Kapitel 9

Metrische Räume und Banachräume

9.1 Metrik und Norm

Definition 9.1. *(Metrik auf einer Menge X) Eine Metrik (oder Abstandsfunktion) ist eine Abbildung* $d : X \times X \to \mathbb{R}, \quad (x, y) \mapsto d(x, y)$ *mit drei Eigenschaften*

(i) $d(x, y) \geq 0$ *für alle* $x, y \in X$ *und* $d(x, y) = 0 \Leftrightarrow x = y$ *(Positivität).*

(ii) $d(x, y) = d(y, x)$ *(Symmetrie).*

(iii) $d(x, y) \leq d(x, z) + d(z, y)$ *für alle* $x, y, z \in X$ *(Dreiecksungleichung).*

Beispiel 9.2. (i) *auf jeder Menge X definiert* $d(x, y) = \begin{cases} 0 & \text{für } x = y \\ 1 & \text{für } x \neq y \end{cases}$

die sogenannte diskrete Metrik.

(ii) *Auf* $\mathbb{K}$ *definiert* $d(x, y) = |x - y|$ *eine Metrik.*

(iii) *Auf jeder nicht leeren Teilmenge* $A \subset X$ *eines metrischen Raumes* (X, d) *definiert die Einschränkung von d auf* $A \times A \to \mathbb{R}$ *eine Metrik.*

(iv) *Auf dem kartesischen Produkt zweier metrischer Räume definiert die Summe beider Metriken eine Metrik. Sie heißt Metrik des kartesischen Produktes.*

(v) *Die Einschränkung der Metrik (ii) auf die Vereinigung der inversen der natürlichen Zahlen mit* $\{0\}$ *definiert eine Metrik auf* $\bar{\mathbb{N}} = \mathbb{N} \cup \{\infty\} \simeq \{\frac{1}{n} \mid n \in \mathbb{N}\} \cup \{0\}$*:*

$$d(n, m) = \frac{|n - m|}{nm} \quad d(\infty, n) = d(n, \infty) = \frac{1}{n} \quad d(\infty, \infty) = 0 \quad \text{für alle } n, m \in \mathbb{N}.$$

Die Menge $V = \mathbb{K}^n$ erfüllt mit $(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$ und $0 = (0, \ldots, 0)$ die Axiome $A1$. Außerdem besitzt sie eine Skalarmultiplikation

$$\mathbb{K} \times V \to V, \qquad (\lambda, (x_1, \ldots, x_n)) \mapsto \lambda \cdot (x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n),$$

so dass sie einen Vektorraum über dem Körper $\mathbb{K}$ bildet:

Definition 9.3. *Ein $\mathbb{K}$-Vektorraum V ist eine Menge V zusammen mit Abbildungen:*

$$+ : V \times V \to V, \qquad (v, w) \mapsto v + w \qquad \cdot : \mathbb{K} \times V \to V, \qquad (\lambda, v) \mapsto \lambda \cdot v,$$

die die Axiome $A1$ der Addition und für alle $\lambda, \mu \in \mathbb{K}$ und $v, w \in V$ folgendes erfüllt:

$$\begin{aligned} \lambda \cdot (v + w) &= \lambda \cdot v + \lambda \cdot w, & (\lambda + \mu) \cdot v &= \lambda \cdot v + \mu \cdot v, \\ (\lambda \cdot \mu) \cdot v &= \lambda \cdot (\mu \cdot v), & 1 \cdot v &= v. \end{aligned}$$

Definition 9.4. *Eine Norm auf einem reellen bzw. komplexen Vektorraum V ist eine Abbildung $\| \cdot \| : V \to \mathbb{R}$, $\quad x \mapsto \|x\|$ mit folgenden drei Eigenschaften:*

(i) $\|x\| \geq 0$ *für alle $x \in V$ und $\|x\| = 0 \iff x = 0$*

(ii) $\|\lambda \cdot x\| = |\lambda| \cdot \|x\|$ *für $\lambda \in \mathbb{K}$ und $x \in V$*

(iii) $\|x + y\| \leq \|x\| + \|y\|$ *für alle $x, y \in V$*

Satz 9.5. *Für jede Norm ist $d : V \times V \to \mathbb{R}$, $\quad (x, y) \mapsto d(x, y) = \|x - y\|$ eine Metrik.*

Beweis:(i) folgt aus (i) der Definition einer Norm.
(ii) $d(y, x) = \|y - x\| = \|(-1)(x - y)\| = |-1|\|x - y\| = \|x - y\| = d(x, y)$
(iii) $d(x, y) = \|x - y\| = \|x - z + z - y\| \leq \|x - z\| + \|z - y\| = d(x, z) + d(z, y)$.**q.e.d.**

Im folgenden werden wir mit der Vorgabe einer Norm auf einem Vektorraum immer auch die entsprechende induzierte Metrik auf diesem Vektorraum betrachten. Deshalb fassen wir jeden normierten Vektorraum auch als metrischen Raum auf.

Definition 9.6. *(Euklidische Norm und Metrik) Auf $\mathbb{R}^n$ definiert*

$$\|(x_1, \ldots, x_n)\| = \sqrt{|x_1|^2 + \ldots + |x_n|^2} = \sqrt{x_1^2 + \ldots + x_n^2}$$

die euklidische Norm. Die entsprechende Metrik heißt dann euklidische Metrik.

Die Eigenschaft (iii) heißt dabei Minkowski–Ungleichung:

$$\sqrt{(x_1 + y_1)^2 + \ldots + (x_n + y_n)^2} \leq \sqrt{x_1^2 + \ldots + x_n^2} + \sqrt{y_1^2 + \ldots + y_n^2}.$$

Um diese zu beweisen zeigen wir zuerst die

Cauchy–Schwarzsche Ungleichung 9.7.

$$|x_1y_1| + \ldots + |x_ny_n| \le \sqrt{x_1^2 + \ldots + x_n^2}\sqrt{y_1^2 + \ldots + y_n^2}$$

Beweis*: Wegen $|x_iy_i| = |x_i| \cdot |y_i|$, $x_i^2 = |x_i|^2$ und $y_i^2 = |y_i|^2$ genügt es offenbar

$$x_1y_1 + \ldots + x_ny_n \le \sqrt{x_1^2 + \ldots + x_n^2}\sqrt{y_1^2 + \ldots + y_n^2}$$

zu zeigen. Das ist äquivalent zu

$$(x_1y_1 + \ldots + x_ny_n)^2 \le (x_1^2 + \ldots + x_n^2)(y_1^2 + \ldots + y_n^2).$$

Für $y = (y_1, \ldots, y_n) = 0$ ist die Aussage trivial. Für $y \neq 0$ gilt

$$\left\| \|y\|x - \frac{x_1y_1 + \ldots + x_ny_n}{\|y\|}y \right\|^2 = \sum_{\imath=1}^n \left(\|y\|x_\imath - \frac{x_1y_1 + \ldots + x_ny_n}{\|y\|}y_\imath \right)^2$$
$$= \sum_{\imath=1}^n \left(\|y\|^2x_\imath^2 + \frac{(x_1y_1 + \ldots + x_ny_n)^2}{\|y\|^2}y_\imath^2 - 2(x_1y_1 + \ldots + x_ny_n)x_\imath y_\imath \right)$$
$$= \|y\|^2\|x\|^2 + \frac{(x_1y_1 + \ldots + x_ny_n)^2}{\|y\|^2}\|y\|^2 - 2(x_1y_1 + \ldots + x_ny_n)^2$$
$$= \|y\|^2\|x\|^2 - (x_1y_1 + \ldots + x_ny_n)^2.$$

Weil diese Ausdrücke nicht negativ sind folgt $(x_1y_1 + \ldots + x_ny_n)^2 \le \|x\|^2\|y\|^2$. **q.e.d.**
Beweis der Minkowski Ungleichung:

$$\begin{aligned}(x_1 + y_1)^2 + \ldots + (x_n + y_n)^2 &= \|x\|^2 + \|y\|^2 + 2(x_1y_1 + \ldots + x_ny_n) \\ &\le \|x\|^2 + \|y\|^2 + 2\|x\| \cdot \|y\| \le (\|x\| + \|y\|)^2. \quad \textbf{q.e.d.}\end{aligned}$$

Analog definiert $\|(x_1, \ldots, x_n)\| = \sqrt{x_1\bar{x}_1 + \ldots + x_n\bar{x}_n}$ eine Norm auf $\mathbb{C}^n$. Identifizieren wir $\mathbb{C}$ mit $\mathbb{R}^2$ (Realteil und Imaginärteil), dann ist $\mathbb{C}^n \simeq (\mathbb{R}^2)^n \simeq \mathbb{R}^{2n}$.

