \;\le g(x_1),$$

mit Konstanten $a \le b$ und stetigen Funktionen $f, g : [a, b] \longrightarrow \mathbb{R}$, sowie $f(x_1) \le g(x_1)$ für alle $x_1 \in [a, b]$. Zeige die Jordan-messbarkeit von B und berechne seinen Inhalt. Benutze dies, um den Flächeninhalt eines Kreises zu berechnen.

Aufgabe 7.3.6 *Verallgemeinere die vorausgegangene Aufgabe auf drei Dimensionen, und benutze dies, um den Inhalt, also hier das Volumen, einer Kugel in $\mathbb{R}^3$ zu berechnen.*

Aufgabe 7.3.7 *Berechne den Inhalt von Mengen B wie in* Beispiel 7.2.5.

7.4 Bewegungsinvarianz des Inhalts

Der folgende Satz soll hier nicht bewiesen werden. Sein Beweis kann in dem Buch von *W. Walter* [11] nachgelesen werden.

Satz 7.4.1 *Sei A eine quadratische n-reihige reelle Matrix, und sei $b \in \mathbb{R}^n$. Sei ferner $M \subset \mathbb{R}^n$ Jordan-messbar, und sei M_f das Bild von M unter der affinen Abbildung $f(x) = A\,x + b$. Dann ist M_f ebenfalls Jordan-messbar, und es gilt*

$$|M_f| \;=\; |\det A|\;|M|\,.$$

Definition 7.4.2 *Eine affine Abbildung $f(x) = A\,x + b$, mit A, b wie im vorausgegangenen Satz, heißt* eine Bewegung, *falls $|\det A| = 1$ ist.*

Aus dem obigen Satz folgt unmittelbar, dass eine Bewegung den Jordan-Inhalt von Mengen nicht verändert. Man spricht deshalb auch von der *Bewegungsinvarianz* des Jordan-Inhalts.

Aufgabe 7.4.3 *Seien $a_1, \ldots, a_n \in \mathbb{R}^n$ die Spalten einer n-reihigen Matrix quadratischen A, und sei*

$$M \;=\; M(A) \;=\; \{\lambda_1\,a_1 + \ldots + \lambda_n\,a_n :\; 0 \le \lambda_j \le 1 \quad \forall\; j = 1, \ldots, n\}\,.$$

Skizziere $M(A)$ für $n = 2$ und $n = 3$.

Aufgabe 7.4.4 *Berechne den Jordan-Inhalt von $M(A)$, für eine beliebige quadratische Matrix A.*

Aufgabe 7.4.5 *Seien $a, b, c \in \mathbb{R}_+$. Berechne den Jordan-Inhalt des Ellipsoids*

$$M \;=\; \{(x, y, z) \in \mathbb{R}^3 :\; (x/a)^2 + (y/b)^2 + (z/c)^2 \;\le\; 1\;\}\,.$$

Kapitel 8

Das Riemann-Integral

8.1 Die Definition

Im Folgenden sei immer $B \subset \mathbb{R}^n$ Jordan-messbar und nicht leer, und $f : B \longrightarrow \mathbb{R}$ sei auf B beschränkt.

Definition 8.1.1 *Wir nennen $\mathcal{Z} = \{B_1, \ldots, B_N\}$* Zerlegung *von B, wenn alle B_k Jordan-messbar und paarweise fremd sind, und wenn $B = \cup_{k=1}^{N} B_k$ gilt. Als* Feinheit *von $\mathcal{Z}$ bezeichnen wir das Maximum der Durchmesser der B_k. Eine Folge $(\mathcal{Z}_n)$ von Zerlegungen von B heißt* zulässig, *wenn die Feinheit von $\mathcal{Z}_n$ für $n \to \infty$ gegen 0 geht. Eine Zerlegung $\mathcal{Z}_1 = \{B_1^{(1)}, \ldots, B_N^{(1)}\}$ heißt* feiner als, *oder* Verfeinerung von *$\mathcal{Z}_2 = \{B_1^{(2)}, \ldots, B_M^{(2)}\}$, falls es zu jedem $k = 1, \ldots, N$ ein $j \in \{1, \ldots, M\}$ gibt mit $B_k^{(1)} \subset B_j^{(2)}$. Wenn $\mathcal{Z}_1 = \{B_1^{(1)}, \ldots, B_N^{(1)}\}$ und $\mathcal{Z}_2 = \{B_1^{(2)}, \ldots, B_M^{(2)}\}$ beliebig gegeben sind, dann heißt die Zerlegung $\mathcal{Z}_1 + \mathcal{Z}_2$, welche aus allen $B_k^{(1)} \cap B_j^{(2)}$, $1 \leq k \leq N$, $1 \leq j \leq M$, besteht, die* gemeinsame Verfeinerung *oder* Überlagerung *von $\mathcal{Z}_1$ und $\mathcal{Z}_2$. Zu einer Zerlegung $\mathcal{Z} = \{B_1, \ldots, B_N\}$ setzen wir $m_k = \inf\{f(x) : x \in B_k\}$, $M_k = \sup\{f(x) : x \in B_k\}$, und nennen $U(\mathcal{Z}) = \sum_{k=1}^{N} m_k \, |B_k|$ bzw. $O(\mathcal{Z}) = \sum_{k=1}^{N} M_k \, |B_k|$ die zu $\mathcal{Z}$ gehörige* Untersumme *bzw.* Obersumme *von $\mathcal{Z}$. Für beliebige Zwischenpunkte $\xi_k \in B_k$ heißt $S(\mathcal{Z}, \xi) = \sum_{k=1}^{N} f(\xi_k) \, |B_k|$ auch die zu $\mathcal{Z}$ und dem Zwischenpunktvektor $\xi = (\xi_1, \ldots, \xi_N)^T$ gehörige* Riemannsumme von *f. Offenbar gilt immer $U(\mathcal{Z}) \leq S(\mathcal{Z}, \xi) \leq O(\mathcal{Z})$.*

Um Ober- und Unterintegral definieren zu können, zeigen wir:

Lemma 8.1.2 *Für beliebige Zerlegungen $\mathcal{Z}_j$ von B gilt stets $U(\mathcal{Z}_1) \leq O(\mathcal{Z}_2)$.*

Beweis: Es ist leicht zu sehen, dass bei Verfeinerung einer Zerlegung die zugehörigen Obersummen (Untersummen) nicht größer (nicht kleiner) werden können. Deshalb gilt für die Überlagerung $\mathcal{Z}_1 + \mathcal{Z}_2$:

$$U(\mathcal{Z}_1) \;\leq\; U(\mathcal{Z}_1 + \mathcal{Z}_2) \;\leq\; O(\mathcal{Z}_1 + \mathcal{Z}_2) \;\leq\; O(\mathcal{Z}_2).$$

Darum gilt die Behauptung. $\square$

Definition 8.1.3 *Wir nennen*

$$I_* \;=\; \underline{\int_B} f(x)\,dx \;=\sup\{U(\mathcal{Z})\}\,, \qquad I^* \;=\; \overline{\int_B} f(x)\,dx \;=\; \inf\{O(\mathcal{Z})\}\,,$$

wobei das Supremum bzw. Infimum jeweils über alle Zerlegungen von B gebildet wird, das Unter- bzw. Oberintegral von f *über den Bereich B. Offenbar ist $I_* \le I^*$, und wir nennen f* über B integrierbar, *wenn $I_* = I^*$ gilt. Der gemeinsame Wert von Unter- und Oberintegral wird dann mit $\int_B f(x)\,dx$ bezeichnet und heißt* das Bereichsintegral *oder einfach* das Integral von f *über B.*

Beispiel 8.1.4 *Ist f konstant, also $f(x) \equiv c$, so ist f über B integrierbar, und $\int_B f(x)\,dx = c\,|B|$. Wenn $f(x) \ge 0$ ist für $x \in B$, dann ist das Ober- bzw. Unterintegral von f über B gleich dem äußeren bzw. inneren Inhalt der $(n+1)$-dimensionalen Menge $M_{f,B} = \{(x,y)^T : x \in B,\ 0 \le y \le f(x)\}$. Also existiert das Integral von f über B genau dann, wenn $M_{f,B}$ Jordan-messbar ist, und dann ist*

$$\int_B f(x)\,dx \;=\; |M_{f,B}|\,.$$

Dies alles folgt unmittelbar aus der Definition des Bereichsintegrals.

Aufgabe 8.1.5 *Gib für $|B| > 0$ ein Beispiel einer nicht über B integrierbaren Funktion f.*

Aufgabe 8.1.6 (Integration von Ungleichungen) *Zeige: Sind die Funktionen f und g über B integrierbar, und gilt $f(x) \le g(x)$ für $x \in B$, so folgt $\int_B f(x)\,dx \le \int_B g(x)\,dx$.*

8.2 Eigenschaften des Bereichsintegrals

Wir geben jetzt einige Resultate an, deren Beweise entweder direkt aus der entsprechenden Definition folgen oder weitgehend analog zu denen im eindimensionalen Fall und deshalb hier ausgelassen sind:

Proposition 8.2.1

(a) **(Fundamentalabschätzung)** *Für $m = \inf\{\, f(x) :\ x \in B \,\}$, $M = \sup\{\, f(x) :\ x \in B \,\}$ gilt*

$$m\,|B| \;\le\; \underline{\int}_B f(x)\,dx \;\le\; \overline{\int}_B f(x)\,dx \;\le\; M\,|B|\,.$$

Falls f über B integrierbar ist, gilt insbesondere

$$\left| \int_B f(x)\,dx \right| \;\le\; |B|\,\sup\{\,|f(x)| :\ x \in B \,\}\,.$$

(b) **(Linearität des Integrals)** *Sind f und g über B integrierbar, und sind $\alpha, \beta \in \mathbb{R}$, so ist auch $\alpha\,f + \beta\,g$ über B integrierbar, und es gilt*

$$\int_B (\alpha\,f(x) + \beta\,g(x))\,dx \;=\; \alpha \int_B f(x)\,dx \;+\; \beta \int_B g(x)\,dx\,.$$

Korollar zu Proposition 8.2.1 (Integrale über Nullmengen) *Ist B eine Jordansche Nullmenge, so ist jede auf B beschränkte Funktion auch über B integrierbar, und es ist $\int_B f(x)\,dx = 0$.*

Satz 8.2.2

(a) *Für jede zulässige Zerlegungsfolge $(\mathcal{Z}_n)$ konvergieren die Folgen $(U(\mathcal{Z}_n))$ und $(O(\mathcal{Z}_n))$, und es gilt*

$$\underline{\int}_B f(x)\,dx \;=\; \lim_{n\to\infty} U(\mathcal{Z}_n), \qquad \overline{\int}_B f(x)\,dx \;=\; \lim_{n\to\infty} O(\mathcal{Z}_n).$$

(b) *Die Funktion f ist genau dann integrierbar über B, wenn für jede zulässige Folge von Zerlegungen von B und jede Wahl von Zwischenpunktvektoren die zugehörigen Riemannsummen einen Grenzwert haben. Ist dies so, dann ist dieser Grenzwert gleich $\int_B f(x)\,dx$, also unabhängig von der Wahl der Zerlegungsfolge und der Zwischenpunktvektoren.*

Proposition 8.2.3 (Riemannsches Integrabilitätskriterium) *Die Funktion f ist genau dann über B integrierbar, wenn es zu jedem $\varepsilon > 0$ eine Zerlegung $\mathcal{Z}$ von B gibt mit $O(\mathcal{Z}) - U(\mathcal{Z}) < \varepsilon$.*

Mit diesem Kriterium können wir nun folgende wichtige Resultate beweisen:

Satz 8.2.4 (Additivität des Integrals) *Seien A_1, A_2 Jordan-messbar und fremd, und sei $B = A_1 \cup A_2$. Dann gilt*

$$\begin{aligned} \underline{\int}_B f(x)\,dx &= \underline{\int}_{A_1} f(x)\,dx + \underline{\int}_{A_2} f(x)\,dx\,, \\ \overline{\int}_B f(x)\,dx &= \overline{\int}_{A_1} f(x)\,dx + \overline{\int}_{A_2} f(x)\,dx\,. \end{aligned}$$

Insbesondere ist f über B integrierbar genau dann, wenn es sowohl über A_1 als auch über A_2 integrierbar ist, und dann gilt

$$\int_B f(x)\,dx = \int_{A_1} f(x)\,dx + \int_{A_2} f(x)\,dx\,,$$

Beweis: Für zulässige Zerlegungsfolgen $\mathcal{Z}_{n,j}$ von A_j, für $1 \le j \le 2$, ist $\mathcal{Z}_n = \mathcal{Z}_{n,1} \cup \mathcal{Z}_{n,2}$ eine zulässige Zerlegungsfolge von B. Mit Satz 8.2.2 folgen dann die ersten beiden Aussagen. Aus diesen ergibt sich dann direkt die Integrierbarkeit von f über B unter Annahme der Integrierbarkeit über A_1 und A_2. Umgekehrt sei jetzt f über B integrierbar. Dann gibt es nach dem Riemannschen Integrabilitätskriterium eine Zerlegung $\mathcal{Z} = \{B_1, \ldots, B_N\}$ von B mit $O(\mathcal{Z}) - U(\mathcal{Z}) < \varepsilon$. Für $\mathcal{Z}_j = \{B_1 \cap A_j, \ldots, B_N \cap A_j\}$ folgt dann $O(\mathcal{Z}_j) - U(\mathcal{Z}_j) \le O(\mathcal{Z}) - U(\mathcal{Z})$, und hieraus folgt die Integrierbarkeit über A_j, für $1 \le j \le 2$. □

Satz 8.2.5 *Ist f auf B beschränkt, und gibt es eine Jordansche Nullmenge $N \subset B$, für welche f auf $B \setminus N$ stetig ist, so ist f über B integrierbar.*

Beweis: Sei zunächst angenommen, dass f gleichmäßig stetig auf B ist. Sei $\varepsilon > 0$, und sei $\delta > 0$ derart, dass $|f(x_1) - f(x_2)| < \varepsilon$ ist für alle $x_1, x_2 \in B$ mit $\|x_1 - x_2\| < \delta$. Sei $\mathcal{Z} = \{B_1, \ldots, B_N\}$ eine Zerlegung von B, deren Feinheit kleiner als δ ist. Dann folgt $O(\mathcal{Z}) - U(\mathcal{Z}) < \varepsilon\,|B|$, und deshalb ist f über B integrierbar.