Übungsaufgabe 9.8. *Die euklidische Norm auf $\mathbb{R}^{2n}$ induziert durch diese Identifikation auf $\mathbb{C}^n$ die Norm $\|(x_1, \ldots, x_n)\| = \sqrt{x_1\bar{x}_1 + \ldots + x_n\bar{x}_n}$.*

Definition 9.9. *Für $x = (x_1, \ldots, x_n) \in \mathbb{K}^n$ und $1 \le p < \infty$ sei*

$$\|x\|_p = (|x_1|^p + \ldots + |x_n|^p)^{1/p}$$

Wir hatten in einer Übungsaufgabe gesehen, dass $\|x\|_p$ im Grenzwert $p \to \infty$ gegen

$$\|x\|_\infty = \max\{|x_1|, \ldots, |x_n|\}$$

konvergiert. Das sogenannte hermitische Skalarprodukt wird definiert als

$$\langle x, y \rangle = x_1\bar{y}_1 + \ldots + x_n\bar{y}_n \in \mathbb{K}, \qquad \text{für } x, y \in \mathbb{K}^n.$$

Für $y \in \mathbb{R}^n$ setzen wir hierbei $(\bar{y}_1, \ldots, \bar{y}_n) = \bar{y} = y$.

Satz 9.10. *(Höldersche Ungleichung) Seien $p \geq 1$ und $q \geq 1$ mit $\frac{1}{p} + \frac{1}{q} = 1$.*

Dann gilt $\quad |\langle x, y\rangle| \leq |x_1 y_1| + \ldots + |x_n y_n| \leq \|x\|_p \cdot \|y\|_q$ *für alle $x, y \in \mathbb{K}^n$.*

Für $p = q = 2$ erhalten wir wieder die Cauchy-Schwarzsche Ungleichung

$$|x_1 y_1| + \ldots + |x_n y_n| \leq \|x\| \cdot \|y\|.$$

Beweis: Wenn $p = 1$ und $q = \infty$ gilt

$$|x_1 y_1| + \ldots + |x_n y_n| \leq (|x_1| + \ldots + |x_n|) \max\{|y_1|, \ldots, |y_n|\}.$$

Den Fall $p = \infty$ und $q = 1$ erhalten wir durch vertauschen von x und y. Wir können also $1 < p, q < \infty$ annehmen, und dass $\|x\|_p \neq 0 \neq \|y\|_q$ gilt, weil die Ungleichung sonst offensichtlich ist. Dann folgt aus der Youngschen Ungleichung für alle $k = 1, \ldots, n$

$$\frac{|x_k y_k|}{\|x\|_p \|y\|_q} = \frac{|x_k|}{\|x\|_p} \frac{|y_k|}{\|y\|_q} \leq \frac{1}{p} \frac{|x_k|^p}{\|x\|_p^p} + \frac{1}{q} \frac{|y_k|^q}{\|y\|_q^q}$$

Nach Summation über $k = 1, \ldots, n$ erhalten wir

$$\frac{|x_1 y_1| + \ldots + |x_n y_n|}{\|x\|_p \|y\|_q} \leq \frac{1}{p} + \frac{1}{q} = 1.$$ **q.e.d.**

Satz 9.11. *(Minkowski Ungleichung) Sei $p \geq 1$ und $x, y \in \mathbb{K}^n$, dann gilt*

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p$$

Korollar 9.12. *Für alle $1 \leq p \leq \infty$ und $n \in \mathbb{N}$ ist $\|\cdot\|_p : \mathbb{K}^n \to \mathbb{R}$ eine Norm.* **q.e.d.**

Beweis der Minkowski Ungleichung: Für $p = 1$ oder $p = \infty$ folgt sie aus der Dreiecksungleichung. Sei also $1 < p, q < \infty$ mit $\frac{1}{p} + \frac{1}{q} = 1 \Leftrightarrow p + q = pq \Leftrightarrow p = (p-1)q$.

$$\begin{aligned}
|x_1 + y_1|^p + \ldots + |x_n + y_n|^p &= |x_1 + y_1| \cdot |x_1 + y_1|^{p-1} + \ldots + |x_n + y_n| \cdot |x_n + y_n|^{p-1} \\
&\leq (|x_1| + |y_1|)|x_1 + y_1|^{p-1} + \ldots + (|x_n| + |y_n|)|x_n + y_n|^{p-1} \\
&\leq (\|x\|_p + \|y\|_p)(|x_1 + y_1|^{(p-1)q} + \ldots + |x_n + y_n|^{(p-1)q})^{1/q} \\
&= (\|x\|_p + \|y\|_p)\|x + y\|_p^{p/q}
\end{aligned}$$

Die zweite Zeile wurde ausmultipliziert und jeweils die Höldersche Ungleichung benutzt. Es folgt $\|x + y\|_p^p \leq (\|x\|_p + \|y\|_p)\|x + y\|_p^{p/q}$. Für $\|x + y\|_p = 0$ ist die Aussage trivial. Aus $\|x + y\|_p \neq 0$ folgt $\|x + y\|_p = \|x + y\|_p^{p(1-\frac{1}{q})} = \|x + y\|_p^{(p-\frac{p}{q})} \leq \|x\|_p + \|y\|_p$. **q.e.d.**

Definition 9.13. *(offener Ball, Umgebung, offene Menge) Ein offener Ball in (X, d) mit Zentrum $x \in X$ und Radius $r > 0$ ist die Menge $B(x, r) = \{y \in X \mid d(x, y) < r\}$. Eine Umgebung eines Punktes $x \in X$ ist eine Menge $O \subset X$, die für ein $r > 0$ einen Ball $B(x, r)$ enthält. Eine offene Menge $O \subset X$ ist eine Teilmenge, die eine Umgebung aller ihrer Punkte ist, d.h. für alle $x \in O$ gibt es ein $\epsilon > 0$ mit $B(x, \epsilon) \subset O$.*

Beispiel 9.14. *In $\mathbb{R}$ besteht der Ball $B(x,r)$ aus $(x-r,x+r)$. Im $\mathbb{R}^n$ besteht der Ball $B(x,r)$ aus allen Punkten, deren euklidischer Abstand zu x kleiner ist als r.*

Alle offenen Bälle $B(x,r)$ sind tatsächlich Umgebungen von x: Für $y \in B(x,r)$ ist $d(x,y) < r$. Sei $z \in B(y, r-d(x,y))$. Dann gilt $d(x,z) \leq d(x,y) + d(y,z) < r$, also auch $B(y, r-d(x,y)) \subset B(x,r)$. Deshalb sind die offenen Bälle tatsächlich offen.

Offenbar ist die beliebige Vereinigung von offenen Mengen wieder offen. Für zwei offene Mengen O und O' und $x \in O \cap O'$ gibt es $r > 0$ und $r' > 0$ mit $B(x,r) \subset O$ und $B(x,r') \subset O'$. Also ist $B(x, \min\{r,r'\}) \subset B(x,r) \cap B(x,r') \subset O \cap O'$, und $O \cap O'$ offen. Damit ist auch die Schnittmenge von endlich vielen offenen Mengen wieder offen.