Im allgemeinen Fall sei $A = B \setminus N$. Sei S eine Intervallsumme, welche ganz in A liegt. Dann ist f auf S stetig, und da S kompakt ist, folgt gleichmäßige Stetigkeit, also Integrierbarkeit, von f über S. Aus der Additivität des Integrals folgt dann, dass

$$\underline{\int}_B f(x)\,dx - \underline{\int}_S f(x)\,dx = \underline{\int}_{B \setminus S} f(x)\,dx$$

und genauso für die Oberintegrale. Daher ist wegen der Integrierbarkeit von f über S und der Fundamentalabschätzung

$$\begin{aligned} 0 &\le \overline{\int}_B f(x)\,dx - \underline{\int}_B f(x)\,dx = \overline{\int}_{B \setminus S} f(x)\,dx - \underline{\int}_{B \setminus S} f(x)\,dx \\ &\le \left| \overline{\int}_{B \setminus S} f(x)\,dx \right| + \left| \underline{\int}_{B \setminus S} f(x)\,dx \right| \le 2\,M\,|B \setminus S|\,, \end{aligned}$$

mit $M = \sup_{x\in B} |f(x)|$. Wegen $|B \setminus S| = |B| - |S| = |A| - |S|$ und weil man S so wählen kann, dass $|A| - |S|$ beliebig klein wird, folgt die Behauptung. $\square$

Satz 8.2.6 *Sind f und g über B integrierbar, so gilt dasselbe auch für $f\,g$.*

Beweis: Da integrierbare Funktionen beschränkt sein müssen, gibt es ein $K \in \mathbb{R}$ mit $|f(x)|, |g(x)| \le K$ für alle $x \in B$. Also gilt

$$\begin{aligned} |f(x)\,g(x) - f(y)\,g(y)| &\le |f(x)|\,|g(x)-g(y)| + |g(y)|\,|f(x)-f(y)| \\ &\le K\,(|g(x)-g(y)| + |f(x)-f(y)|)\,. \end{aligned}$$

Ist $\mathcal{Z}$ eine beliebige Zerlegung, und stehen U_f, O_f, bzw. U_g, O_g, bzw. U_{fg}, O_{fg} für die zugehörigen Unter- und Obersummen von f, bzw. g, bzw. $f\,g$, so folgt hieraus

$$O_{fg} - U_{fg} \le K\,(O_f - U_f + O_g - U_g).$$

Daraus folgt die Behauptung mit dem Riemannschen Kriterium. $\square$

Lemma 8.2.7 *Sei $f : B \longrightarrow A$, mit $A \subset \mathbb{R}$, über B integrierbar, und gelte für $g : A \to \mathbb{R}$ eine Lipschitzbedingung auf A. Dann ist $g \circ f$ über B integrierbar.*

Beweis: Wird genauso bewiesen wie das entsprechende Lemma in Analysis 1. $\square$

Proposition 8.2.8 (Dreiecksungleichung für Integrale) *Ist f über B integrierbar, dann gilt dasselbe auch für $|f|$, und es gilt*

$$\left|\int_B f(x)\,dx\right| \le \int_B |f(x)|\,dx.$$

Beweis: Für $g(x) = |x|$ folgt aus dem obigen Lemma die Integrierbarkeit von $|f|$, und die Ungleichung folgt aus der Dreiecksungleichung für die Riemannsummen. $\square$

Aufgabe 8.2.9 *Sei f über B integrierbar, sei $N \subset B$ eine Jordansche Nullmenge, und sei g auf B beschränkt und $g(x) = f(x)$ für $x \in B \setminus N$. Zeige, dass auch g über B integrierbar ist und dass $\int_B f(x)\,dx = \int_B g(x)\,dx$ ist.*

Aufgabe 8.2.10 *Sei f über B integrierbar. Zeige die Integrierbarkeit von f^+ und f^-, mit*

$$f^+(x) = \max\{f(x), 0\}, \qquad f^-(x) = \max\{-f(x), 0\} \quad \forall\, x \in B\,.$$

8.3 Mittelwertsätze und gliedweise Integration

Definition 8.3.1 *Sei f über B integrierbar, und sei $|B| > 0$. Die Zahl*

$$\mu = \frac{1}{|B|}\int_B f(x)\,dx$$

***heißt** der Mittelwert von f über B.*

Satz 8.3.2 (Mittelwertsatz der Integralrechnung) *Sei f über B integrierbar, sei $|B| > 0$, und sei μ der Mittelwert von f über B. Dann ist*

$$m = \inf\{f(x) : x \in B\} \;\leq\; \mu \;\leq\; \sup\{f(x) : X \in B\} = M.$$

Beweis: Es ist $m \leq f(x) \leq M$ für alle $x \in B$, und durch Integration folgt die Behauptung. □

Satz 8.3.3 (Erweiterter Mittelwertsatz der Integralrechnung) *Seien g und $f\,g$ über B integrierbar, und sei $g(x) \geq 0$ für alle $x \in B$. Dann existiert ein $\tilde{\mu} \in [m, M]$, mit m, M wie im vorigen Satz, so dass*

$$\int_B f(x)\,g(x)\,dx \;=\; \tilde{\mu} \int_B g(x)\,dx\,.$$

Beweis: Es gilt

$$m \int_B g(x)\,dx \;\leq \int_B f(x)\,g(x)\,dx \;\leq M \int_B g(x)\,dx\,,$$

und daraus folgt die Behauptung. □

Satz 8.3.4 (Gliedweise Integration) *Gegeben seien $f_n, g_k : B \longrightarrow \mathbb{R}$ für alle $n, k \in \mathbb{N}$. Dann gilt*

(a) *Sind alle f_n über B integrierbar, und ist die Funktionenfolge (f_n) auf B gleichmäßig konvergent, so ist auch die Grenzfunktion f über B integrierbar, und es gilt*

$$\int_B f(x)\,dx \;=\; \lim_{n\to\infty} \int_B f_n(x)\,dx.$$

(b) *Sind alle g_k über B integrierbar, und ist die Funktionenreihe $\sum_{k=1}^{\infty} g_k$ auf B gleichmäßig konvergent, so ist auch die Grenzfunktion f über B integrierbar, und es gilt*

$$\int_B f(x)\,dx \;=\; \sum_{k=1}^{\infty} \int_B g_k(x)\,dx.$$

Beweis: Zu (a): O. B. d. A. sei $|B| > 0$. Zu $\varepsilon > 0$ gibt es nach Definition der gleichmäßigen Konvergenz ein $n \in \mathbb{N}$ mit $|f(x) - f_n(x)| \leq \varepsilon/|B|$ für alle $x \in B$. Daraus folgt für jede Zerlegung $\mathcal{Z}$ von B für die Obersummen $O_f(\mathcal{Z}), O_{f_n}(\mathcal{Z})$ und Untersummen $U_f(\mathcal{Z}), U_{f_n}(\mathcal{Z})$ von f bzw. f_n, dass

$$|U_f(\mathcal{Z}) - U_{f_n}(\mathcal{Z})| \;\leq \varepsilon\,, \qquad |O_f(\mathcal{Z}) - O_{f_n}(\mathcal{Z})| \;\leq \varepsilon\,.$$

Nach dem Riemannschen Integrabilitätskriterium existiert eine Zerlegung $\mathcal{Z}$ mit $O_{f_n}(\mathcal{Z}) - U_{f_n}(\mathcal{Z}) < \varepsilon$, und daher folgt

$$\begin{aligned} O_f(\mathcal{Z}) - U_f(\mathcal{Z}) &= O_f(\mathcal{Z}) - O_{f_n}(\mathcal{Z}) + O_{f_n}(\mathcal{Z}) - U_f(\mathcal{Z}) + O_{f_n}(\mathcal{Z}) - U_{f_n}(\mathcal{Z}) \\ &< 3\,\varepsilon\,. \end{aligned}$$

Daraus folgt die Integrierbarkeit von f mit dem Riemannschen Integrabilitätskriterium. Die behauptete Gleichung folgt sofort wegen

$$\left|\int_B (f(x) \;-\; f_n(x))\,dx\right| \;\leq |B| \;\max_{x\in B} |f(x) \;-\; f_n(x)|$$

aus der gleichmäßigen Konvergenz. Der Teil (b) kann wie üblich mit $f_n = \sum_{k=1}^{n} g_k$ auf (a) zurückgeführt werden. □

Aufgabe 8.3.5 *Gib eine Formulierung des Mittelwertsatzes, welche auch für $|B| = 0$ richtig bleibt.*

Aufgabe 8.3.6 *Diskutiere, unter welchen Voraussetzungen über B ein $x_0 \in B$ existiert, für das $\mu = f(x_0)$ ist.*

8.4 Der Satz von Fubini

In diesem Abschnitt seien wieder $B \subset \mathbb{R}^n$ Jordan-messbar und nicht leer, und $f : B \longrightarrow \mathbb{R}$ sei auf B beschränkt. Weiter sei ein $j \in \{1, \ldots, n\}$ fest gewählt.

Definition 8.4.1 *Für ein $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$ setzen wir*

$$y \;=\; (x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)^T$$

und identifizieren x mit dem Paar (x_j, y). Statt $f(x)$ schreiben wir dann auch $f(x_j, y)$. Weiter nennen wir die Menge

$$B(x_j) \;=\; \{y : \quad x = (x_j, y) \in M\} \;\subset\; \mathbb{R}^{n-1}$$

einen Schnitt von B. *Offenbar ist ein Schnitt $B(x_j)$ für beliebiges $x_j \in \mathbb{R}$ eine beschränkte Teilmenge von $\mathbb{R}^{n-1}$, und es gibt ein Intervall $[\alpha, \beta]$ so, dass $B(x_j) = \emptyset$ ist für $x_j \notin [\alpha, \beta]$. Wir nennen die Jordan-messbare Menge B einen* zulässigen Bereich, *wenn alle Schnitte $B(x_j)$ Jordan-messbare Teilmengen von $\mathbb{R}^{n-1}$ sind.*

Beispiel 8.4.2 *Die in* Beispiel 7.2.5 *betrachtete Menge B ist ein zulässiger Bereich für den Fall $j = 1$, denn dann sind alle nicht-leeren Schnitte Jordan-messbare Teilmengen in $\mathbb{R}^{n-1}$, weil sie wieder durch die entsprechenden Ungleichungen, aber mit einem festen Wert für x_1, beschrieben sind.*

Satz 8.4.3 (Satz von Fubini) *Sei B ein zulässiger Bereich in obigem Sinn, und sei f über B integrierbar. Seien ferner $\alpha, \beta \in \mathbb{R}$ so, dass $B(x_j) = \emptyset$ für $x_j \notin [\alpha, \beta]$. Dann gilt*

$$\int_B f(x)\, dx \;=\; \int_\alpha^\beta \left(\underline{\int}_{B(x_j)} f(x_j, y)\, dy \right) dx_j \;=\; \int_\alpha^\beta \left(\overline{\int}_{B(x_j)} f(x_j, y)\, dy \right) dx_j\, .$$

Beweis: Falls $f(x) \geq 0$ ist, ist $\int_B f(x)\, dx$ gleich dem Inhalt der in Beispiel 8.1.4 definierten Menge $M_{f,B}$, und das Unter- bzw. Oberintegral von $f(x_j, y)$ über $B(x_j)$ ist gleich dem inneren bzw. äußeren Inhalt des entsprechenden Schnittes von $M_{f,B}$ mit derjenigen Hyperebene, bei der wir x_j festhalten. Deshalb folgt die Behauptung in diesem Fall aus Satz 7.3.3. Außerdem sieht man leicht, dass die Behauptung ebenfalls gilt, wenn f konstant ist. Den allgemeinen Fall erhält man dann durch Betrachten von $f(x) + c$ mit einer hinreichend großen Konstanten $c \in \mathbb{R}$. □

Aufgabe 8.4.4 *Berechne das Integral $\int_B f(x)\, dx$, wenn $B \subset \mathbb{R}^2$ die Einheitskreisscheibe und $f(x_1, x_2) = x_2 \sin x_1$ ist.*

Aufgabe 8.4.5 *Sei B wie in* Beispiel 7.2.5, *und sei f stetig und beschränkt auf B. Zeige:*

$$\int_B f(x)\, dx \;=\; \int_{a_1}^{b_1} \Big(\int_{a_2(x_1)}^{b_2(x_1)} \ldots \Big(\int_{a_n(x_1,\ldots,x_{n-1})}^{b_n(x_1,\ldots,x_{n-1})} f(x_1, \ldots, x_n)\, dx_n \Big) \ldots dx_2 \Big) dx_1\, .$$

Der rechtsstehende Ausdruck heißt auch ein n-fach iteriertes Integral.

8.5 Die Substitutionsregel

Der folgende Satz soll hier nicht bewiesen werden; der Beweis kann im Buch von *H. Heuser* [7, Abschnitt 205] nachgelesen werden.

Satz 8.5.1 *Sei $O \subset \mathbb{R}^n$ offen, und sei $g : O \longrightarrow \mathbb{R}^n$ stetig partiell differenzierbar. Sei weiter $B \subset O$ kompakt, $N \subset B$ eine Jordan-Nullmenge, und sei g auf $O \setminus N$ injektiv, und $\det g'(q) > 0$ oder $\det g'(q) < 0$ für alle $q \in O \setminus N$. Dann ist $g(B)$ ebenfalls Jordan-messbar, und für jede auf $g(B)$ stetige Funktion f gilt*

$$\int_{g(B)} f(x)\,dx \;=\; \int_B f(g(q))\,\left|\det g'(q)\right|dq.$$

(Ohne Beweis)

Aufgabe 8.5.2 *Zeige: Für $r > 0$ und $x_0 \in \mathbb{R}^n$ gilt $|K(x_0, r)| = r^n\,|K(0, 1)|$.*

Aufgabe 8.5.3 *Finde die Beziehung zwischen den Inhalten der Kugeln vom Radius $r = 1$ von Dimension n und $n + 1$.*

Kapitel 9

Kurvenintegrale und Stammfunktionen

9.1 Kurven, Rektifizierbarkeit, Wege

Definition 9.1.1 *Sei $I = [a, b] \subset \mathbb{R}$ ein abgeschlossenes Intervall mit $a < b$. Eine stetige Abbildung $x : I \longrightarrow \mathbb{R}^n$, $t \mapsto x(t)$, heißt* Parameterdarstellung einer Kurve *in $\mathbb{R}^n$. Das Intervall I heißt auch das* Parameterintervall *der Darstellung. Der Punkt $x(a)$ heißt* Anfangspunkt, *der Punkt $x(b)$ heißt* Endpunkt *der Kurve, und wir sagen manchmal auch, dass die Kurve die Punkte $x(a)$ und $x(b)$* verbindet, *oder sprechen von einer Kurve* von $x(a)$ nach $x(b)$. *Falls $x(a) = x(b)$ ist, sprechen wir von* einer geschlossenen Kurve. *Falls aus $t_1 \leq t_2$ und $x(t_1) = x(t_2)$ folgt, dass $t_1 = t_2$ oder $t_1 = a$ und $t_2 = b$ ist, dann heißt die Kurve* doppelpunktfrei. *Zwei Parameterdarstellungen $x_1 : I_1 \longrightarrow \mathbb{R}^n$ und $x_2 : I_2 \longrightarrow \mathbb{R}^n$ heißen* äquivalent, *wenn es eine stetige, surjektive und streng monoton wachsende Funktion $\phi : I_1 \longrightarrow I_2$ gibt, für welche $x_1 = x_2 \circ \phi$ gilt. Dies ist eine Äquivalenzrelation auf der Menge aller Parameterdarstellungen von Kurven, und jede Äquivalenzklasse heißt* eine Kurve *in $\mathbb{R}^n$. Ist $x : I \longrightarrow \mathbb{R}^n$ Parameterdarstellung einer Kurve γ, so heißt die Menge $\gamma^* = \{x(t) : \ a \leq t \leq b\}$ auch der* Träger *von γ. Beachte, dass äquivalente Parameterdarstellungen den gleichen Träger haben, so dass es berechtigt ist, vom Träger* der Kurve *zu sprechen. Kurven, deren Träger nur aus einem Punkt besteht, heißen auch* Einpunktkurven.

Beispiel 9.1.2 *Alle Strecken und Polygone sind spezielle Kurven. Die Parameterdarstellung $x(t) = (\cos t, \sin t)^T$, $0 \leq t \leq \pi$, beschreibt die obere Hälfte des Einheitskreises, also des Kreises um den Nullpunkt mit Radius $r = 1$, in $\mathbb{R}^2$. Da $\phi(t) = -\cos t$ auf $[0, \pi]$ stetig und streng monoton wachsend ist, ist $\tilde{x}(t) = (-t, \sqrt{1 - t^2}\ )^T$, $t \in [-1, 1]$, eine äquivalente Darstellung.*

Aufgabe 9.1.3 *Zeige: Zu jeder Parameterdarstellung einer Kurve gibt es eine äquivalente Darstellung mit Parameterintervall $[0, 1]$.*

Definition 9.1.4 *Sei $x : [a, b] \longrightarrow \mathbb{R}^n$ Parameterdarstellung einer Kurve γ. Für eine Zerlegung $\mathcal{Z} = \{a = t_0 < t_1 < \ldots < t_N = b\}$ von $[a, b]$ sei*

$$\ell(x, \mathcal{Z}) \ = \ \sum_{k=1}^{N} \|x(t_k) \ - \ x(t_{k-1})\| \,.$$

Falls das Supremum $\ell(\gamma) = \sup_{\mathcal{Z}} \ \ell(x, \mathcal{Z})$, genommen über alle Zerlegungen $\mathcal{Z}$ von $[a, b]$, endlich ist, heißt die Kurve γ rektifizierbar *oder* von endlicher Länge. *Wie die gewählte Bezeichnung schon andeutet, hängt zwar $\ell(x, \mathcal{Z})$ von der gewählten Parameterdarstellung ab, aber $\ell(\gamma)$ ist für äquivalente Parameterdarstellungen immer gleich und heißt* die Länge *der Kurve γ. Die Kurve γ heißt* stetig differenzierbar, *wenn sie* *eine Parameterdarstellung $x : [a,b] \longrightarrow \mathbb{R}^n$ besitzt, welche auf $[a,b]$ stetig differenzierbar ist. Gilt zusätzlich noch $x'(t) \neq 0$ für alle $t \in [a,b]$, so nennen wir γ auch* glatte Kurve *und x glatte Parameterdarstellung von γ.*

Bemerkung 9.1.5 *Sei $x = (x_1, \dots, x_n)^T : [a,b] \longrightarrow \mathbb{R}^n$ Parameterdarstellung einer Kurve γ. Wegen*

$$\begin{aligned} |x_j(t_k) - x_j(t_{k-1})| &\leq \|x(t_k) - x(t_{k-1})\| \\ &\leq \sqrt{n}\ \max\{|x_\nu(t_k) - x_\nu(t_{k-1})| :\ 1 \leq \nu \leq n\} \end{aligned}$$

folgt, dass γ genau dann rektifizierbar ist, wenn alle x_j auf $[a,b]$ von beschränkter Variation sind,[1] *und es gilt die Ungleichung*

$$\max\{V_a^b(x_\nu) :\ 1 \leq \nu \leq n\} \ \leq\ \ell(\gamma)\ \leq \sqrt{n}\ \max\{V_a^b(x_\nu) :\ 1 \leq \nu \leq n\}\,.$$

Beispiel 9.1.6 *Setzt man $t\,\sin(1/t) = 0$ für $t = 0$, so ist*

$$x(t) = \begin{pmatrix} t\,\sin(1/t) \\ t \end{pmatrix}, \qquad \forall\ t \in [0, 2/\pi]$$

Parameterdarstellung einer Kurve in $\mathbb{R}^2$. Die Punkte

$$t_0 \ =\ 0, \qquad t_k \ =\ \frac{1}{\pi\,(N - k + 1/2)}\,, \quad 1 \leq k \leq N,$$

bilden eine Zerlegung $\mathcal{Z}$ von $[0, 2/\pi]$, und wegen $\sin(1/t_k) = (-1)^{n-k}$ folgt

$$\ell(x, \mathcal{Z})\ \geq \sum_{k=1}^{N} |t_k\,\sin(1/t_k) - t_{k-1}\,\sin(1/t_{k-1})|\ \geq \sum_{k=1}^{N}(t_k + t_{k-1}).$$

Die rechte Seite ist größer als $(2/\pi)\sum_{k=2}^{N} 1/k$, und dies geht gegen ∞ für $N \to \infty$. Daher ist diese Kurve nicht rektifizierbar.

Proposition 9.1.7 *Jede stetig differenzierbare Kurve γ ist rektifizierbar. Genauer: Ist $x : [a,b] \longrightarrow \mathbb{R}^n$ eine stetig differenzierbare Parameterdarstellung von γ, so gilt*

$$\ell(\gamma) \ =\ \int_a^b \|x'(t)\|\,dt\,. \tag{9.1.1}$$

Beweis: Sei $\mathcal{Z} = \{a = t_0 < t_1 < \ldots < t_N = b\}$ eine Zerlegung von $[a,b]$ und sei $\tau = (\tau_1, \dots, \tau_n)^T$ ein zugehöriger Zwischenpunktvektor. Dann ist also $S(\mathcal{Z}, \tau) = \sum_{k=1}^{N} \|x'(\tau_k)\|\,(t_k - t_{k-1})$ die entsprechende Riemannsumme für das Integral in (9.1.1). Wenn man auf jede Komponente von $x(t) = (x_1(t), \dots, x_n(t))^T$ den ersten Mittelwertsatz aus Analysis 1 anwendet, findet man

$$\|x(t_k) - x(t_{k-1})\|^2 \ \leq\ (t_k - t_{k-1})^2 \ \sum_{j=1}^{n}[x_j'(\theta_{kj})]^2\,,$$

mit geeigneten Werten $\theta_{kj} \in (t_{k-1}, t_k)$. Sei jetzt ein $\varepsilon > 0$ betrachtet. Da die x_k' auf $[a,b]$ gleichmäßig stetig ist, folgt: Wenn die Zerlegung $\mathcal{Z}$ nur genügend fein ist, dann ist

$$\left|\ \|x'(\tau_k)\| - \sqrt{\sum_{j=1}^{n}[x_j'(\theta_{kj})]^2}\ \right| \ <\ \varepsilon \qquad \forall\ k = 1, \dots, N\,.$$

Also ist $|\ell(x, \mathcal{Z}) - S(\mathcal{Z}, \tau)| \ \leq (b-a)\,\varepsilon$. Hieraus und aus der Definition der Kurvenlänge sowie des Riemann-Integrals folgt die Behauptung. □

[1] Für die Definition dieses Begriffes wird auf die Literatur verwiesen

Definition 9.1.8 *Seien γ_1 und γ_2 Kurven mit Parameterdarstellungen $x_j : [0,1] \longrightarrow \mathbb{R}^n$, und sei der Endpunkt $x_1(1)$ von γ_1 gleich dem Anfangspunkt $x_2(0)$ von γ_2. Dann ist*

$$x(t) \;=\; \begin{cases} x_1(2\,t) & (0 \le t \le 1/2), \\ x_2(2\,t-1) & (1/2 \le t \le 1), \end{cases}$$

Parameterdarstellung einer neuen Kurve, welche wir mit $\gamma_1+\gamma_2$ bezeichnen und die Summe von γ_1 und γ_2 *nennen wollen. Umgekehrt, ist γ eine Kurve mit Parameterdarstellung $x : [a,b] \longrightarrow \mathbb{R}^n$, so sind für irgend ein $c \in (a,b)$ die Restriktionen von x auf die beiden Intervalle $[a,c]$ und $[c,b]$ Parameterdarstellungen von Kurven γ_1 und γ_2, und es gilt $\gamma = \gamma_1 + \gamma_2$. Wir sagen dann auch: γ_1 und γ_2 bilden eine* Zerlegung *von γ. Eine Kurve γ heißt* stückweise stetig differenzierbar *oder kürzer* ein Weg, *wenn γ in endlich viele stetig differenzierbare Teilkurven zerlegt werden kann. Entsprechend definieren wir* stückweise glatte Kurven. *Es ist wichtig zu beachten, dass die Begriffe Kurven und Wege in der Literatur unterschiedlich definiert werden! Ist γ eine Kurve mit Parameterdarstellung $x : [0,1] \longrightarrow \mathbb{R}^n$, so bezeichnen wir die Kurve mit der Parameterdarstellung $x(1-t)$, $0 \le t \le 1$, mit $-\gamma$.*

Bemerkung 9.1.9 *Aus der Definition der Länge von Kurven folgt unmittelbar, dass $\ell(\gamma_1+\gamma_2) = \ell(\gamma_1)+\ell(\gamma_2)$ ist, falls $\gamma_1+\gamma_2$ definiert ist. Daher folgt, dass Wege immer endliche Länge haben, und dass die Gleichung (9.1.1) auch für Wege gilt, wenn man davon absieht, dass der Integrand an endlich vielen Punkten undefiniert sein kann.*

Aufgabe 9.1.10 *Zeige: Äquivalente Parameterdarstellungen haben den gleichen Träger, aber Parameterdarstellungen von Kurven mit gleichem Träger müssen nicht unbedingt äquivalent sein.*

Aufgabe 9.1.11 *Zeige: Einpunktkurven sind stetig differenzierbar, aber nicht glatt.*

Aufgabe 9.1.12 *Zeige: Eine glatte Kurve besitzt immer auch eine Parameterdarstellung, welche nicht an allen Stellen differenzierbar ist.*

Aufgabe 9.1.13 *Berechne die Länge des Kreisbogens γ_s gegeben durch die Parameterdarstellung $x(t) = (r\,\cos t, r\,\sin t)^T$, $0 \le t \le s$.*

9.2 Kurvenintegrale von Vektorfunktionen

In diesem Abschnitt betrachten wir eine Kurve γ mit Träger $\gamma^* \subset \mathbb{R}^n$ und eine Funktion $f : \gamma^* \longrightarrow \mathbb{R}^n$.

Definition 9.2.1 *Sei $x : [a,b] \longrightarrow \mathbb{R}^n$ Parameterdarstellung der Kurve γ. Für eine beliebige Zerlegung $\mathcal{Z} = \{a = t_0 < t_1 < \ldots < t_N = b\}$ von $[a,b]$ und beliebigen Zwischenpunktvektor $\xi = (\xi_1, \ldots, \xi_N)^T$ heißt die Zahl*

$$\sum_{k=1}^{N} f^T(x(\xi_k))\;(x(t_k)\;-\;x(t_{k-1}))$$

die zugehörige Riemannsumme. *Wir sagen dann,*[2]*, dass das* Kurvenintegral

$$I \;=\; \int_\gamma f^T(x)\,dx \;=\; \int_\gamma \Big(\sum_{j=1}^{n} f_j(x)\,dx_j\Big)$$

[2] Vergleiche dies mit der Definition des Riemann-Stieltjes-Integrals – siehe hierzu etwa das Buch von *Walter* [11].

existiert, *wenn es zu jedem $\varepsilon > 0$ eine Zerlegung $\mathcal{Z}_\varepsilon$ gibt, so dass für jede Verfeinerung $\mathcal{Z}$ von $\mathcal{Z}_\varepsilon$ und jeden zu $\mathcal{Z}$ gehörigen Zwischenpunktvektor ξ gilt*

$$\Big| I - \sum_{k=1}^{N} f^T(x(\xi_k)) \; (x(t_k) \; - \; x(t_{k-1}))\Big| < \varepsilon.$$

Wir sehen aus der Definition: Falls die Riemann-Stieltjes-Integrale $\int_a^b f_j(x(t))\, dx_j(t)$ alle existieren, dann existiert auch das Kurvenintegral, und es gilt

$$\int_\gamma f^T(x)\, dx \;=\; \sum_{j=1}^{n} \int_a^b f_j(x(t))\, dx_j(t)\,. \tag{9.2.1}$$

Beachte aber, dass die Existenz des Kurvenintegrals nicht unbedingt die der einzelnen Riemann-Stieltjes-Integrale impliziert. Um das hier eingeführte Kurvenintegral von dem später folgenden Kurvenintegral einer skalarwertigen Funktion zu unterscheiden, sprechen wir auch vom Kurvenintegral einer Vektorfunktion.

Satz 9.2.2 (Existenz und Berechnung von Kurvenintegralen) *Sei γ eine rektifizierbare Kurve, und sei f auf dem Träger γ^* stetig. Dann existiert das Kurvenintegral von f entlang γ, und es gilt* (9.2.1)*. Falls γ sogar eine stetig differenzierbare Parameterdarstellung $x : [a, b] \longrightarrow \mathbb{R}^n$ besitzt, dann gilt*

$$\int_\gamma f^T(x)\, dx \;=\; \int_a^b f^T(x(t))\, x'(t)\, dt\,. \tag{9.2.2}$$

Beweis: Folgt aus dem entsprechenden Satz über Riemann-Stieltjes- Integrale – siehe hierzu etwa das Buch von *Walter* [11]. □

Satz 9.2.3 (Rechenregeln) *Für beliebige Kurven gelten folgende Aussagen, wobei in* (a), (b) *und* (d) *aus der Existenz des Kurvenintegrals/der Kurvenintegrale auf der rechten Seite die Existenz der linken Seite folgt, während in* (c) *die Existenz der Integrale rechts äquivalent zur Existenz des links stehenden Kurvenintegrals ist:*

(a) $\quad \forall\, \alpha \in \mathbb{R} : \quad \int_\gamma \alpha\, f^T(x)\, dx \;=\; \alpha \int_\gamma f^T(x)\, dx.$

(b) $\quad \int_\gamma (f(x) + g(x))^T\, dx \;=\; \int_\gamma f^T(x)\, dx \;+\; \int_\gamma g^T(x)\, dx.$

(c) $\quad \int_{\gamma_1+\gamma_2} f^T(x)\, dx \;=\; \int_{\gamma_1} f^T(x)\, dx \;+\; \int_{\gamma_2} f^T(x)\, dx.$

(d) $\quad \int_{-\gamma} f^T(x)\, dx \;=\; -\int_\gamma f^T(x)\, dx.$

(e) $\quad \Big|\int_\gamma f^T(x)\, dx\Big| \;\le\; \ell(\gamma)\; \sup\{\, \|f(x)\| : x \in \gamma^*\}.$

Beweis: Alle Behauptungen zeigt man unmittelbar mit der Definition; für Einzelheiten siehe z. B. die angegebene Literatur. □

Aufgabe 9.2.4 *Berechne das Kurvenintegral folgender Vektorfunktionen entlang des positiv orientierten Einheitskreises, d. h., für die Kurve mit der Parameterdarstellung wie in* Aufgabe 9.1.13*, mit $s = 2\pi$:*

$$f(x_1, x_2) \;=\; (2\,x_1\,x_2, x_1^2)^T, \quad f(x_1, x_2) \;=\; \Big(-\,\frac{x_2}{x_1^2 + x_2^2}\,, \frac{x_1}{x_1^2 + x_2^2}\Big)^T .$$

Aufgabe 9.2.5 *Sei γ ein doppelpunktfreier Weg. Definiere eine Vektorfunktion $f : \gamma^* \longrightarrow \mathbb{R}^n$ so, dass das Kurvenintegral von f entlang γ gerade die Kurvenlänge $\ell(\gamma)$ ergibt.*

9.3 Wegunabhängigkeit und Stammfunktionen

In diesem Abschnitt sei immer $G \subset \mathbb{R}^n$ ein Gebiet, und $f = (f_1, \ldots, f_n)^T : G \longrightarrow \mathbb{R}^n$ eine Funktion. Wir können uns f als ein *Vektorfeld* veranschaulichen: Jedem $x \in G$ ist ein Vektor $f(x)$ zugeordnet, der z. B. die (Anziehungs-) Kraft angibt, welche auf einen im Punkt x befindlichen Körper der Masse 1 einwirkt, oder die Geschwindigkeit, mit der sich ein im Punkt x befindliches Flüssigkeitsteilchen weiterbewegt. Im ersten Fall spricht man auch von einem Kraftfeld, im zweiten von einem Geschwindigkeitsfeld. Bei der Interpretation als Kraftfeld hat das Kurvenintegral physikalisch die Bedeutung der Arbeit, welche bei einer Verschiebung eines Körpers mit Masse 1 entlang der Kurve γ gewonnen wird.