Definition 9.15. *(abgeschlossene Mengen, Abschluss) Die Komplemente von offenen Mengen heißen abgeschlossen. Der Abschluss $\bar{A}$ einer Menge A ist die Schnittmenge aller abgeschlossenen Mengen, die A enthalten.*

Wegen der Regel von de Morgan, sind beliebige Schnitte und endliche Vereinigungen von abgeschlossenen Mengen wieder abgeschlossen. Deshalb ist eine Menge genau dann abgeschlossen, wenn sie mit ihrem Abschluss übereinstimmt.

Definition 9.16. *Zwei Normen $\| \quad \|_1$ und $\| \quad \|_2$ heißen äquivalent, wenn es Konstanten $C_1, C_2 > 0$ gibt, so dass für alle $v \in V$ folgendes gilt*

$$\|v\|_1 \leq C_1 \|v\|_2 \qquad \textit{und} \qquad \|v\|_2 \leq C_2 \|v\|_1.$$

Diese Relation zwischen Normen ist eine Äquivalenzrelation, denn aus

$$\begin{array}{cccc} \|v\|_1 \leq C_1\|v\|_2 \quad \|v\|_2 \leq C_2\|v\|_1 & & \|v\|_2 \leq C_3\|v\|_3 & \|v\|_3 \leq C_4\|v\|_2 \\ \text{folgt} \qquad \|v\|_1 \leq C_1C_3\|v\|_3 & \text{und} & \|v\|_3 \leq C_4C_2\|v\|_1. & \end{array}$$

Wenn $\|v\|_1 \leq C_1\|v\|_2$ für alle $v \in V$ gilt, dann ist C_1 wegen der Positivität von $\| \cdot \|_1$ positiv und die Bälle bezüglich der Normen mit den gleichen Indizes erfüllen $B_2(w,r) \subset B_1(w, C_1 r)$ für alle $w \in V$ und $r > 0$. Also sind alle Umgebungen bezüglich $\| \cdot \|_1$ auch Umgebungen bezüglich $\| \cdot \|_2$. Wenn umgekehrt $B_1(0,1)$ auch eine Umgebung von 0 bezüglich $\| \cdot \|_2$ ist, folgt $\frac{\|v\|_1}{C_1\|v\|_2+\epsilon} = \|\frac{v}{C_1\|v\|_2+\epsilon}\|_1 < 1$ wegen $\frac{v}{C_1\|v\|_2+\epsilon} \in B_2(0, \frac{1}{C_1})$ für $\epsilon > 0$ und $v \in V$ aus $B_2(0, \frac{1}{C_1}) \subset B_1(0,1)$, und im Grenzwert $\epsilon \downarrow 0$ wieder $\|v\|_1 \leq C_1\|v\|_2$. Also stimmen die Umgebungen und damit die offenen und abgeschlossenen Mengen zweier Normen auf V genau dann überein, wenn die Normen äquivalent sind.

Beispiel 9.17. *Auf den Vektorräumen $\mathbb{K}^n$ haben wir für $1 \leq p \leq \infty$ die Normen*

$$\| \cdot \|_p : \mathbb{K}^n \to \mathbb{R}, \quad v \mapsto \|v\|_p = \begin{cases} \left(\sum_{i=1}^n |v_i|^p \right)^{1/p} & \textit{für } p < \infty \\ \max\{|v_1|, \ldots |v_n|\} & \textit{für } p = \infty. \end{cases}$$

eingeführt. Sie sind alle äquivalent, weil für $1 \leq p < \infty$ und $v \in \mathbb{K}^n$ folgendes gilt:

$$\|v\|_\infty \leq \|v\|_p \leq (n\|v\|_\infty^p)^{1/p} = n^{1/p}\|v\|_\infty.$$

9.2 Vollständigkeit und Kompaktheit

Zunächst verallgemeinern wir einige Aussagen über Zahlenfolgen auf allgemeine Folgen in metrischen Räumen.

Definition 9.18. *(Folgen und Cauchyfolgen) Eine Folge $(x_n)_{n\in\mathbb{N}}$ in einem metrischen Raum (X,d) ist eine Abbildung von $\mathbb{N}$ nach X, mit $n \mapsto x_n$.*

Eine Folge $(x_n)_{n\in\mathbb{N}}$ in einem metrischen Raum (X,d) konvergiert gegen $x \in X$, wenn es für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass $d(x_n, x) < \epsilon$ für alle $n \geq N$ gilt.

Eine Folge $(x_n)_{n\in\mathbb{N}}$ heißt Cauchyfolge, wenn es für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass $d(x_n, x_m) < \epsilon$ für alle $n, m \geq N$ gilt.

Offenbar konvergiert eine Folge $(x_n)_{n\in\mathbb{N}}$ genau dann gegen einen Punkt x, wenn jede Umgebung von x alle bis auf endlich viele Folgenglieder enthält. Deshalb hängt der Begriff der Konvergenz nur von der Wahl der offenen Mengen ab.

Ein Punkt x gehört genau dann zu dem Abschluss $\bar{A}$, wenn es keine offene Menge gibt, die x enthält und einen leeren Schnitt mit A hat. Dies ist wiederum äquivalent dazu, dass es für jedes $n \in \mathbb{N}$ ein Element a_n in dem Ball $B(x, \frac{1}{n}) \cap A$ gibt, oder auch dazu, dass es eine Folge in A gibt, die gegen x konvergiert. Damit haben wir gezeigt:

Lemma 9.19. *Der Abschluss einer Teilmenge eines metrischen Raumes besteht aus allen Grenzwerten von Folgen innerhalb der Teilmenge, die in dem metrischen Raum konvergieren. Und eine Teilmenge ist genau dann abgeschlossen, wenn die Grenzwerte von allen konvergenten Folgen in der Teilmenge auch zu der Menge gehören.* **q.e.d.**

Wenn die Folge $(x_n)_{n\in\mathbb{N}}$ gegen x konvergiert, dann gibt es für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$, so dass $d(x_n, x) < \frac{\epsilon}{2}$ und $d(x_m, x) < \frac{\epsilon}{2}$ für alle $n, m \geq N$ gilt. Dann gilt auch $d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \epsilon$. Damit haben wir gezeigt:

Satz 9.20. *In einem metrischen Raum sind konvergente Folgen Cauchyfolgen.* **q.e.d.**

Definition 9.21. *Ein metrischer Raum (X,d) heißt vollständig, wenn auch jede Cauchyfolge konvergiert.*

Definition 9.22. *Ein vollständiger normierter Vektorraum heißt Banachraum.*

Wegen dem Vollständigkeitsaxiom sind $\mathbb{R}^n$ und $\mathbb{C}^n$ für alle $n \in \mathbb{N}$ Banachräume. Wegen Lemma 9.19 ist eine Teilmenge eines vollständigen metrischen Raumes genau dann ein vollständiger metrischer Raum, wenn sie abgeschlossen ist.