Definition 9.3.1 *Falls eine partiell differenzierbare Funktion $F : G \longrightarrow \mathbb{R}$ existiert, für welche $f^T(x) =$* grad $F(x)$ *für alle $x \in G$ ist, dann nennen wir F* eine Stammfunktion von f (auf G)*, und in diesem Fall heißt f auch* ein Gradientenfeld *oder* konservatives Feld*, und $-F$ heißt dann auch* potenzielle Energie *des Feldes. Falls f auf G stetig ist, und falls für jeden geschlossenen Weg γ mit Träger in G gilt*

$$\int_\gamma f^T(x)\,dx \;=\; 0\,,$$

dann sagen wir: Das Kurvenintegral über f ist in G wegunabhängig*.*

Bemerkung 9.3.2 *Wenn γ_1, γ_2 zwei Wege in G mit demselben Anfangs- und Endpunkt sind, ist $\gamma = \gamma_1 + (-\gamma_2)$ ein geschlossener Weg, und nach den Rechenregeln für Kurvenintegrale gilt*

$$\int_\gamma f(x)^T\,dx \;=\; \int_{\gamma_1} f(x)^T\,dx \;-\; \int_{\gamma_2} f(x)^T\,dx\,.$$

Daraus liest man ab, dass die Wegunabhängigkeit des Kurvenintegrals äquivalent ist dazu, dass das Kurvenintegral über einen beliebigen Weg in G nicht von dessen Verlauf, sondern höchstens von seinem Anfangs- und Endpunkt abhängt.

Satz 9.3.3 *Das Kurvenintegral über ein stetiges Vektorfeld $f : G \longrightarrow \mathbb{R}^n$ ist genau dann wegunabhängig, wenn es ein Gradientenfeld ist. Ist dies der Fall, und ist F eine Stammfunktion von f, so gilt für jeden in G gelegenen Weg γ mit Anfangs- bzw. Endpunkt x_0 bzw. x_1, dass*

$$\int_\gamma f^T(x)\,dx \;=\; F(x_1) \;-\; F(x_0)\,. \tag{9.3.1}$$

Beweis: Sei das Kurvenintegral wegunabhängig. Für festes $x_0 \in G$ und einen Weg $\gamma(x)$ in G mit Anfangspunkt x_0 und irgend einem Endpunkt $x \in G$ hängt das Kurvenintegral über f nur von x ab, und wir setzen

$$F(x) \;=\; \int_{\gamma(x)} f(y)^T\,dy \qquad \forall\, x \in G\,. \tag{9.3.2}$$

Beachte, dass G offen und zusammenhängend ist, und deshalb existiert nach Satz 2.6.7 für jedes $x \in G$ immer ein Polygon (also ein ganz spezieller Weg), das x_0 und x verbindet. Für genügend kleine $h \in \mathbb{R}\setminus\{0\}$ und ein $k \in \{1, \ldots, n\}$ ist dann

$$\frac{F(x + h\, e_k) \;-\; F(x)}{h} \;-\; f_k(x) \;=\; \frac{1}{h} \int_{\gamma(x,h)} (f(y) - f(x))^T\, dy\,,$$

wobei $\gamma(x, h)$ die Verbindungsstrecke von x nach $x + h\, e_k$ ist. Aus der Abschätzung in Satz 9.2.2 (e) und der Stetigkeit von f folgt, dass die rechte Seite dieser Gleichung für $h \to 0$ gegen 0 geht, und deshalb ist F nach x_k partiell differenzierbar, und $F_{x_k} = f$. Mit anderen Worten: f ist ein Gradientenfeld, und die oben definierte Funktion F ist eine Stammfunktion. Umgekehrt, sei F Stammfunktion zu f. Ein Weg γ in G mit Anfangspunkt x_0 und Endpunkt x_1 ist per Definition die Summe von endlich vielen stetig differenzierbaren Kurven $\gamma_1, \ldots, \gamma_N$, und somit gilt

$$\int_\gamma f^T(x)\, dx \;=\; \sum_{j=1}^N \int_{\gamma_j} f^T(x)\, dx\,.$$

Wenn $x_j : [0, 1] \longrightarrow \mathbb{R}^n$ eine stetig differenzierbare Parameterdarstellung von γ_j ist, dann folgt aus (9.2.2) und dem ersten Hauptsatz, dass

$$\int_{\gamma_j} f^T(x)\, dx \;=\; \int_0^1 f(x_j(t))^T\, x_j'(t)\, dt \;=\; F(x_j(1)) \;-\; F(x_j(0))$$

für alle $j = 1, \ldots, N$. Da $\gamma = \gamma_1 + \ldots + \gamma_N$ ist, muss $x_j(0) = x_{j-1}(1)$ sein, und $x_1(0) = x_0$ sowie $x_N(1) = x_1$. Deshalb folgt

$$\int_\gamma f^T(x)\, dx \;=\; \sum_{j=1}^N \big(F(x_j(1)) \;-\; F(x_j(0))\big) \;=\; F(x_1) \;-\; F(x_0)\,.$$

Das impliziert (9.3.1) und insbesondere auch die Wegunabhängigkeit des Kurvenintegrals über f in G. □

Beispiel 9.3.4 *Die Funktion $k(x) = -c\, \|x\|^{-3}\, x$, für $x \neq 0$ und geeignetes $c > 0$, beschreibt das Schwerkraftfeld der Erde, oder genauer, das eines beliebigen Massepunktes im Nullpunkt. Es ist ein Gradientenfeld, denn die Funktion $U(x) \;=\; c\, \|x\|^{-1}$ ist Stammfunktion.*

Aufgabe 9.3.5 *Zeige dass das Vektorfeld $f(x_1, x_2) = (2\, x_1\, \mathrm{e}^{x_2}, x_1^2\, \mathrm{e}^{x_2})^T$ in ganz $\mathbb{R}^2$ eine Stammfunktion hat.*

Aufgabe 9.3.6 *Zeige, dass $f(x) = (x_1, x_1\, x_2, x_3)^T$, $x = (x_1, x_2, x_3)^T \in \mathbb{R}^3$, kein Gradientenfeld ist.*

9.4 Die Integrabilitätsbedingungen

In einer Dimension hat jede stetige Funktion auch eine Stammfunktion. Dies ist nicht so für $n \geq 2$:

Satz 9.4.1 (Die Integrabilitätsbedingungen) *Sei $f = (f_1, \ldots, f_n)^T$ ein Gradientenfeld in G, und sei $n \geq 2$. Falls dann f in G stetig partiell differenzierbar ist, gilt für alle $x \in G$:*

$$\frac{\partial f_k}{\partial x_j}(x) \;=\; \frac{\partial f_j}{\partial x_k}(x)\,, \qquad 1 \leq j, k \leq n \quad (j \neq k).$$

Beweis: Nach Voraussetzung hat f eine Stammfunktion F, und wegen

$$\frac{\partial^2 F}{\partial x_k \partial x_j}(x) \;=\; \frac{\partial f_j}{\partial x_k}(x)\,, \qquad 1 \le j,k \le n,$$

folgt die Behauptung aus dem Satz von Schwarz über die Vertauschbarkeit der Differenziationsreihenfolge bei zweiten partiellen Ableitungen. □

Beispiel 9.4.2 *Sei $G = \mathbb{R}^2 \setminus \{0\}$, und*

$$f(x) \;=\; f(x_1,x_2) \;=\; \Big(-\,\frac{x_2}{x_1^2+x_2^2}\;,\;\frac{x_1}{x_1^2+x_2^2}\Big)^T.$$

Dann kann man leicht überprüfen, dass die Integrabilitätsbedingungen gelten. Für den Einheitskreis γ mit der Parameterdarstellung $x(t) = (\cos t, \sin t)^T$, $0 \le t \le 2\pi$, ist aber

$$\int_\gamma f(x)^T\,dx \;=\; \int_0^{2\pi} (\sin^2 t + \cos^2 t)\,dt \;=\; 2\,\pi.$$

Also ist das Kurvenintegral über f in G nicht wegunabhängig, und deshalb kann f in G keine Stammfunktion haben. Dies zeigt, dass die Integrabilitätsbedingungen im Allgemeinen nicht hinreichend sind für die Existenz einer Stammfunktion.

Satz 9.4.3 *Sei G sternförmig bezüglich eines Punktes $x_0 \in G$, und erfülle f die Integrabilitätsbedingungen in G. Dann ist f ein Gradientenfeld.*

Beweis: Wir definieren $F(x)$ als das Kurvenintegral über f entlang der Verbindungsstrecke von x_0 nach x, also

$$F(x) \;=\; \int_0^1 f(t\,x + (1-t)\,x_0)^T\,(x - x_0)\,dt\,.$$

Aus Lemma 5.1.8 folgt, dass F nach allen Variablen partiell differenzierbar ist, und dass man die partielle Ableitung nach x_j durch Differenzieren unter dem Integral erhält (vergleiche auch die an das Lemma anschließende Bemerkung). Auf diese Weise folgt

$$F_{x_j}(x) \;=\; \int_0^1 t\,f_{x_j}(t\,x + (1-t)\,x_0)^T\,(x-x_0)\,dt \;+\; \int_0^1 f(t\,x + (1-t)\,x_0)^T\,e_j\,dt\,.$$

Die Integrabilitätsbedingungen bedeuten genau, dass die Ableitung $f'(x)$ eine symmetrische Matrix ist. Wenn man dies ausnutzt, sieht man, dass

$$\begin{aligned} t\,f_{x_j}(t\,x + (1-t)\,x_0)^T\,(x-x_0) \;&=\; t\,\frac{d}{dt}\,f(t\,x + (1-t)\,x_0)^T\,e_j \\ &=\; \frac{d}{dt}\,t\,f(t\,x + (1-t)\,x_0)^T\,e_j \;-\; f(t\,x + (1-t)\,x_0)^T\,e_j\,. \end{aligned}$$

gilt. Daraus folgt aber $F_{x_j}(x) = f_j(x)$, was zu beweisen war. □

Bemerkung 9.4.4 *Der obige Satz gilt allgemeiner auch in sogenannten* einfach zusammenhängenden Gebieten*; darauf wird hier nicht näher eingegangen.*

Aufgabe 9.4.5 *Überprüfe die Integrabilitätsbedingungen für das Vektorfeld in obigem Beispiel.*

9.5 Kurvenintegrale von skalaren Funktionen

Definition 9.5.1 *Sei $x : [a,b] \longrightarrow \mathbb{R}^n$ Parameterdarstellung einer rektifizierbaren Kurve γ der Länge $\ell(\gamma)$. Für $t \in [a,b]$ ist die Restriktion von x auf das Intervall $[a,t]$ wieder Parameterdarstellung einer Kurve γ_t, deren Länge höchstens gleich ℓ ist. Wir nennen die Funktion $s : [a,b] \longrightarrow [0,\ell]$, $t \mapsto s(t) = \ell(\gamma_t)$, die* Bogenlänge *der Kurve γ. Falls x stückweise stetig differenzierbar ist, dann ist*

$$s(t) \;=\; \int_a^t \|x'(u)\|\, du\,, \qquad \forall\; t \in [a,b],$$

und dann folgt $s'(t) = \|x'(t)\|$ außer an endlich vielen Stellen $t \in [a,b]$. Sei $f : \gamma^* \longrightarrow \mathbb{R}$. Wenn das Riemann-Stieltjes-Intgral

$$\int_\gamma f(x)\, ds \;=\; \int_a^b f(x(t))\, ds(t)$$

existiert, heißt es das Kurvenintegral der skalaren Funktion f. *Für stückweise stetig differenzierbares x ist also*

$$\int_\gamma f(x)\, ds \;=\; \int_a^b f(x(t))\, s'(t)\, dt \;=\; \int_a^b f(x(t))\, \|x'(t)\|\, dt\,.$$

Eine Interpretation dieses Kurvenintegrals ist wie folgt: Wenn wir uns die Kurve γ als ein sehr dünnes Drahtstück denken, und wenn $f(x)$ das Gewicht des Drahts pro Längeneinheit an der Stelle x bedeutet, dann ist das Kurvenintegral gerade das Gesamtgewicht des Drahtes.

Aufgabe 9.5.2 *Mit den Bezeichungen wie in der Definition sei $\tilde{x} : [c,d] \longrightarrow \mathbb{R}^n$ eine zu x äquivalente Parameterdarstellung derselben Kurve γ. Nach Definition gibt es also eine streng monoton wachsende und surjektive Funktion $\phi : [c,d] \longrightarrow [a,b]$ mit $\tilde{x}(t) = x(\phi(t))$ für alle $t \in [c,d]$. Ist $\tilde{s}$ die zur neuen Parameterdarstellung $\tilde{x}$ gehörige Bogenlänge, so sieht man aus der Definition, dass $\tilde{s}(t) = s(\phi(t))$ für alle $t \in [c,d]$ gilt. Zeige durch Vergleich der entsprechenden Riemann-Stieltjes-Summen, dass*

$$\int_c^d f(\tilde{x}(t))\, d\tilde{s}(t) \;=\; \int_a^b f(x(t))\, ds(t)$$

gilt. Das bedeutet also, dass das hier definierte Kurvenintegral von der gewählten Parameterdarstellung unabhängig ist.