Definition 9.23. *(kompakt) Eine Teilmenge der Menge der offenen Teilmengen von (X,d) heißt offene Überdeckung von X, wenn jedes Element von X in mindestens einer der offenen Mengen enthalten ist. Der metrische Raum heißt kompakt, wenn jede offene Überdeckung eine endliche Teilüberdeckung besitzt, d.h. jede Menge von offenen Teilmengen von X, die X überdeckt, enthält eine endliche Teilmenge, die X überdeckt.*

Satz 9.24. *Für einen metrischen Raum (X, d) sind folgende Aussagen äquivalent:*

(i) *(X, d) ist kompakt.*

(ii) *Jede Folge in (X, d) besitzt eine konvergente Teilfolge.*

(iii) *(X, d) ist vollständig und für jedes $\epsilon > 0$ besitzt (X, d) eine endliche Überdeckung mit offenen Bällen vom Radius ϵ.*

Beweis: (i)$\Rightarrow$(ii): Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge ohne Häufungspunkt. Dann sind für alle $n \in \mathbb{N}$ die Mengen $F_n = \{x_m \mid m \geq n\}$ abgeschlossen, weil ihr Abschluss neben den Elementen von F_n nur aus Häufungspunkten von $(x_n)_{n\in\mathbb{N}}$ besteht. Weil $\bigcap_{n\in\mathbb{N}} F_n$ nur Häufungspunkte von $(x_n)_{n\in\mathbb{N}}$ enthält, bilden die Mengen $(X \setminus F_n)_{n\in\mathbb{N}}$ eine offene Überdeckung von X. Offenbar ist der Schnitt von endlich vielen Mengen der Mengen $(F_n)_{n\in\mathbb{N}}$ nicht leer. Also besitzt die Überdeckung $(X \setminus F_n)_{n\in\mathbb{N}}$ keine endliche Teilüberdeckung.

(ii)$\Rightarrow$(iii): Sei also (X, d) ein metrischer Raum, der (ii) erfüllt. Dann besitzt jede Cauchyfolge $(x_n)_{n\in\mathbb{N}}$ einen Häufungspunkt x. Dann gibt es für jedes $\epsilon > 0$ eine natürliche Zahl $N \in \mathbb{N}$, so dass alle $m, n \geq N$ auch $d(x_n, x_m) < \frac{\epsilon}{2}$ erfüllen. Weil x ein Häufungspunkt ist, gibt es ein $m \geq N$, so dass $d(x_m, x) < \frac{\epsilon}{2}$ gilt. Also gilt

$$d(x_n, x) \leq d(x_n, x_m) + d(x_m, x) < \epsilon \quad \text{für alle } n \geq N.$$

Also konvergiert $(x_n)_{n\in\mathbb{N}}$ gegen x und damit ist (X, d) vollständig. Sei $\epsilon > 0$ so gewählt, dass es keine endliche Überdeckung von X mit Bällen vom Radius ϵ gibt. Wir definieren induktiv eine Folge $(x_m)_{m\in\mathbb{N}}$, so dass $x_{n+1} \in X \setminus \bigcup_{m=1}^{n} B(x_m, \epsilon)$ für jedes $n \in \mathbb{N}$ gilt. Daraus folgt $d(x_n, x_m) \geq \epsilon$ für alle $n \neq m \in \mathbb{N}$. Also besitzt sie keine Teilfolge, die eine Cauchyfolge ist, und damit auch keinen Häufungspunkt im Widerspruch zu (ii).

(iii)$\Rightarrow$(i): Wir nehmen an (X, d) erfüllt Bedingung (iii) und $(U_\lambda)_{\lambda\in L}$ sei eine Überdeckung von (X, d), die keine endliche Teilüberdeckung besitzt. Dann definieren wir induktiv eine Folge $(x_n)_{n\in\mathbb{N}}$, so dass die Bälle $B(x_n, 2^{-n})$ nicht durch endlich viele $(U_\lambda)_{\lambda\in L}$ überdeckt werden und für alle $n \in \mathbb{N}$ die Bälle $B(x_{n+1}, 2^{-(n+1)})$ und $B(x_n, 2^{-n})$ nicht disjunkt sind. Weil nämlich (X, d) durch endliche viele Bälle vom Radius $2^{-(n+1)}$ überdeckt wird, aber $B(x_n, 2^{-n})$ nicht durch endlich viele $(U_\lambda)_{\lambda\in L}$ überdeckt wird, gibt es mindestens einen dieser endlich vielen Bälle vom Radius $2^{-(n+1)}$, der nichtleeren Schnitt mit $B(x_n, 2^{-n})$ hat und nicht durch endlich viele $(U_\lambda)_{\lambda\in L}$ überdeckt wird. Wegen

$$d(x_n, x_{n+1}) < 2^{-n} + 2^{-(n+1)} = \frac{3}{2 \cdot 2^n} \quad \text{und} \quad \sum_{n=m}^{\infty} \frac{3}{2 \cdot 2^n} = \frac{3}{2 \cdot 2^m} \cdot \frac{1}{1 - \frac{1}{2}} = \frac{3}{2^m}$$

ist die Folge $(x_n)_{n\in\mathbb{N}}$ eine Cauchyfolge. Der Grenzwert x erfüllt $d(x_n, x) \leq \frac{3}{2^n}$ für alle $n \in \mathbb{N}$ und gehört für ein $\lambda \in L$ zu der offenen Menge U_λ, die $B(x, \epsilon)$ für ein $\epsilon > 0$ enthält. Für genügend großes m ist dann $\frac{1}{2^{m-2}} \leq \epsilon$, und wegen $d(x_m, x) \leq \frac{3}{2^m}$ auch

$$B(x_m, 2^{-m}) \subset B(x, 2^{-m} + 3 \cdot 2^{-m}) = B(x, 2^{2-m}) \subset B(x, \epsilon).$$

Also wird der Ball $B(x_m, 2^{-m})$ durch U_λ für ein $\lambda \in L$ überdeckt im Widerspruch zu der Annahme, dass $(U_\lambda)_{\lambda \in L}$ keine endliche Teilüberdeckung besitzt. **q.e.d.**

Dieser Satz hat einige wichtige Folgerungen:

Korollar 9.25. **(i)** *Kompakte Mengen eines metrischen Raums sind abgeschlossen.*

(ii) *Abgeschlossene Teilmengen einer kompakten Menge sind wieder kompakt.*

Beweis:(i) Kompakte Mengen sind wegen (iii) des vorangehenden Satzes vollständig, und stimmen wegen Lemma 9.19 mit ihrem Abschluss überein.
(ii) Abgeschlossene Teilmengen einer kompakten Menge erfüllen wegen Lemma 9.19 wieder die Bedingung (ii) des vorangehenden Satzes. **q.e.d.**

Definition 9.26. *Eine Teilmenge A eines metrischen Raumes (X, d) heißt beschränkt, wenn für ein $x \in X$, die Menge der Abstände $\{d(x, y)|y \in A\}$ beschränkt ist.*

Wegen der Dreiecksungleichung ist diese Bedingung äquivalent dazu, dass für alle $x \in X$ die Mengen $\{d(x, y) \mid y \in A\}$ beschränkt sind, aber nicht uniform in $x \in X$.

Der zweite Teil des Beweises vom Satz 5.9 zeigt dass alle kompakten Teilmengen eines metrischen Raumes beschränkt sind. Im $\mathbb{K}^n$ konvergiert eine Folge genau dann, wenn für alle $i = 1, \ldots, n$ die jeweilige Folge der i-ten Komponenten konvergieren. Deshalb überträgt sich auch der erste Teil des Beweises des Satzes 5.9 und zeigt

Satz 9.27. *In jedem metrischen Raum ist eine kompakte Teilmenge beschränkt.*

In $\mathbb{K}^n$ ist eine Teilmenge bezüglich einer der äquivalenten Normen $\|\cdot\|_p$ mit $p \in [1, \infty]$ genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.

Beispiel 9.28. **(i)** *Die Intervalle $[a, b]$ sind kompakt.*
(ii) *$\bar{\mathbb{N}}$ aus Beispiel (vi) ist kompakt.*
(iii) *Sei $(a_n)_{n \in \mathbb{N}}$ konvergent mit Grenzwert a. Dann ist $\{a\} \cup \{a_n | n \in \mathbb{N}\}$ kompakt.*

9.3 Stetigkeit

Definition 9.29. *Eine Abbildung $f : X \to Y, \quad x \mapsto f(x)$ von einem metrischen Raum (X, d) in den metrischen Raum (Y, d) heißt stetig in $x \in X$, wenn es für jedes $\epsilon > 0$ ein $\delta > 0$ gibt, so dass alle $y \in B(x, \delta) \subset X$ auch $f(y) \in B(f(x), \epsilon) \subset Y$ erfüllen. Die Abbildung f heißt stetig, wenn sie in allen Punkten von X stetig ist.*

Stetig im Punkt x heißt also, dass alle Punkte, die hinreichend nahe bei x liegen, auf Werte abgebildet werden, die beliebig nahe bei $f(x)$ liegen.

Satz 9.30. *Für eine Abbildung $f : X \to Y, \quad x \mapsto f(x)$ zwischen den metrischen Räumen (X, d) und (Y, d) ist folgendes äquivalent:*

(i) *f ist stetig in x.*

(ii) *Das Urbild jeder Umgebung von $f(x)$ ist eine Umgebung von x.*

(iii) *Für jede Folge $(x_n)_{n\in\mathbb{N}}$ in (X,d), die gegen x konvergiert, konvergiert auch die Folge $(f(x_n))_{n\in\mathbb{N}}$ gegen $f(x)$.*

Beweis:(i)$\Leftrightarrow$(ii) Die Umgebungen von x sind gerade die Mengen, die einen δ-Ball um x enthalten. Also ist (ii) äquivalent zu der Aussage, dass das Urbild jedes ϵ-Balles um $f(x)$ einen δ-Ball um x enthält. Diese Aussage ist nur eine Umformulierung von (i).
(ii)$\Leftrightarrow$(iii) Die Folgen $(x_n)_{n\in\mathbb{N}}$ und $(f(x_n))_{n\in\mathbb{N}}$ konvergieren genau dann gegen x bzw. $f(x)$, wenn jede Umgebung von x bzw. $f(x)$ alle bis auf endlich viele Folgenglieder enthält. Wenn also $(x_n)_{n\in\mathbb{N}}$ gegen x konvergiert und f (ii) erfüllt, dann konvergiert auch $(f(x_n))_{n\in\mathbb{N}}$ gegen $f(x)$. Also folgt aus (ii) auch (iii). Wenn es umgekehrt einen ϵ-Ball von $f(x)$ gibt, dessen Urbild keinen δ-Ball von x enthält, dann gibt es eine Folge $(x_n)_{n\in\mathbb{N}}$ von Punkten $x_n \in B\left(x, \frac{1}{n}\right)$, so dass die Folge der entsprechenden Werte $f(x_n)$ im Komplement dieses ϵ-Balles von $f(x)$ liegt: $f(x_n) \notin B(f(x), \epsilon)$. Also konvergiert $(x_n)_{n\in\mathbb{N}}$ gegen x aber $f(x_n)$ nicht gegen $f(x)$. **q.e.d.**

Korollar 9.31. *Für eine Abbildung $f : X \to Y$ zwischen den metrischen Räumen (X,d) und (Y,d) ist folgendes äquivalent:*

(i) *f ist stetig.*

(ii) *Das Bild $(f(x_n))_{n\in\mathbb{N}}$ jeder konvergenten Folge $(x_n)_{n\in\mathbb{N}}$ ist konvergent und es gilt*

$$\lim_{n\to\infty} f(x_n) = f\left(\lim_{n\to\infty} x_n\right).$$

(iii) *Das Urbild jeder offenen Menge ist offen.*

(iv) *Das Urbild jeder abgeschlossenen Menge ist abgeschlossen.*

Beweis: Wegen dem vorangehenden Satz sind (i) und (ii) äquivalent. Weil eine Menge genau dann offen ist, wenn sie eine Umgebung von allen ihren Punkten ist, zeigt der vorangehende Satz, dass aus (i) bzw. (ii) auch (iii) folgt. Weil jede Umgebung eines Punktes auch eine offene Umgebung des Punktes enthält, folgt wieder wegen dem vorangehenden Satz aus (iii) auch (i) bzw. (ii). Weil nun die abgeschlossenen Mengen gerade die Komplemente der offenen Mengen sind und das Urbild eines Komplementes gerade gleich dem Komplement des Urbildes ist, ist (iii) zu (iv) äquivalent. **q.e.d.**

Korollar 9.32. *Die Komposition zweier stetiger Abbildungen ist stetig. Die analoge punktweise Aussage gilt auch.*

Beweis: Für die Verkettung von $g \circ f$ von stetigen Funktionen gilt wegen (ii) im Korollar 9.31 bzw. (iii) in Satz 9.30 für eine konvergente Folge $(x_n)_{n\in\mathbb{N}}$ die Gleichung

$$\lim_{n\to\infty} g(f(x_n) = g\left(\lim_{n\to\infty} f(x_n)\right) = g\left(f\left(\lim_{n\to\infty} x_n\right)\right).$$

Beispiel 9.33. **(i)** *Auf jedem metrischen Raum ist die identische Abbildung $\mathbb{1}_X$ stetig.*

(ii) *Die konstante Abbildung, die alle $x \in X$ auf einen Punkt y abbildet ist stetig.*

(iii) *Wegen der Dreiecksungleichung gilt*

$$d(x,y) \le d(x,u) + d(u,y) \le d(x,u) + d(u,v) + d(v,y).$$

Also gilt auch $d(x,y) - d(u,v) \le d(x,u) + d(v,y)$. Durch vertauschen $(x,y) \leftrightarrow (u,v)$ und unter Benutzung der Symmetrie erhalten wir

$$d(u,v) - d(x,y) \le d(x,u) + d(v,y) \Rightarrow |d(x,y) - d(u,v)| \le d(x,u) + d(v,y).$$

Mit der Metrik aus dem Beispiel 9.2 (v) auf $X \times X$ ist $d : X \times X \to \mathbb{R}$ also stetig.

(iv) *Auf einem normierten Vektorraum V folgt $\|v\| - \|w\| \le \|v - w\|$ für $v, w \in V$ aus $\|v\| \le \|v-w\| + \|w\|$. Das Vertauschen von v und w ergibt $|\|v\| - \|w\|| \le \|v-w\|$. Also ist $\|\cdot\|$ auf jedem normierten Vektorraum eine stetige reelle Funktion.*

(v) *Wegen der Dreiecksungleichung sind für jeden normierten Vektorraum V*

$$+ : V \times V \to V, \qquad (v,w) \mapsto v + w \qquad \cdot : \mathbb{K} \times V \to \mathbb{K}, \qquad (\lambda, v) \mapsto \lambda \cdot v$$

stetige Abbildungen. Das gilt auch für die Abbildungen

$$- : \mathbb{K} \to \mathbb{K}, \qquad x \mapsto -x \qquad \textit{und} \qquad {}^{-1} : \mathbb{K} \setminus \{0\} \to \mathbb{K} \setminus \{0\}, \qquad x \mapsto x^{-1}.$$

(vi) *Eine Folge $(x_n)_{n\in\mathbb{N}}$ in einem metrischen Raum (X,d) läßt sich genau dann zu einer stetigen Abbildung von $\bar{\mathbb{N}}$ nach X fortsetzen, wenn sie konvergiert. Dann wird ∞ auf $\lim\limits_{n\to\infty} x_n$ abgebildet.*

Korollar 9.34. *Das Bild einer kompakten Menge unter einer stetigen Abbildung ist kompakt.*

Beweis: Sei $f : X \to Y, \quad x \mapsto f(x)$ eine stetige Abbildung und $A \subset X$ eine kompakte Menge. Dann ist das Urbild einer beliebig offenen Überdeckung von dem Bild

$$f[A] = \{y \in Y \mid \exists x \in A \text{ mit } f(x) = y\}$$

eine offene Überdeckung von A. Diese besitzt, wenn A kompakt ist, eine endliche Teilüberdeckung. Die entsprechende endliche Teilmenge der ursprünglichen offenen Überdeckung von $f[A]$ überdeckt auch $f[A]$. Also ist $f[A]$ kompakt. **q.e.d.**

Korollar 9.35. *Das Bild einer kompakten Menge unter einer stetigen Abbildung ist beschränkt. Das Bild eine solchen reellen Funktion besitzt Minimum und Maximum.*

Beweis: Sei (X, d) kompakt und $f : X \to Y, \quad x \mapsto f(x)$ stetig. Wegen Korollar 9.34 ist $f[X]$ kompakt und wegen Satz 9.27 beschränkt. Wegen Korollar 5.11 besitzen die kompakten Teilmengen von $\mathbb{R}$ sowohl ein Minimum als auch ein Maximum. **q.e.d.**

Eine bijektive stetige Abbildung $f : X \to Y$ zwischen metrischen Räumen (X, d) und (Y, d) heißt homöomorph oder Homöomorphismus, wenn f^{-1} stetig ist.