Aufgabe 9.5.3 *Zeige: Jede glatte Kurve γ mit Länge ℓ besitzt eine Parameterdarstellung $x : [0,\ell] \longrightarrow \mathbb{R}^n$, für die die Ableitung der zugehörigen Bogenlänge konstant gleich 1 ist.*

Kapitel 10

Fourierreihen

10.1 Orthogonalsysteme

Definition 10.1.1 *Ein Vektorraum X über $\mathbb{K}$ heißt* ein Prä-Hilbert-Raum, *wenn eine Abbildung $\langle \cdot , \cdot \rangle$: $X \times X \longrightarrow \mathbb{K}$ gegeben ist, welche die Axiome eines* inneren Produktes *oder* Skalarproduktes *erfüllt; d. h., wenn gilt:*

(S1) $\forall\ x \in X: \quad \langle x, x\rangle \ \geq\ 0\,; \quad \langle x, x\rangle \ =\ 0 \ \Longleftrightarrow\ x = 0.$

(S2) $\forall\ x_1, x_2 \in X: \quad \langle x_1, x_2\rangle \ =\ \overline{\langle x_2, x_1\rangle}.$

(S3) $\forall\ x, x_1, x_2 \in X: \quad \langle x, x_1 + x_2\rangle \ =\ \langle x, x_1\rangle \ +\ \langle x, x_2\rangle.$

(S4) $\forall\ x_1, x_2 \in X,\ \lambda \in \mathbb{K}: \quad \langle x_1, \lambda\, x_2\rangle \ =\ \lambda\, \langle x_1, x_2\rangle.$

Im Falle von $\mathbb{K} = \mathbb{R}$ nennt man einen Prä-Hilbert-Raum auch euklidischen Raum, *im anderen Fall auch einen* unitären Raum. *Auf einem Prä-Hilbert-Raum X definieren wir*

$$\forall\ x \in X: \quad \|x\| \ =\ \sqrt{\langle x, x\rangle}\,.$$

Die Cauchy-Schwarzsche Ungleichung *aus der* Linearen Algebra *besagt, dass dann $|\langle x_1, x_2\rangle| \leq \|x_1\|\ \|x_2\|$ für alle $x_1, x_2 \in X$ gilt. Damit zeigt man leicht, dass die Abbildung $x \mapsto \|x\|$ die Eigenschaften einer Norm auf X besitzt. Ein Prä-Hilbert-Raum, welcher im Sinne von* Abschnitt 3.1 *vollständig ist, heißt* ein Hilbert-Raum.

Aufgabe 10.1.2 *Sei X ein Prä-Hilbertraum. Zeige mit der Cauchy-Schwarzschen Ungleichung, dass folgendes richtig ist:*

$$\forall x, y, x_n, y_n \in X\ :\ \lim_{n\to\infty} \|x_n - x\| \ =\ 0\,, \quad \|y_n - y\| \ =\ 0 \quad \Longrightarrow \quad \lim_{n\to\infty} \langle x_n, y_n\rangle \ =\ \langle x, y\rangle\ .$$

Dies bedeutet, dass die Abbildung $(x, y) \mapsto \langle x, y\rangle$ von $X \times X \to \mathbb{K}$ stetig ist. Wiederhole Aufgabe 1.3.3, *um zu sehen dass $X \times X$ ein normierter Raum ist.*

Beispiel 10.1.3 *Auf dem Raum $C[a, b]$ der auf $[a, b]$ stetigen komplexwertigen Funktionen, mit $a < b$, wird durch die Festsetzung*

$$\forall\ f, g \in C[a, b]: \quad \langle f, g\rangle \ =\ \int_a^b \overline{f(t)}\, g(t)\, dt$$

ein inneres Produkt definiert. Dadurch wird $C[a, b]$ ein Prä-Hilbert-Raum, aber kein Hilbert-Raum.

Definition 10.1.4 *Eine Familie $(x_j)_{j \in J}$ aus einem Prä-Hilbert-Raum X heißt ein* Orthogonalsystem, *falls kein x_j gleich dem Nullvektor ist, und falls gilt*

$$\forall\ j, k \in J\,,\ \ j \neq k: \quad \langle x_j, x_k \rangle \;=\; 0\,.$$

Man zeigt leicht, dass ein Orthogonalsystem immer linear unabhängig ist. Wenn zusätzlich noch alle Vektoren x_j normiert sind, d. h. wenn $\|x_j\| = 1$ ist für alle $j \in J$, dann sprechen wir auch von einem Orthonormalsystem.

Aufgabe 10.1.5 *Zeige mit Hilfe des Schmidtschen Orthogonalisierungsverfahrens, dass es in jedem unendlichdimensionalen Prä-Hilbert-Raum ein abzählbar unendliches Orthonormalsystem gibt.*

Satz 10.1.6 (Beste Approximation, orthogonale Projektion) *Sei $(x_1, \ldots, x_n)$ ein Orthogonalsystem in einem Prä-Hilbert-Raum X, und sei $x \in X$. Dann gilt:*

(a) *Für $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ ist $\|x \;-\; \sum_{k=1}^n \alpha_k\, x_k\|$ genau dann minimal, wenn*

$$\alpha_k \;=\; \frac{\langle x_k, x\rangle}{\langle x_k, x_k\rangle} \;=\; \frac{\langle x_k, x\rangle}{\|x_k\|^2} \qquad \forall\ k \;=\; 1, \ldots, n\,. \tag{10.1.1}$$

(b) *Es gilt*

$$0 \;\leq\; \Big\| x \;-\; \sum_{k=1}^n \frac{\langle x_k, x\rangle}{\|x_k\|^2}\, x_k \Big\|^2 \;=\; \|x\|^2 \;-\; \sum_{k=1}^n \frac{|\langle x_k, x\rangle|^2}{\langle x_k, x_k\rangle}\,. \tag{10.1.2}$$

Beweis: Mit den Axiomen des inneren Produktes folgt

$$\begin{aligned} 0 \;&\leq\; \Big\| x \;-\; \sum_{k=1}^n \alpha_k\, x_k \Big\|^2 \\ &=\; \|x\|^2 \;-\; \sum_{k=1}^n \overline{\alpha_k}\, \langle x_k, x\rangle \;-\; \sum_{k=1}^n \alpha_k\, \langle x, x_k\rangle \;+\; \sum_{j,k=1}^n \overline{\alpha_j}\, \alpha_k\, \langle x_j, x_k\rangle \\ &=\; \|x\|^2 \;-\; \sum_{k=1}^n \frac{|\langle x_k, x\rangle|^2}{\langle x_k, x_k\rangle} \;+\; \sum_{k=1}^n \Big| \alpha_k \;-\; \frac{\langle x_k, x\rangle}{\langle x_k, x_k\rangle} \Big|^2 \; \|x_k\|^2\,. \end{aligned}$$

Daraus folgt (a). Wenn man dann α_k wie in (10.1.1) einsetzt, erhält man auch die Ungleichung (10.1.2).
□

Im Folgenden sei jetzt X ein unendlich-dimensionaler Prä-Hilbert-Raum über $\mathbb{K}$, und $(x_k)_{k=0}^\infty$ sei ein beliebiges, aber fest gewähltes abzählbar unendliches Orthogonalsystem in X.

Definition 10.1.7 *Für beliebige $\alpha_k \in \mathbb{K}$ heißt eine Reihe der Form*

$$\sum_{k=0}^\infty \alpha_k\; x_k$$

eine Orthogonalreihe *mit Koeffizienten α_k. Wir nennen die Reihe* konvergent, *wenn es ein $x \in X$ gibt, für welches die Partialsummen $s_n = \sum_{k=0}^n \alpha_k\; x_k$ gegen x konvergieren, also wenn*

$$\lim_{n \to \infty} \|x \;-\; s_n\| \;=\; 0\,.$$

Die Reihe heißt Cauchy-Reihe, *wenn die Folge der Partialsummen eine Cauchy-Folge ist.*

Für ein $x \in X$ heißen die Zahlen

$$\frac{\langle x_k, x\rangle}{\langle x_k, x_k\rangle} \;=\; \frac{\langle x_k, x\rangle}{\|x_k\|^2} \qquad (k \geq 0) \tag{10.1.3}$$

die allgemeinen Fourier-Koeffizienten *des Vektors x bezüglich des gegebenen Orthogonalsystems $(x_k)_{k=0}^{\infty}$. Die Reihe*

$$\sum_{k=0}^{\infty} \frac{\langle x_k, x\rangle}{\|x_k\|^2} \quad x_k \tag{10.1.4}$$

heißt dann auch die allgemeine Fourier-Reihe *des Vektors x.*

Proposition 10.1.8 *Eine Orthogonalreihe $\sum_{k=0}^{\infty} \alpha_k\, x_k$ ist genau dann eine Cauchy-Reihe, wenn*

$$\sum_{k=0}^{\infty} |\alpha_k|^2 \;\|x_k\|^2 \;<\; \infty\ .$$

Beweis: Mit den Regeln für ein inneres Produkt folgt für beliebiges $m, p \in \mathbb{N}_0$:

$$\Big\| \sum_{k=m}^{m+p} \alpha_k\ x_k \Big\|^2 \;=\; \sum_{k=m}^{m+p} |\alpha_k|^2\ \|x_k\|^2\ ,$$

woraus die Behauptung folgt. □

Satz 10.1.9 (Normkonvergenz von Orthogonalreihen)

(a) *Falls eine Orthogonalreihe $\sum_{k=0}^{\infty} \alpha_k\, x_k$ gegen einen Vektor $x \in X$ konvergiert, dann sind die Koeffizienten α_k genau die allgemeinen Fourier-Koeffizienten von x; d. h., es gilt*

$$\alpha_k \;=\; \frac{\langle x_k, x\rangle}{\|x_k\|^2} \qquad (k \geq 0)\,.$$

(b) *Falls die allgemeine Fourier-Reihe* (10.1.4) *eines Vektors $x \in X$ gegen einen Vektor $\tilde{x} \in X$ konvergiert, dann folgt*

$$\langle x\,,\, x_k\rangle - \langle \tilde{x}\,,\, x_k\rangle \;=\; \langle x - \tilde{x}\,,\, x_k\rangle \;=\; 0 \qquad \forall\ k \in \mathbb{N}_0$$

d. h., die Vektoren x und $\tilde{x}$ haben die gleichen allgemeinen Fourier-Koeffizienten, oder anders ausgedrückt: die Differenz $x - \tilde{x}$ ist zu allen x_k orthogonal.

Beweis: Zu (a): Aus $x = \lim_{n\to\infty} \sum_{k=0}^{n} \alpha_k\, x_k$ folgt mit Hilfe der Stetigkeit des Skalarproduktes (vergl. Aufgabe 10.1.2) und der Orthogonalität der x_k:

$$\langle x_j\ ,\ x\rangle \;=\; \lim_{n\to\infty}\ \Big\langle x_j\,,\, \sum_{k=0}^{n} \alpha_k\, x_k\ \Big\rangle \;=\; \alpha_j\ \|x_j\|^2\ .$$

Zu (b): Wenn man in der obigen Gleichung für α_k die Fourier-Koeffizienten von x einsetzt und dann x durch $\tilde{x}$ ersetzt, so folgt die Behauptung. □

Definition 10.1.10 *Wir nennen das Orthogonalsystem* (x_k) maximal*, wenn die lineare Hülle* $LH(x_k)$ *in* X *dicht liegt; d. h., wenn* $\overline{LH(x_k)} = X$ *ist. Wir sagen weiter, dass* (x_k) vollständig *ist, wenn aus* $\langle x\,,\, x_k\rangle = 0$ *für alle* $k \in \mathbb{N}_0$ *folgt, dass* $x = 0$ *ist. Das bedeutet also, dass die Fourierreihe eines Vektors* x *bei einem maximalen Orthogonalsystem, wenn überhaupt, nur gegen* x *konvergieren kann.*

Satz 10.1.11 (Normkonvergenz der Fourierreihe)

(a) *Für jedes* $x \in X$ *gilt die* Besselsche Ungleichung

$$\sum_{k=0}^{\infty} \frac{|\langle x_k, x\rangle|^2}{\|x_k\|^2} \;\leq\; \|x\|^2\,. \tag{10.1.5}$$

(b) *Genau dann konvergiert die Fourierreihe eines* $x \in X$ *gegen* x*, wenn die* Parsevalsche Gleichung

$$\sum_{k=0}^{\infty} \frac{|\langle x_k, x\rangle|^2}{\|x_k\|^2} \;=\; \|x\|^2 \tag{10.1.6}$$

gilt, d. h. also, wenn in der Besselschen Ungleichung das Gleichheitszeichen eintritt.

(c) *Wenn* (x_k) *maximal ist, gilt* (10.1.6) *für alle* $x \in X$*, und* (x_k) *ist auch vollständig.*

(d) *Wenn* X *ein Hilbert-Raum und* (x_k) *ein vollständiges Orthogonalsystem ist, dann ist* (x_k) *auch maximal.*

Beweis: Aus (10.1.2) folgen für $n \to \infty$ sofort (a) und (b). Falls (x_k) maximal ist, gibt es per Definition zu jedem $\varepsilon > 0$ eine Linearkombination $\tilde{x} = \sum_{k=0}^{n} \alpha_k\, x_k$, mit $n = n(\varepsilon)$, für welche $\|x - \tilde{x}\| < \varepsilon$ ist. Nach Satz 10.1.6 gilt dies erst recht, wenn wir α_k als die allgemeinen Fourier-Koeffizienten von x wählen. Daher folgt (c). Die allgemeine Fourierreihe von x ist wegen der Besselschen Ungleichung und Proposition 10.1.8 ein Cauchy-Reihe, konvergiert also gegen ein $\tilde{x} \in X$, falls X ein Hilbert-Raum ist. Nach Satz 10.1.9 ist $x - \tilde{x}$ zu allen x_k orthogonal, und daraus folgt $\tilde{x} = x$ wegen der Vollständigkeit. Also gilt (d). □

Beispiel 10.1.12 *Wir setzen* X *gleich der Menge aller Linearkombinationen über* $\mathbb{C}$ *der Folgen* $x^{(k)} = (x_n^{(k)})_{n=0}^{\infty}$*, mit*

$$x_n^{(0)} \;=\; \frac{1}{n+1}\,, \quad x_n^{(k)} \;=\; \delta_{nk} \;=\; \begin{cases} 1 & (n = k) \\ 0 & (n \neq k) \end{cases} \qquad \forall\; k \geq 1,\; n \geq 0\,.$$