Korollar 9.36. *Auf kompaktem X ist jedes stetige bijektive $f : X \to Y$ homöomorph.*

Beweis: Wegen Korollar 9.34 ist das Bild $f[X] = Y$ kompakt. Wegen Korollar 9.25 ist eine Teilmenge eines kompakten metrischen Raumes genau dann abgeschlossen, wenn sie kompakt ist. Weil das Urbild unter der Umkehrabbildung gleich dem Bild unter f ist, folgt die Aussage aus Korollar 9.34 und Korollar 9.31 (iv). **q.e.d.**

Satz 9.37. *Auf $\mathbb{K}^n$ sind alle Normen paarweise äquivalent.*

Beweis: Es genügt zu zeigen, dass alle Normen äquivalent sind zu $\|\cdot\|_1$. Sei also $\|\cdot\|$ eine beliebige Norm. Sei $e_1, \ldots, e_n$ die Basis von $\mathbb{K}^n$, deren i–tes Element nur an der i–ten Stelle eine nicht verschwindende Komponente hat, die dann jeweils gleich Eins ist. Wegen der Dreiecksungleichung gilt dann für jedes $v = (v_1, \ldots, v_n) \in \mathbb{K}^n$

$$\|v\| \leq |v_1| \cdot \|e_1\| + \ldots + |v_n| \cdot \|e_n\| \leq \max\{\|e_1\|, \ldots, \|e_n\|\}\|v\|_1.$$

Aus der Dreiecksungleichung folgt dann für alle $v, w \in \mathbb{K}^n$

$$|\|v\| - \|w\|| \leq \|v - w\| \leq \max\{\|e_1\|, \ldots, \|e_n\|\}\|v - w\|_1.$$

Also ist die reelle Funktion $v \mapsto \|v\|$ stetig bezüglich der Norm $\|\cdot\|_1$, und wegen dem Satz 9.27 von Heine-Borel ist die Teilmenge $\{v \in \mathbb{K}^n \mid \|v\|_1 = 1\}$ mit der von $\|\cdot\|_1$ induzierten Metrik kompakt. Wegen Korollar 9.35 nimmt diese Funktion auf dieser Menge das Minimum C an. Wegen der Positivität von $\|\cdot\|$ gilt $C > 0$. Daraus folgt

$$C\|v\|_1 \leq \left\|\frac{v}{\|v\|_1}\right\| \|v\|_1 = \|v\| \leq \max\{\|e_1\|, \ldots, \|e_n\|\}\|v\|_1 \text{ für alle } v \in \mathbb{K}^n \setminus \{0\}.$$ **q.e.d.**

Definition 9.38. *(Gleichmäßige Stetigkeit, Lipschitzstetigkeit) Eine Abbildung $f : X \to Y$ zwischen metrischen Räumen heißt gleichmäßig stetig, wenn es für alle $\epsilon > 0$ ein $\delta > 0$ gibt, so dass $d(f(x), f(y)) < \epsilon$ für alle $x, y \in X$ mit $d(x, y) < \delta$ gilt.*

Die Abbildung heißt lipschitzstetig auf A, wenn es eine Konstante $L > 0$ (Lipschitzkonstante) gibt, so dass für alle $x, y \in A$ gilt $d(f(x), f(y)) \leq Ld(x, y)$.

Offenbar ist jede lipschitzstetige Abbildung auch gleichmäßig stetig und jede gleichmäßig stetige Abbildung auch stetig. Es gilt auch folgende Umkehrung:

Satz 9.39. *Sei $f : X \to Y, \quad x \mapsto f(x)$ eine stetige Abbildung zwischen metrischen Räumen. Dann ist f auf jeder kompakten Menge A auch gleichmäßig stetig.*

Auf kompakten Mengen sind also gleichmäßige und einfache Stetigkeit äquivalent.

Beweis: Sei also $f : X \to Y, \quad x \mapsto f(x)$ stetig und $A \subset X$ kompakt. Dann gibt es für jedes $\epsilon > 0$ und jedes $x \in A$ ein $\delta(x)$, so dass $f(y) \in B(f(x), \frac{\epsilon}{2})$ aus $y \in B(x, 2\delta(x))$ folgt. Wir wählen eine endliche Teilüberdeckung von der offenen Überdeckung $\{B(x, \delta(x)) \mid x \in A\}$ von A. Sei δ das Minimum der Radien dieser endlichen Teilüberdeckung. Dann gibt es für alle $y, z \in A$ mit $d(y, z) < \delta$ einen Ball $B(x, \delta(x))$ der endlichen Teilüberdeckung mit $y \in B(x, \delta(x))$. Dann folgt

$$d(x, z) \leq d(x, y) + d(y, z) < \delta(x) + \delta \leq 2\delta(x) \quad \text{also } z \in B(x, 2\delta(x)).$$

Daraus folgt $$d(f(y), f(z)) \leq d(f(x), f(y)) + d(f(x), f(z)) < \epsilon.$$ **q.e.d.**

Übungsaufgabe 9.40. *In dieser Aufgabe konstruieren wir $\mathbb{R}$ als die Vervollständigung von dem angeordneten Körper $\mathbb{Q}$. Sei dazu $\mathfrak{C}$ die Menge der Cauchyfolgen in $\mathbb{Q}$.*

Für zwei Cauchyfolgen $(x_n)_{n\in\mathbb{N}}, (\widetilde{x}_n)_{n\in\mathbb{N}} \in \mathfrak{C}$ definieren wir die Relation

$$(x_n) \sim (\widetilde{x}_n)_{n\in\mathbb{N}} \qquad \Longleftrightarrow \qquad (x_n - \tilde{x}_n)_{n\in\mathbb{N}} \text{ ist Nullfolge in } \mathbb{Q}.$$

(i) *Zeige, dass $\sim$ eine Äquivalenzrelation auf $\mathfrak{C}$ ist.*

(ii) *Wir bezeichnen die Menge der Äquivalenzklassen $\mathfrak{C}/\sim$ suggestiv mit $\mathbb{R}$. Zeige, dass die folgendermaßen definierte Addition und Multiplikation*

$$[(x_n)_{n\in\mathbb{N}}] + [(y_n)_{n\in\mathbb{N}}] := [(x_n + y_n)_{n\in\mathbb{N}}], \quad [(x_n)_{n\in\mathbb{N}}] \cdot [(y_n)_{n\in\mathbb{N}}] := [(x_n \cdot y_n)_{n\in\mathbb{N}}]$$

*jeweils wohldefiniert sind, d.h. dass erstens die Addition und die Multiplikation von zwei Cauchyfolgen selbst eine Cauchyfolge ist und zweitens, dass diese Verknüpfungen nicht von den jeweiligen Repräsentanten (x_n) und (y_n) der Äquivalenzklassen abhängen. (*Tipp: *Benutze: Cauchyfolgen sind beschränkt)*

(iii) *Zeige, dass $\mathbb{R} = \mathfrak{C}/\sim$ mit diesen Verknüpfungen die Körperaxiome A1–A3 erfüllt. (*Tipp: *Benutze, dass $\mathbb{Q}$ die Axiome A1–A3 erfüllt.)*

(iv) *Zeige, dass folgende Relation auf $\mathbb{R}$ eine wohldefinierte Ordnungsrelation ist, die das Axiom A4 erfüllt:*

$$[(x_n)_{n\in\mathbb{N}}] > [(y_n)_{n\in\mathbb{N}}] \qquad \Longleftrightarrow \qquad \exists N \in \mathbb{N} : x_n - y_n \geq \tfrac{1}{N} \ \forall n \geq N.$$

(v) *Definiere die Einbettungsabbildung,*

$$\Phi : \mathbb{Q} \hookrightarrow \mathbb{R}, \qquad\qquad q \mapsto [q],$$

wobei für jedes $q \in \mathbb{Q}$, $[q]$ die konstante Folge $q_n = q$ bezeichnet. Diese Abbildung ist offenbar injektiv. Zeige, dass $\mathbb{R}$ archimedisch ist. Hierbei kann ohne Beweis benutzt werden, dass die natürlichen Zahlen in $\mathbb{R}$ das Bild von $\mathbb{N} \subset \mathbb{Q}$ unter Φ sind. Zeige in einem zweiten Schritt, dass das Bild $\Phi[\mathbb{Q}]$ dicht in $\mathbb{R}$ liegt.

(vi) *Zeige, dass* $\mathbb{R}$ *vollständig ist. (*Tipp: *Nimm eine Cauchyfolge* $(\xi_n)_{n\in\mathbb{N}}$ *in* $\mathbb{R}$ *als gegeben an. Da das Bild* $\Phi[\mathbb{Q}]$ *dicht in* $\mathbb{R}$ *liegt, existiert für jedes* $n \in \mathbb{N}$ *jeweils ein* $x_n \in \mathbb{Q}$ *mit* $\Phi(-\frac{1}{n}) < \Phi(x_n) - \xi_n < \Phi(\frac{1}{n})$*. Zeige, dass* $(x_n)_{n\in\mathbb{N}}$ *eine Cauchyfolge in* $\mathbb{Q}$ *ist, und* $(\xi_n)_{n\in\mathbb{N}}$ *in* $\mathbb{R}$ *gegen das entsprechende* $\xi := [(x_n)_{n\in\mathbb{N}}]$ *konvergiert.)*

9.4 Funktionenräume

In diesem Abschnitt sei (Y, d) ein metrischer Raum, ein normierter Vektorraum oder eine normierte Algebra und später auch (X, d) ein metrischer Raum:

Definition 9.41. *Eine normierte Algebra ist ein normierter Vektorraum* V *mit einer assoziativen und distributiven Multiplikation* $\cdot : V \times V \to V$*, die folgendes erfüllt:*

$$\begin{aligned} &(v + v'') \cdot v' = v \cdot v' + v'' \cdot v && (\lambda v) \cdot v' = \lambda(v \cdot v') && \|v \cdot v'\| \le \|v\| \cdot \|v'\| \\ &v \cdot (v' + v'') = v \cdot v' + v \cdot v'' && v \cdot (\lambda v') = \lambda(v \cdot v') && \text{für alle } v, v', v'' \in V, \lambda \in \mathbb{K}. \end{aligned}$$

Wenn V *vollständig ist, heißt* V *Banachalgebra.*

Wir betrachten in diesem Abschnitt Mengen von Abbildungen von X nach Y. Wenn Y ein normierter Vektorraum ist, können wir solche Abbildungen punktweise miteinander addieren und mit Elementen von $\mathbb{K}$ multiplizieren, und wenn Y eine Algbera ist, auch punktweise miteinander multiplizieren:

$$\begin{aligned} f + g &: X \to Y, & x &\mapsto f(x) + g(x), & \lambda f &: X \to Y, & x &\mapsto \lambda f(x) \\ f \cdot g &: X \to Y, & x &\mapsto f(x) \cdot g(x). \end{aligned}$$

Die Addition erfüllt die Axiome $A1$ und mit der Skalarmultiplikation das Distributivgesetz. Dadurch wird die Menge aller Abbildungen in einen Vektorraum Y zu einem Vektorraum, und zu einer Algebra, wenn Y eine Algebra ist. Das Inverse einer Funktion f in einer Algebra mit Eins $\mathbb{1} \in Y$ existiert nur, wenn $f(x)$ für alle $x \in X$ invertierbar ist. Indem wir die Elemente von $\mathbb{K}$ mit den entsprechenden Vielfachen der Eins identifizieren, wird die Skalarmultiplikation zu einem Spezialfall der Multiplikation.

Definition 9.42. *Eine Folge von Funktionen* $(f_n)_{n\in\mathbb{N}}$ *von* X *nach* Y *heißt*

punktweise konvergent, *wenn die Folgen* $(f_n(x))_{n\in\mathbb{N}}$ *für jedes* $x \in X$ *konvergieren. Die Grenzwerte definieren wieder eine Funktion* $f : X \to Y, \quad x \mapsto \lim_{n\to\infty} f_n(x)$.

gleichmäßig konvergent, *wenn es eine Funktion* $f : X \to Y, \quad x \mapsto f(x)$ *gibt, und für alle* $\epsilon > 0$ *ein* $N \in \mathbb{N}$*, so dass* $d(f_n(x), f(x)) < \epsilon$ *für* $n \ge N$ *und* $x \in X$ *gilt.*

Offenbar ist jede gleichmäßige konvergente Folge (f_n) auch punktweise konvergent, aber nicht umgekehrt (siehe Beispiel 5.24).

Definition 9.43. *Eine Abbildung $f : X \to Y$ in einen metrischen Raum Y heißt beschränkt, wenn das Bild $f[X]$ in Y beschränkt ist. $B(X,Y)$, bezeichne die Menge aller beschränkten Abbildungen von X nach Y. Auf $B(X,Y)$ bezeichne d folgende Abbildung:*

$$d : B(X,Y) \times B(X,Y) \to \mathbb{R}, \quad (f,g) \mapsto d(f,g) = \sup\{d(f(x),g(x)) \mid x \in X\}.$$

Wenn Y ein normierter Vektorraum ist, dann bezeichne $\|\cdot\|_\infty$ folgende Abbildung:

$$\|\cdot\|_\infty : B(X,Y) \to \mathbb{R}, \quad f \mapsto \|f\|_\infty = \sup\{\|f(x)\| \mid x \in X\}.$$

Satz 9.44. **(i)** *Für einen metrischen Raum Y ist d eine Metrik auf $B(X,Y)$.*

(ii) *Wenn Y ein normierter Vektorraum (Algebra) ist, ist $B(X,Y)$ ein normierter Vektorraum (Algebra) mit Norm $\|\cdot\|_\infty$, die die Metrik aus (i) induziert.*

(iii) *Wenn Y ein vollständiger metrischer Raum ist, dann auch $(B(X,Y),d)$.*

Beweis:(i) und **(ii)** folgen aus den Eigenschaften der Metrik bzw. Norm $\|\cdot\|$ von Y, und weil wegen der Dreiecksungleichung und wegen $\|v \cdot w\| \leq \|v\| \cdot \|w\|$ die Summe und das Produkt zweier beschränkter Abbildungen wieder beschränkt ist.
(iii) Sei $(f_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in $B(X,Y)$. Für alle $\epsilon > 0$ gibt es ein $N \in \mathbb{N}$ mit

$$d(f_n(x), f_m(x)) \leq d(f_n, f_m) < \frac{\epsilon}{2} \quad \text{für alle } n,m \geq N \text{ und alle } x \in X.$$