Es folgt, dass für jedes $x = (x_n) \in X$ *entweder nur endlich viele Glieder* $x_n \neq 0$ *sind, oder dass gilt* $x_n = 1/n$ *für alle* $n \geq n_0$*. Als inneres Produkt zweier Folgen* $x = (x_n)$*,* $y = (y_n)$ *setzen wir*

$$\langle x, y\rangle \;=\; \sum_{n=0}^{\infty} x_n\, y_n\,.$$

Beachte, dass die Reihe für $x, y \in X$ *immer absolut konvergent ist, denn entweder sind nur endlich viele ihrer Glieder von 0 verschieden, oder es gilt* $x_n\, y_n = 1/n^2$ *für alle* $n \geq n_0$*. Die* $(x^{(k)})_{k=1}^{\infty}$ *sind ein Orthogonalsystem in* X*, und aus* $\langle x, x^{(k)}\rangle = 0$ *für* $k \geq 1$ *folgt, dass alle Glieder* x_n *der Folge* x *verschwinden müssen, ausser evtl. dem Anfangsglied* x_0*. Nach Definition von* X *kann dies aber nur gelten, wenn* x *die Nullfolge ist. Das bedeutet, dass das Orthogonalsystem vollständig ist. Es ist aber nicht maximal, denn ist* $\tilde{x}$ *eine beliebige Linearkombination der* $x^{(k)}$*,* $k \geq 1$*, so beginnt die Folge* $\tilde{x}$ *mit* 0*, und deshalb ist* $\|x^{(0)} - \tilde{x}\| \geq 1$*.*

Beispiel 10.1.13 *In $C[a,b]$, $a < b$, mit dem üblichen inneren Produkt kann man mit Hilfe des Schmidtschen Orthogonalisierungsverfahrens zu der Familie aller Monome $(t^k)_{k=0}^{\infty}$ ein äquivalentes Orthogonalsystem $(p_k)_{k=0}^{\infty}$ konstruieren. Dabei ist jedes p_k ein Polynom vom Grade k. Die Menge aller Polynome ist* dicht *in $C[a,b]$, d. h., in jeder Kugel um jeden beliebigen Vektor liegt immer ein Polynom. Dies folgt aus dem sogenannten* Weierstraßschen Approximationssatzes, *der in dieser Vorlesung nicht behandelt wurde. Daraus wiederum ergibt sich, dass (p_k) ein maximales Orthogonalsystem ist. Orthogonalsysteme von Polynomen sind in verschiedenen Anwendungen wichtig. Wenn z. B. $[a,b] = [-1,1]$ ist, erhält man das Orthogonalsystem der* Legendre-Polynome

$$p_k(t) \; = \; \frac{1}{2^k \, k!} \, \frac{d^k}{dt^k} \, (t^2 \; - \; 1)^k \; , \qquad k \in \mathbb{N}_0 \, .$$

Für den Rest dieses Kapitels ist das Orthogonalsystem der trigonometrischen Funktionen *wichtig: Wir setzen*

$$f_0(t) \; \equiv \; 1 \, , \quad f_{2k-1}(t) \; = \; \sin(k\,t) \, , \quad f_{2k}(t) \; = \; \cos(k\,t) \, , \qquad \forall \; k \in \mathbb{N} \, .$$

Diese Funktionen bilden auf jedem Intervall der Länge 2π ein Orthogonalsystem, welches wir das trigonometrische System *nennen wollen. Wir setzen weiter für eine auf dem Intervall $[-\pi,\pi]$ integrierbare Funktion f*

$$a_k \; = \; \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \, \cos(k\,t) \; dt \qquad \forall k \in \mathbb{N}_0 \, , \tag{10.1.7}$$

$$b_k \; = \; \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \, \sin(k\,t) \; dt \qquad \forall k \in \mathbb{N} \, , \tag{10.1.8}$$

und nennen diese Zahlen auch die Fourierkoeffizienten *der Funktion f. Die Reihe*

$$\frac{a_0}{2} \; + \; \sum_{k=1}^{\infty} (a_k \, \cos(k\,t) \; + \; b_k \, \sin(k\,t)) \tag{10.1.9}$$

heißt die Fourierreihe der Funktion f. *Beachte, dass diese Reihe nichts anderes ist als die Orthogonalreihe von f bezüglich des trigonometrischen Systems. Siehe dazu auch untenstehende Aufgaben.*

Aufgabe 10.1.14 *Zeige die Orthogonalität des oben eingeführten trigonometrischen Systems $(f_k)_{k=0}^{\infty}$ auf dem Intervall $[-\pi,\pi]$. Berechne die Normen der Funktionen f_k und vergleiche die Formeln* (10.1.7) *und* (10.1.8) *mit* (10.1.3).

Aufgabe 10.1.15 *Zeige, dass die folgenden Funktionen f die angegebenen Fourierkoeffizienten haben:*

1. $f(t) \equiv 1 \implies a_0 = 2$, $a_k = b_k = 0$ für alle $k \in \mathbb{N}$.
2. $f(t) = t \implies a_k = 0$ für alle $k \in \mathbb{N}_0$, $b_k = 1/(2k)$ für alle $k \in \mathbb{N}$.
3. $f(t) = |t| \implies a_0 = \pi$, $a_k = \frac{2}{k^2\pi} \, [(-1)^k - 1]$, $b_k = 0$ für alle $k \in \mathbb{N}$.

Was kann man zur Konvergenz der jeweiligen Fourierreihe sagen?

Aufgabe 10.1.16 *Zeige: Wenn f eine gerade bzw. ungerade Funktion ist, dann verschwinden alle b_k bzw. alle a_k. Man sagt dann: Die Fourierreihe ist eine Cosinus- bzw. eine Sinusreihe.*

10.2 Die komplexe Form der Fourierreihe

Definition 10.2.1 *Für $k \in \mathbb{Z}$ setzen wir*

$$f_k(x) \;=\; \mathrm{e}^{i\,k\,x} \;=\; (\mathrm{e}^{i\,x})^k \;=\; \cos(k\,x) \;+\; i\,\sin(k\,x)\,.$$

Dann folgt für $j \neq k$:

$$\int_{-\pi}^{\pi} \overline{f_j(x)}\, f_k(x)\,dx \;=\; \int_{-\pi}^{\pi} \mathrm{e}^{i\,(k-j)\,x}\,dx \;=\; \frac{\mathrm{e}^{i\,(k-j)}}{i\,(k-j)}\Bigg|_{-\pi}^{\pi} \;=\; 0\,.$$

Also sind die Funktionen f_k ein Orthogonalsystem, und man rechnet aus, dass $\|f_k\|^2 = 2\,\pi$, für alle $k \in \mathbb{Z}$. Für eine über $[-\pi,\pi]$ integrierbare Funktion f setzen wir

$$c_k \;=\; \frac{\langle f_k, f\rangle}{2\,\pi} \;=\; \frac{1}{2\,\pi}\int_{-\pi}^{\pi} \mathrm{e}^{-i\,k\,x}\; f(x)\,dx \qquad \forall\; k \in \mathbb{Z}$$

und nennen die Reihe

$$\sum_{k=-\infty}^{\infty} c_k\;f_k(x) \;=\; \sum_{k=-\infty}^{\infty} c_k\;\mathrm{e}^{i\,k\,x}$$

die komplexe Form *der Fourierreihe von f.*

Bemerkung 10.2.2 *Wegen $\mathrm{e}^{-i\,k\,x} = \cos(k\,x) - i\,\sin(k\,x)$ folgt, dass $c_0 = a_0/2$ und*

$$c_k \;=\; \frac{a_k - i\,b_k}{2}\,, \quad c_{-k} \;=\; \frac{a_k + i\,b_k}{2} \qquad \forall\; k \in \mathbb{N}\,.$$

Daraus folgt sofort, dass

$$a_0 \;=\; 2\,c_0\,, \quad a_k \;=\; c_k + c_{-k}\,, \quad b_k \;=\; i\,(c_k - c_{-k}) \qquad \forall\; k \in \mathbb{N}\,.$$

Falls f eine Funktion mit reellen Werten ist, was normalerweise der Fall ist, dann sind a_k und b_k reelle Zahlen, und

$$a_k \;=\; 2\;\mathrm{Re}\;c_k \;=\; 2\;\mathrm{Re}\;c_{-k}\,, \quad b_k \;=\; -2\;\mathrm{Im}\;c_k \;=\; 2\;\mathrm{Im}\;c_{-k} \qquad \forall\; k \in \mathbb{N}\,.$$

10.3 Punktweise Konvergenz von Fourierreihen

Im Folgenden betrachten wir immer Funktionen $f : \mathbb{R} \longrightarrow \mathbb{R}$, welche $2\,\pi$-periodisch sind, d. h., für welche $f(x+2\,\pi) = f(x)$ gilt für alle $x \in \mathbb{R}$. Weiter sei f über das Intervall $[-\pi,\pi]$ Riemann-integrierbar. Beachte, dass aus diesen Voraussetzungen für jedes $k \in \mathbb{N}$ die Integrierbarkeit der Funktionen $f(t)\,\cos(k\,t)$ und $f(t)\,\sin(k\,t)$ über jedes abgeschlossene Intervall $[a,b]$ folgt. Insbesondere existieren immer die Fourierkoeffizienten a_k, b_k der Funktion f. Ob die Fourierreihe von f punktweise konvergiert, ist dabei im allgemeinen nicht gesichert und soll im Folgenden genauer untersucht werden. Wir zeigen jetzt, dass die Glieder der Fourierreihe jedenfalls eine Nullfolge bilden. Es gilt sogar allgemeiner:

Proposition 10.3.1 *Für beliebiges $a < b$ und f wie oben gilt immer*

$$\lim_{k\to\infty}\int_a^b f(t)\,\cos(k\,t)\,dt \;=\; 0\,, \qquad \lim_{k\to\infty}\int_a^b f(t)\,\sin(k\,t)\,dt \;=\; 0\,.$$

Beweis: Für ein Intervall der Länge 2π folgt die Behauptung aus der Besselschen Ungleichung; beachte allerdings folgende Feinheit: Die Menge aller Funktionen f mit den oben angegebenen Eigenschaften ist zwar ein Vektorraum, aber da die Funktionen nicht stetig sein müssen, ist das innere Produkt auf diesem Raum nicht positiv definit, d. h., das Axiom (S1) ist verletzt. Setzt man aber $f \sim g$ wenn $\|f - g\| = 0$ ist, so wird die Menge der Äquivalenzklassen in natürlicher Weise zu einem Prä-Hilbert-Raum. Siehe dazu auch die Aufgaben 1.2.4 und 1.2.5 sowie Aufgabe 2; dies und weitere Einzelheiten werden in *Maßtheorie* und/oder der Vorlesung *Funktionalanalysis* genauer besprochen. Da ein allgemeines Intervall in endlich viele Intervalle einer Länge $\le 2\pi$ zerlegt werden kann, ist jetzt noch der Fall $b - a < 2\pi$ zu behandeln. In dieser Situation setzen wir

$$g(t) \;=\; \begin{cases} f(t) & (a \le t \le b) \\ 0 & (b < t \le a + 2\pi) \end{cases}$$

Diese Funktion kann auf eindeutige Weise zu einer 2π-periodischen Funktion auf $\mathbb{R}$ fortgesetzt werden, und die Fourierkoeffizienten von g stimmen mit den Integralen in der Behauptung überein. Aus der Besselschen Ungleichung, angewandt auf g, folgt deshalb die Behauptung. □

Definition 10.3.2 *Wir setzen*

$$s_n(t) \;=\; \frac{a_0}{2} \;+\; \sum_{k=1}^{n} (a_k \cos(k\,t) \;+\; b_k \sin(k\,t))\;,$$

$$D_n(t) \;=\; \frac{1}{2} \;+\; \sum_{k=1}^{n} \cos(k\,t) \;=\; \frac{\sin(2\,n+1)\frac{t}{2}}{2\,\sin\frac{t}{2}}\;,$$

wobei der rechts stehende Quotient für $t = 0$ als der entprechende Grenzwert zu interpretieren und die Gleichheit der beiden Darstellungen durch Induktion zu zeigen ist. Die Funktion $D_n(t)$ heißt der Dirichlet-Kern*; seine Bedeutung ergibt sich aus der folgenden Behauptung.*

Aufgabe 10.3.3 *Berechne $\int_0^\pi D_n(t)\,dt$.*

Behauptung 10.3.4 *Es gilt die folgende Integraldarstellung für die Partialsummen der Fourierreihe von f:*

$$s_n(t) \;=\; \frac{1}{\pi} \int_0^\pi [f(t+\tau) \;+\; f(t-\tau)]\; D_n(\tau)d\tau \qquad \forall\; t \in \mathbb{R}\,.$$

Beweis: Unter Benutzung von (10.1.7) und (10.1.8) sowie der Additionstheoreme für die trigonometrischen Funktionen folgt

$$\begin{aligned} s_n(t) \;&=\; \frac{1}{\pi} \int_{-\pi}^{\pi} f(\tau) \left[\frac{1}{2} \;+\; \sum_{k=1}^{n} \cos(k\,\tau)\,\cos(k\,t) \;+\; \sin(k\,\tau)\,\sin(k\,t)\right] d\tau \\ &=\; \frac{1}{\pi} \int_{-\pi}^{\pi} f(\tau) \left[\frac{1}{2} \;+\; \sum_{k=1}^{n} \cos(k\,(\tau - t)\right] d\tau \\ &=\; \frac{1}{\pi} \int_{-\pi}^{\pi} f(\tau + t)\, D_n(\tau)\, d\tau\,, \end{aligned}$$

wobei für die letzte Gleichung eine Substitution $\tau \to \tau + t$ sowie die Periodizität von f benutzt wurde. Wenn man das Integrationsintervall noch in zwei Teilintervalle zerlegt und eine weitere Substitution für das eine Teilintegral durchführt, so ergibt sich die Behauptung. □

Proposition 10.3.5 *Genau dann existiert $s(t) = \lim_{n\to\infty} s_n(t)$ für ein $t \in \mathbb{R}$, wenn*

$$\int_0^\pi [f(t+\tau) \; + \; f(t-\tau) \; - \; 2\,s(t)]\; D_n(\tau)d\tau \;\longrightarrow\; 0 \qquad (n\to\infty)\,.$$

Beweis: Folgt aus der obenstehenden Behauptung und der Tatsache, dass $\int_0^\pi D_n(t)\,dt = \pi/2$ ist; vergleiche hierzu auch Aufgabe 10.3.3. □

Satz 10.3.6 (Riemannscher Lokalisationssatz)
Genau dann existiert $s(t) = \lim_{n\to\infty} s_n(t)$ für ein $t \in \mathbb{R}$, wenn es ein $\delta > 0$ gibt, für welches

$$\int_0^\delta [f(t+\tau) \; + \; f(t-\tau) \; - \; 2\,s(t)]\; D_n(\tau)d\tau \;\longrightarrow\; 0 \qquad (n\to\infty)\,.$$

Dies bedeutet, dass das Konvergenzverhalten der Fourierreihe an der Stelle t nur von den Werten von f in der Nähe von t abhängt, obwohl bei der Berechnung der Koeffizienten alle Werte von f auf dem Intervall $[-\pi,\pi]$ eingehen.

Beweis: Sei t festgehalten, und sei $g(\tau) = f(t+\tau) \; + \; f(t-\tau) \; - \; 2\,s(t)$ gesetzt. Dann ist

$$\int_\delta^\pi g(\tau)\,D_n(\tau)d\tau \;=\; \int_\delta^\pi \frac{g(\tau)}{2\,\sin(\tau/2)}\;[\cos(\tau/2)\,\sin(n\,\tau) \;+\; \sin(\tau/2)\,\cos(n\tau)]\;d\tau\,.$$

Aus Proposition 10.3.1, angewandt auf die Funktion $f(\tau) = g(\tau)/\sin(\tau/2)$ für $\delta \le \tau \le \pi$, bzw. $f(\tau) = 0$ für $0 \le \tau < \delta$, folgt $\int_\delta^\pi g(\tau)\,D_n(\tau)d\tau \to 0$ für $n\to\infty$. Daraus folgt mit Hilfe von Proposition 10.3.5 die Behauptung. □

Satz 10.3.7 (Dirichletsche Regel) *Falls f auf einem abgeschlossenen Intervall $[a,b]$ von beschränkter Variation ist, dann gilt*

$$\lim_{n\to\infty} s_n(t) \;=\; \frac{f(t+)+f(t-)}{2} \qquad \forall\; t\in[a,b]\,.$$

Insbesondere konvergiert die Fourierreihe gegen $f(t)$ an jeder in $[a,b]$ gelegenen Stetigkeitsstelle von f.

Satz 10.3.8 *Ist f auf $[-\pi,\pi]$ stückweise stetig differenzierbar, so konvergiert die Fourierreihe gleichmäßig gegen f.*

Wir geben noch eine Reihe von Beispielen:

Beispiel 10.3.9 *Für $f(t) = t$, $-\pi < t < \pi$, $f(\pm\pi) = 0$ folgt mit der Dirichletschen Regel, dass*

$$f(t) \;=\; 2\,\sum_{k=1}^\infty (-1)^{k+1}\,\frac{\sin(k\,t)}{k} \qquad \forall\; t\in[-\pi,\pi]\,.$$

Für $t = \pi/2$ folgt (wegen $\sin(k\,\pi/2) = (-1)^j$ für $k = 2\,j+1$ bzw. $= 0$ für $k = 2\,j$)

$$\frac{\pi}{4} \;=\; \sum_{j=0}^\infty \frac{(-1)^j}{2\,j+1}\,.$$

Beispiel 10.3.10 *Für $f(t) = |t|$, $-\pi < t < \pi$, folgt*

$$f(t) \;=\; \frac{\pi}{2} \;-\; \frac{4}{\pi} \sum_{k=1}^{\infty} \frac{\cos(2\,k-1)t}{(2\,k-1)^2} \qquad \forall\; t \in (-\pi, \pi]\,.$$

Für $t = 0$ folgt hieraus

$$\frac{\pi^2}{8} \;=\; \sum_{k=1}^{\infty} \frac{1}{(2\,k-1)^2}\;.$$

Beispiel 10.3.11 *Es gilt*

$$|\sin t| \;=\; \frac{4}{\pi} \left(\frac{1}{2} \;-\; \sum_{k=1}^{\infty} \frac{\cos(2\,k\,t)}{(2\,k-1)(2\,k+1)} \right) \qquad \forall\; t \in [-\pi, \pi]\,.$$

Für $t = \pi/2$ folgt hieraus

$$\frac{\pi}{4} \;=\; \frac{1}{2} \;+\; \sum_{k=1}^{\infty} \frac{(-1)^k}{(2\,k-1)(2\,k+1)}\;.$$

Beispiel 10.3.12 *Es gilt*

$$\frac{\pi^2}{12} \;-\; \frac{t^2}{4} \;=\; \sum_{k=1}^{\infty} (-1)^k\; \frac{\cos(k\,t)}{k^2} \qquad \forall\; t \in [-\pi, \pi]\,.$$

Für $t = 0$ folgt hieraus

$$\frac{\pi^2}{12} \;=\; \sum_{k=1}^{\infty} \frac{(-1)^k}{k^2}\;.$$

Beispiel 10.3.13 *Es gilt für alle $\alpha \neq 0$ (reell)*

$$\cosh(\alpha\,t) \;=\; \frac{\sinh(\alpha\,\pi)}{\pi} \left(\frac{1}{\alpha} \;+\; \sum_{k=1}^{\infty} (-1)^k\; \frac{2\,\alpha\,\cos(k\,t)}{\alpha^2 + k^2} \right) \qquad \forall\; t \in [-\pi, \pi]\,.$$

Für $t = \pi$ folgt hieraus die sogenannte Partialbruchzerlegung des hyperbolischen Cotangens

$$\pi\,\coth(\alpha\,\pi) \;=\; \frac{1}{\alpha} \;+\; \sum_{k=1}^{\infty} \frac{2\,\alpha}{\alpha^2 + k^2}\;.$$

Aufgabe 10.3.14 *Bestätige die in den obigen Beispielen angegebenen Fourierreihen.*

10.4 Die Größenordnung der Fourierkoeffizienten

Bisher wissen wir nur, dass die Fourierkoeffizienten immer Nullfolgen sind, aber nicht, mit welcher Geschwindigkeit sie gegen 0 konvergieren. Jetzt wollen wir einige Resultate in dieser Richtung kennen lernen.

Satz 10.4.1 *Ist f auf $[-\pi, \pi]$ von beschränkter Variation, so sind die Folgen $(n\,a_n)$ und $(n\,b_n)$ beide beschränkt.*

Beweis: Aus *Analysis 1* wissen wir, dass eine Funktion von beschränkter Variation als Differenz von monoton wachsenden Funktionen geschrieben werden kann. Wir können deshalb o. B. d. A. annehmen, dass f monoton ist. Nach dem zweiten Mittelwertsatz der integralrechnung gibt es dann zu jedem $n \in \mathbb{N}$ ein $\xi = \xi_n \in (-\pi, \pi)$ mit

$$\pi\, a_n \;=\; f(-\pi) \int_{-\pi}^{\xi} \cos(n\,x)\,dx \;+\; f(\pi) \int_{\xi}^{\pi} \cos(n\,x)\,dx\,.$$

Wenn man diese beiden Integrale ausrechnet und berücksichtigt, dass die trigonometrischen Funktionen beschränkt sind, folgt die Behauptung. □

Allgemeiner gilt folgender Satz:

Satz 10.4.2 *Für ein $m \in \mathbb{N}$ sei f m-mal differenzierbar auf $[-\pi, \pi]$, und $f^{(m)}$ sei dort von beschränkter Variation. Weiter gelte*

$$f^{(k)}(-\pi) \;=\; f^{(k)}(\pi) \qquad \forall\; k \;=\; 0, \ldots, m-1\,.$$

Dann sind die Folgen $(n^{m+1}\, a_n)$ und $(n^{m+1}\, b_n)$ beide beschränkt.

Beweis: Folgt durch partielle Integration aus dem vorherigen Satz. □

Aufgabe 10.4.3 *Finde die Größenordnung der Fourierkoeffizienten der Funktion $f(x) = x^2$; d. h., finde heraus, für welches m die Folgen $(n^{m+1}\, a_n)$ und $(n^{m+1}\, b_n)$ beide beschränkt sind.*

Kapitel 11

Vektoranalysis und Integralsätze

11.1 Krummlinige Koordinaten

Im Folgenden seien O_1, O_2 offene Mengen in $\mathbb{R}^3$, und

$$q(x) = \begin{pmatrix} q_1(x_1, x_2, x_3) \\ q_2(x_1, x_2, x_3) \\ q_3(x_1, x_2, x_3) \end{pmatrix}$$

sei eine bijektive Abbildung von O_1 auf O_2. Für die Umkehrabbildung schreiben wir dann $x(q) = (x_1(q_1, q_2, q_3), x_2(q_1, q_2, q_3), x_3(q_1, q_2, q_3))^T$. Wir setzen weiter voraus, dass alle später vorkommenden partiellen Ableitungen dieser Funktionen existieren und stetig sein sollen. Wir interpretieren die Werte von $q_1(x), q_2(x), q_3(x)$ als *krummlinige Koordinaten* des Punktes $x \in \mathbb{R}^3$. Manchmal betrachten wir in Beispielen auch offene Teilmengen O_1, $O_2 \in \mathbb{R}^2$ und eine entsprechende bijektive Abbildung $q(x) = (q_1(x_1, x_2), q_2(x_1, x_2))^T$. Alle nachfolgenden Begriffe und Gleichungen sind dann entsprechend zu modifizieren.

Definition 11.1.1 *Wir setzen für $j, k, \ell = 1, 2, 3$*

$$\varepsilon_{q_k} = \frac{\partial}{\partial q_k} x \,, \quad \varepsilon_{q_j q_k} = \frac{\partial^2}{\partial q_j \partial q_k} x = \varepsilon_{q_k q_j} \,, \quad g_k = \|\varepsilon_{q_k}\| \,,$$

und, unter der zusätzlichen Annahme von $g_k \neq 0$,

$$e_{q_k} = \frac{1}{g_k} \varepsilon_{q_k} \,, \quad \Gamma^{\ell}_{jk} = \frac{1}{g_\ell^2} \langle \varepsilon_{q_j q_k}, \varepsilon_{q_\ell} \rangle \,.$$

Die Zahlen Γ^{ℓ}_{jk} heißen auch Christoffel-Symbole.

Hält man z. B. die Werte von q_2 und q_3 fest, so beschreibt x als Funktion von q_1 eine Kurve in $\mathbb{R}^3$, und ε_{q_1} ist ein Tangentenvektor an die Kurve. Der Vektor e_{q_1} ist dann ein entsprechender Einheitsvektor. Entsprechendes gilt für alle drei Vektoren ε_{q_j} bzw. e_{q_j}.

Definition 11.1.2 *Ein System von krummlinigen Koordinaten heißt* orthogonal*, wenn die drei Vektoren e_1, e_2, e_3 immer ein Orthonormalsystem bilden.*

Aufgabe 11.1.3 *Zeige für orthogonale krummlinige Koordinaten die Gleichung*

$$\varepsilon_{q_j q_k} \;=\; \sum_{\ell=1}^{3} \Gamma_{jk}^{\ell}\, \varepsilon_{q_\ell}\,, \qquad 1 \;\le\; j,k \;\le\; 3\,.$$

Proposition 11.1.4 *Für orthogonale krummlinige Koordinaten gilt immer*

$$g_\ell^2\; \Gamma_{jk}^{\ell} \;=\; -\, g_k^2\; \Gamma_{j\ell}^{k}\,, \qquad 1 \;\le\; j,k,\ell \;\le\; 3\,.$$

Beweis: Bei orthogonalen Koordinaten gilt immer $0 = \langle \varepsilon_{q_k}, \varepsilon_{q_\ell} \rangle$. Durch Ableiten nach q_j folgt hieraus

$$0 \;=\; \langle \varepsilon_{q_k q_j}, \varepsilon_{q_\ell} \rangle \;+\; \langle \varepsilon_{q_k}, \varepsilon_{q_\ell q_j} \rangle\,.$$

Daraus folgt die Behauptung. $\square$

Beispiel 11.1.5 (Polarkoordinaten in $\mathbb{R}^2$) *Wir schreiben $q_1 = r$, $q_2 = \phi$ und definieren $x(r,\phi) = (r\,\cos\phi, r\,\sin\phi)^T$. Diese Abbildung ist bijektiv von $O_2 = (0,\infty) \times (-\pi,\pi)$ nach $O_1 = \mathbb{R}^2 \setminus \{x_1 \le 0\}$. Es folgt*

$$\varepsilon_r \;=\; \begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix}, \quad g_1 \;=\; 1\,, \quad \varepsilon_\phi \;=\; \begin{pmatrix} -r\,\sin\phi \\ r\,\cos\phi \end{pmatrix}, \quad g_2 \;=\; r\,,$$

$$e_r \;=\; \begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix}, \quad e_\phi \;=\; \begin{pmatrix} -\sin\phi \\ \cos\phi \end{pmatrix},$$

$$\varepsilon_{rr} \;=\; 0\,, \quad \varepsilon_{r\phi} \;=\; \varepsilon_{\phi r} \;=\; \begin{pmatrix} -\sin\phi \\ \cos\phi \end{pmatrix}, \quad \varepsilon_{\phi\phi} \;=\; \begin{pmatrix} -r\,\cos\phi \\ -r\,\sin\phi \end{pmatrix}.$$

Beispiel 11.1.6 (Kugelkoordinaten in $\mathbb{R}^2$) *Wir schreiben jetzt $q_1 = r$, $q_2 = \theta$, $q_3 = \phi$, und $x(r,\theta,\phi) = (r\,\cos\phi\sin\theta, r\,\sin\phi\sin\theta, r\,\cos\theta)^T$. Es folgt*

$$\varepsilon_r = \begin{bmatrix} \cos\phi\sin\theta \\ \sin\phi\sin\theta \\ \cos\theta \end{bmatrix}, \; \varepsilon_\theta \;=\; \begin{bmatrix} r\,\cos\phi\cos\theta \\ r\,\sin\phi\cos\theta \\ -r\,\sin\theta \end{bmatrix}, \; \varepsilon_\phi \;=\; \begin{bmatrix} -r\,\sin\phi\sin\theta \\ r\,\cos\phi\sin\theta \\ 0 \end{bmatrix}.$$

Diese Vektoren sind orthogonal, also sind Kugelkoordinaten orthogonale Koordinaten in $\mathbb{R}^3$.

Es gibt viele weitere krummlinige Koordinatensysteme, die in manchen Anwendungen wichtig sind; siehe hierzu auch Formelsammlungen wie z. B. [3].

Definition 11.1.7 *Gegeben seien krummlinige Koordinaten $q = (q_1, q_2, q_3)^T$. Eine stetige Funktion $q(t)$, $a \le t \le b$, heißt* Parameterdarstellung einer Kurve in den krummlinigen Koordinaten. *Die eigentliche Kurve in $\mathbb{R}^3$ hat dann natürlich die Parameterdarstellung $x(q(t))$.*

Falls die Parameterdarstellung der Kurve entsprechend oft differenzierbar ist, haben die Ableitungen

$$v(t) \;=\; \frac{d}{dt}x(q(t))\,, \quad b(t) \;=\; \frac{d}{dt}v(t) \;=\; \frac{d^2}{dt^2}x(q(t))$$

die Bedeutung von Geschwindigkeit *bzw.* Beschleunigung *der Bewegung eines Punktes, der sich zur Zeit t im Punkt $x(q(t))$ aufhält. Es ist üblich, Ableitungen nach der Zeit durch einen Punkt statt mit einem Strich zu kennzeichnen, und deshalb schreiben wir hier $\dot{x}$ und $\ddot{x}$ für die erste und zweite Ableitung von x nach der Zeit t.*

Behauptung 11.1.8 *In orthogonalen krummlinigen Koordinaten gelten folgende Darstellungen*

$$v \;=\; \sum_{j=1}^{3} \dot q_j \, \varepsilon_{q_j}\,, \quad b \;=\; \sum_{\ell=1}^{3} \Big[\ddot q_\ell \;+\; \sum_{j,k=1}^{3} \Gamma_{jk}^{\ell} \, \dot q_j \, \dot q_k \Big] \, \varepsilon_{q_\ell}\,.$$

für Geschwindigkeit und Beschleunigung.

Beweis: Wende die Kettenregel und Aufgabe 11.1.3 an! □

Definition 11.1.9 *Gegeben seien krummlinige Koordinaten $q = (q_1, q_2, q_3)^T$. Eine Funktion $u(x)$, definiert auf O_1, kann in der Form $u(x(q))$ auch als Funktion der Variablen q aufgefasst werden. Ihr Definitionsbereich ist dann natürlich gleich O_2, d. h. streng genommen haben wir dann eine andere Funktion. Für die folgenden Überlegungen ist es aber praktisch, für diese neue Funktion einfach wieder u zu schreiben.*

Für eine skalare Funktion u, welche entsprechend oft partiell differenzierbar ist, betrachten wir im Folgenden die Ausdrücke

$$\text{grad } u \;=\; (u_{x_1}, u_{x_2}, u_{x_3})\,, \quad \Delta\, u \;=\; \sum_{k=1}^{3} \frac{\partial^2 u}{\partial x_k^2}\,.$$

Der linke Ausdruck ist natürlich der wohlbekannte Gradient von u, während der zweite Ausdruck als Laplaceoperator*, angewandt auf die Funktion u, bezeichnet wird. Für eine vektorwertige Funktion $v = (v_1, v_2, v_3)^T$ mit entsprechenden Differenzierbarkeitseigenschaften setzen wir noch*

$$\text{div } v \;=\; \sum_{j=1}^{3} \frac{\partial v_j}{\partial x_j}\,, \qquad \text{rot}\, v \;=\; \begin{pmatrix} \frac{\partial v_3}{\partial x_2} & - & \frac{\partial v_2}{\partial x_3} \\ \frac{\partial v_1}{\partial x_3} & - & \frac{\partial v_3}{\partial x_1} \\ \frac{\partial v_2}{\partial x_1} & - & \frac{\partial v_1}{\partial x_2} \end{pmatrix}.$$

Hier nennt man den ersten Ausdruck auch die Divergenz*, den zweiten* die Rotation *von v. Allgemein fasst man Ausdrücke, die Ableitungen einer Funktion enthalten, auch oft als das Ergebnis der Anwendung eines* Differentialoperators *auf diese Funktion auf.*

Für krummlinige Koordinaten besteht oft die Aufgabe darin, die oben eingeführten Ausdrücke so umzuschreiben, dass statt der Ableitungen nach den ursprünglichen Variablen x_j die partiellen Ableitungen nach den neuen Veränderlichen q_k auftreten. Bei orthogonalen Koordinaten gelten dabei die Gleichungen

$$\text{grad } u \;=\; \sum_{j=1}^{3} \frac{1}{g_j} \frac{\partial u}{\partial q_j}\, e_{q_j}^T\,, \qquad \Delta\, u \;=\; \frac{1}{g_1 g_2 g_3} \sum_{k=1}^{3} \frac{\partial}{\partial q_k} \Big(\frac{g_1 g_2 g_3}{g_k} \frac{\partial u}{\partial q_k} \Big)\,.$$

Für die Umrechnung von Divergenz und Rotation wird auf die Literatur verwiesen.

Verwendet man die Abbildung $x = x(q)$ zur Substitution eines Bereichsintegrals wie in Abschnitt 8.5, so gilt

$$|\det x'(q)| \;=\; g_1\, g_2\, g_3.$$

Man sagt deshalb auch, dass das Volumenelement in krummlinigen Koordinaten gegeben ist als

$$dx_1\, dx_2\, dx_3 \;=\; g_1\, g_2\, g_3\, dq_1\, dq_2\, dq_2\,.$$

Dies ist aber mehr als Merkregel zu verstehen!

Aufgabe 11.1.10 *Berechne Divergenz, Rotation, Laplaceoperator und Volumenelement in Polar- und Kugelkoordinaten.*

11.2 Der Gaußsche Integralsatz in der Ebene

Im Folgenden werden Punkte aus $\mathbb{R}^2$ immer mit $(x,y)^T$ bezeichnet.

Definition 11.2.1 *Ein $B \subset \mathbb{R}^2$ heißt* ein Normalbereich *bezüglich y, wenn es ein Intervall $[a,b]$ sowie zwei auf $[a,b]$ stetige Funktionen α, β gibt, sodass*

$$\alpha(x) \ < \ \beta(x) \qquad \forall\ x \in (a,b)\,,$$

$$B \ = \ \{(x,y)^T : \quad \alpha(x) \ \le \ y \ \le \ \beta(x)\,, \quad a \ \le x \ \le b \ \}\,.$$

Sinngemäß wird ein Normalbereich bezüglich x definiert.

Ein solcher Normalbereich B bezüglich y ist immer messbar. Als Randkurve *von B bezeichnet man die Kurve $\gamma = \gamma_1 + \gamma_4 + \gamma_3 + \gamma_4$ mit den vier Teilkurven*

$$\gamma_1 \ : \ \begin{pmatrix} x_1(t) \\ y_1(t) \end{pmatrix} \ = \ \begin{pmatrix} t \\ \alpha(t) \end{pmatrix}, \qquad a \le t \le b\,,$$

$$\gamma_2 \ : \ \begin{pmatrix} x_2(t) \\ y_2(t) \end{pmatrix} \ = \ \begin{pmatrix} b \\ t \end{pmatrix}, \qquad \alpha(b) \le t \le \beta(b)\,,$$

$$\gamma_3 \ : \ \begin{pmatrix} x_3(t) \\ y_3(t) \end{pmatrix} \ = \ \begin{pmatrix} -t \\ \beta(-t) \end{pmatrix}, \qquad -b \le t \le -a\,,$$

$$\gamma_4 \ : \ \begin{pmatrix} x_4(t) \\ y_4(t) \end{pmatrix} \ = \ \begin{pmatrix} a \\ -t \end{pmatrix}, \qquad -\beta(a) \le t \le -\alpha(a)\,.$$

Dabei haben wir die Orientierung der Randkurve γ so gewählt, dass der Bereich B beim Durchlaufen immer links liegt. Man spricht dann auch von einer positiv orientierten Randkurve.

Sei B Normalbereich bezüglich y mit stückweise glatter Randkurve, und sei $f \ : B \longrightarrow \mathbb{R}$ so, dass die partielle Ableitung f_y auf B existiert und f sowie f_y auf B stetig sind. Wir setzen $\int_\gamma f(x,y)\,dx$ gleich dem Kurvenintegral des Vektorfeldes $(f(x,y),0)^T$ über die Randkurve γ. Es folgt dann durch Betrachtung der vier Teile des Randes γ, dass

$$-\int_\gamma f(x,y)\,dx \ = \ \int_a^b \Big(f(x,\beta(x)) \ - \ f(x,\alpha(x))\Big)\ dx \ = \ \int_B f_y(x,y)\ d(x,y)\,.$$

Wenn B auch Normalbereich bezüglich x ist, und wenn f_x existiert und auf B stetig ist, so folgt analog

$$\int_\gamma f(x,y)\,dy \ = \ \int_B f_x(x,y)\ d(x,y)\,.$$

Daher gilt der folgende Satz:

Satz 11.2.2 (Gaußscher Integralsatz in der Ebene) *Sei B ein Normalbereich bezüglich x und y, und sei die Randkurve γ von B positiv orientiert und stückweise glatt. Sei ferner G ein Gebiet welches B umfasst, und seien $f,g : G \longrightarrow \mathbb{R}$ so, dass f_x und g_y existieren, und dass f,g,f_x,g_y auf G stetig sind. Dann gilt*

$$\int_B \Big(f_x(x,y) \ - \ g_y(x,y)\ \Big)\ d(x,y) \ = \ \int_\gamma \Big(f(x,y)\,dy \ + \ g(x,y)\,dx\ \Big)\,.$$

Beispiel 11.2.3 *Setzt man $f(x,y) = x/2$, $g(x,y) = -y/2$, so folgt aus dem obigen Satz*

$$|B| \;=\; \frac{1}{2}\int_\gamma \Big(x\,dy \;-\; y\,dx\Big)\;.$$

Also kann man den Flächeninhalt von B durch ein Kurvenintegral berechnen. Ist B etwa die Kreisscheibe um $(0,0)$ mit Radius $r > 0$, also $x(t) = r\,\cos t$, $y(t) = r\,\sin t$, $0 \le t \le 2\pi$, so folgt

$$|B| \;=\; \frac{1}{2}\int_0^{2\pi} \Big(x(t)\,y'(t) \;-\; y(t)\,x'(t)\Big)\;dt \;=\frac{r^2}{2}\int_0^{2\pi}\Big(\cos^2 t \;+\; \sin^2 t\Big)\;dt\,,$$

also $|B| = r^2\,\pi$.

11.3 Vektorprodukt und Flächen im Raum

Definition 11.3.1 *Für $a = (a_1, a_2, a_3)^T$, $b = (b_1, b_2, b_3)^T \in \mathbb{R}^3$ sei das* Vektorprodukt *oder* Kreuzprodukt *oder* äußere Produkt *definiert als*

$$a \times b \;=\; \begin{pmatrix} a_2\,b_3 \;-\; a_3\,b_2 \\ a_3\,b_1 \;-\; a_1\,b_3 \\ a_1\,b_2 \;-\; a_2\,b_1 \end{pmatrix}\;.$$

Sind e_1, e_2, e_3 die drei kanonischen Basisvektoren in $\mathbb{R}^3$, so gilt die Merkregel

$$a \times b \;=\; \det\begin{bmatrix} e_1 & e_2 & e_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}\;.$$

Beachte aber, dass rechts genau genommen keine Determinante steht, da die erste Zeile der Matrix Vektoren enthält.

Man beweist leicht folgende *Rechenregeln* für das Vektorprodukt:

- $a \times b = -b \times a$.
- $a \times a = 0$.
- $(\lambda\,a) \times b = a \times (\lambda\,b) = \lambda\;(a \times b) \quad \forall\;\lambda \in \mathbb{R}$.
- $a \times (b + c) = a \times b + a \times c$.
- $a^T\,(b \times c) = \det\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}\;.$

Aus der letzten Regel folgt, dass $a \times b$ zu a und b orthogonal ist. Weiter gilt folgende Produktregel für die Differenziation eines Vektorproduktes:

Sind $a(t)$, $b(t)$ glatte Parameterdarstellungen von Kurven in $\mathbb{R}^3$, so ist

$$\frac{d}{dt}\,\big(a(t) \times b(t)\big) \;=\; \frac{da(t)}{dt} \times b(t) \;+\; a(t) \times \frac{db(t)}{dt}\;.$$

Hier muss aber unbedingt auf die Reihenfolge der Faktoren geachtet werden.

Definition 11.3.2 *Sei $G \subset \mathbb{R}^2$ ein zweidimensionales Gebiet, und sei $x : G \longrightarrow \mathbb{R}^3$ stetig auf G. Für eine kompakte Teilmenge $K \subset G$ von G heißt die Restriktion von x auf K* Parameterdarstellung einer Fläche *F in $\mathbb{R}^3$. Wir wollen meistens $x(s,t)$ für einen Punkt der Fläche schreiben. Ist x in G nach beiden Variablen partiell differenzierbar, so heißt der Vektor*

$$n(s,t) \;=\; \frac{\partial x(s,t)}{\partial s} \times \frac{\partial x(s,t)}{\partial t} \qquad (s,t)^T \in K$$

der Normalenvektor der Fläche *im Punkt $x(s,t)$. Falls x auf G stetig partiell differenzierbar und der Normalenvektor nie der Nullvektor ist, heißt die Fläche* glatt. *Die Zahl $\|n(s,t)\|$ gibt dann ungefähr an, um wieviel größer ein kleines Stück der Fläche um den Punkt $x(s,t)$ herum im Vergleich zu einem entsprechenden Stück von B um den Punkt $(s,t)^T$ herum ist. Ist f eine auf dem Träger der Fläche F definierte Funktion mit Werten in $\mathbb{R}$, so definieren wir das* Integral von f über die Fläche F *durch*

$$\int_F f(x)\; d\sigma \;=\; \int_K f(x(s,t))\, \|n(s,t)\|\; d(s,t)\,.$$

Wenn $f(x) \equiv 1$ ist, dann ist das Flächenintegral gerade gleich dem Flächeninhalt von F. Allgemeiner kann man $f(x)$ als die Dichte einer auf F verteilten Masse ansehen, und in diesem Fall ist das Integral gleich der Gesamtmasse.

11.4 Der Stokessche Integralsatz

Definition 11.4.1 *Ist F eine Fläche in $\mathbb{R}^3$ mit einer Parameterdarstellung $x : K \longrightarrow \mathbb{R}^3$, und ist der Rand von K eine Kurve γ_K mit Parameterdarstellung $(s(\tau), t(\tau))^T$, $a \le \tau \le b$, so heißt die Kurve γ_F mit Parameterdarstellung $x(s(\tau), t(\tau))$, $a \le \tau \le b$,* die Randkurve von F. *Wir nennen γ_F genau dann positiv orientiert, wenn γ_K positv orientiert ist. Ist $n(s,t)$ der Normalenvektor auf einer glatten Fläche F, so sei $\nu(s,t) = \|n(s,t)\|^{-1} n(s,t)$ der zugehörige* normierte Normalenvektor*.*

Satz 11.4.2 (Stokesscher Integralsatz) *Mit den oben eingeführten Bezeichnungen gilt, unter genügend starken Voraussetzungen an die Fläche F und das Vektorfeld f, die Gleichung*

$$\int_F \nu^T \operatorname{rot} f\, d\sigma \;=\; \int_{\gamma_F} f^T(x)\, dx\,.$$

Für die genauen Voraussetzungen und einen Beweis des Satzes, siehe etwa *W. Walter* [11, Abschnitt 8.12].

11.5 Der Gaußsche Integralsatz

Definition 11.5.1 *Eine Fläche mit Parameterdarstellung $x : K \longrightarrow \mathbb{R}^3$ heißt* doppelpunktfrei*, falls K ein Normalbereich ist und die Parameterdarstellung x im Inneren von K injektiv ist. Wir sagen, dass ein Bereich $B \subset R^3$* einen glatten Rand hat*, wenn wir den Rand ∂B als eine glatte doppelpunktfreie Fläche schreiben können. Wir nennen den Rand* positiv orientiert*, falls der normierte Normalenvektor ν immer ins Äußere von B zeigt.*

Satz 11.5.2 (Gaußscher Integralsatz in $\mathbb{R}^3$)
Mit den oben eingeführten Bezeichnungen gilt, unter genügend starken Voraussetzungen an den Bereich B und das Vektorfeld f, die Gleichung

$$\int_B \operatorname{div} f(x)\, dx \;=\; \int_{\partial B} f^T(x)\, \nu\, d\sigma\,,$$

wobei der Rand ∂B positiv orientiert sei.

Ein Beweis kann ebenfalls in *W. Walter* [11, Abschnitt 8.6] nachgelesen werden.

Index