Dann sind für alle $x \in X$ die Folgen $(f_n(x))_{n\in\mathbb{N}}$ Cauchyfolgen. Also konvergieren sie punktweise gegen eine Funktion $f : X \to Y, \quad x \mapsto f(x)$. Für obiges ϵ und jedes $x \in X$ gibt es ein $M(x) \in \mathbb{N}$, so dass $d(f_m(x), f(x)) < \frac{\epsilon}{2}$ für alle $m \geq M(x)$ gilt. Damit folgt

$$d(f_n(x), f(x)) \leq d(f_n(x), f_{\max\{N,M(x)\}}(x)) + d(f_{\max\{N,M(x)\}}(x), f(x)) < \epsilon$$

für alle $x \in X$ und $n \geq N$. Also konvergiert $(f_n)_{n\in\mathbb{N}}$ gleichmäßig gegen f. Aus $\sup\{d(f_N(x), f(x)) \mid x \in X\} \leq \epsilon$ und $f_N \in B(X,Y)$ folgt $f \in B(X,Y)$. **q.e.d.**

Definition 9.45. *Für metrische Räume X und Y sei $C_b(X,Y)$ sei der Unterraum von $B(X,Y)$ aller stetigen und beschränkten Funktionen von X nach Y.*

Satz 9.46. **(i)** *Für metrische Räume X und Y ist $C_b(X,Y)$ abgeschlossen in $B(X,Y)$.*

(ii) *Wenn Y ein normierter Vektorraum (Algebra) ist, dann auch $C_b(X,Y)$.*

(iii) *Wenn Y vollständig ist, dann auch $C_b(X,Y)$.*

Beweis: Wegen dem vorangehenden Satz und Lemma 9.19 genügt es zu zeigen, dass für jede Folge in $C_b(X,Y)$, die als Folge in $B(X,Y)$ konvergiert, der Grenzwert in $C_b(X,Y)$ liegt. Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge in $C_b(X,Y))$, die in $B(X,Y)$ gegen f konvergiert. Dann gibt es für jedes $\epsilon > 0$ ein $n \in \mathbb{N}$ mit $d(f_n, f) < \frac{\epsilon}{3}$. Weil f_n stetig bei $x \in X$ ist gibt es ein $\delta > 0$, so dass $d(f_n(x), f_n(y)) < \frac{\epsilon}{3}$ für alle $y \in B(x, \delta)$ gilt. Dann folgt

$$d(f(x), f(y)) \leq d(f(x), f_n(x)) + d(f_n(x), f_n(y)) + d(f_n(y), f(y)) < \epsilon.$$

Also ist f bei $x \in X$ stetig. **q.e.d.**

Die gleichmäßige Konvergenz der Folge $(f_n)_{n\in\mathbb{N}}$ ist notwendig (siehe Beispiel 5.24). Wenn Y ein Banachraum ist, dann sind sowohl $B(X, Y)$ als auch $C_b(X, Y)$ Banachräume. Wenn Y eine Banachalgebra ist wie z.B. $\mathbb{K}$, dann sind auch $B(X, Y)$ und $C_b(X, Y)$ Banachalgebren. Der Fall $Y = \mathbb{K}$ wird im folgenden noch öfter vorkommen. Jetzt können wir die Vervollständigungen aller metrischen Räume leicht konstruieren:

Satz 9.47. *Sei X ein metrischer Raum und $x_0 \in X$. Für alle $x \in X$ gehört dann*

$$I(x) : X \to \mathbb{R} \qquad y \mapsto d(x, y) - d(x_0, y) \quad \textit{zu } C_b(X, \mathbb{R}).$$

$$\textit{Die Abbildung} \quad I : X \to C_b(X, \mathbb{R}) \qquad x \mapsto I(x)$$

ist eine isometrische Abbildung von X nach $C_b(X, \mathbb{R})$, d.h. es gilt $d(I(x), I(y)) = d(x, y)$ für alle $x, y \in X$. Für jede gleichmäßig stetige Abbildung f von X in einen vollständigen metrischen Raum Y, gibt es eine stetige Abbildung $g : \overline{I[X]} \to Y$ auf dem Abschluss des Bildes von I in $C_b(X, \mathbb{R})$, so dass f gleich $g \circ I$ ist (vergleiche Übungsaufgabe 9.40).

Beweis: Wegen Beispiel 9.33 (iii) sind die reellen Funktionen $I(x)$ für alle $x \in X$ stetig. Wegen der Dreiecksungleichung gilt für $x, y, z \in X$

$$\begin{aligned}|d(x, z) - d(y, z)| &= \max\{d(x, z) - d(y, z), d(y, z) - d(x, z)\} \leq \\ &\leq \max\{d(x, y) + d(y, z) - d(y, z), d(y, x) + d(x, z) - d(x, z)\} = d(x, y) \quad \text{und}\end{aligned}$$
$$d(x, y) = d(x, y) - d(y, y) \leq d(I(x), I(y)) = \sup\{|d(x, z) - d(y, z)| \mid z \in X\} \leq d(x, y).$$

Mit $I(x_0) = 0$ und $y = x_0$ folgt $\|I(x)\|_\infty = d(x, x_0)$ und $I(x) \in C_b(X, \mathbb{R})$. Also ist I eine isometrische Abbildung. Sei $h \in \overline{I[X]} \subset C_b(X, \mathbb{R})$ und $f : X \to Y$ eine gleichmäßig stetige Abbildung in einen vollständigen metrischen Raum Y. Für jedes $\epsilon > 0$ gibt es ein $\delta > 0$, so dass $d(f(x), f(y)) < \epsilon$ aus $d(x, y) < \delta$ folgt. Also haben für $u \in \overline{I[X]}$ alle Elemente von $\{f(x) \in Y \mid x \in X \text{ mit } d(I(x), u) < \frac{\delta}{2}\}$ paarweise einen Abstand kleiner als ϵ. Deshalb sind $(f(x_n))_{n\in\mathbb{N}}$ für alle Folgen $(x_n)_{n\in\mathbb{N}}$ in X, für die $(I(x_n))_{n\in\mathbb{N}}$ gegen ein u konvergiert, Cauchyfolgen in Y, die alle gegen das gleiche Element von Y konvergieren. Wir definieren $g(u)$ als $\lim\limits_{n\to\infty} f(x_n)$. Wenn $(I(x_n))_{n\in\mathbb{N}}$ und $(I(y_n))_{n\in\mathbb{N}}$ in $\overline{I[X]}$ gegen u bzw. v konvergieren mit $d(u, v) < \delta$, dann ist $\lim\limits_{n\to\infty} d(x_n, y_n) = \lim\limits_{n\to\infty} d(I(x_n), I(y_n)) < \delta$, also $d(g(u), g(v)) = \lim\limits_{n\to\infty} d(f(x_n), f(y_n)) \leq \epsilon$, und g stetig. **q.e.d.**

Satz 9.48. *(Satz von Stone–Weierstraß) Sei (X, d) ein kompakter metrischer Raum und $A \subset C_b(X, \mathbb{R})$ eine Unteralgebra, die $1 \in C_b(X, \mathbb{R})$ enthält und die Punkte trennt, d.h. für alle $x \neq y \in X$ gibt es $f \in A$ mit $f(x) \neq f(y)$. Dann ist $\bar{A} = C_b(X, \mathbb{R})$.*

Lemma 9.49. *Auf* $[0,1] \subset \mathbb{R}$ *konvergiert die induktiv definierte Folge von Polynomen* $p_{n+1}(x) = p_n(x) + \frac{1}{2}(x - p_n^2(x))$ *mit* $p_0 = 0$, *gleichmäßig gegen die Funktion* $x \mapsto \sqrt{x}$.

Beweis : Wir zeigen zunächst mit vollständiger Induktion, dass $0 \le p_n(x) \le p_{n+1}(x)$ und $0 \le p_n^2(x) \le x$ für $x \in [0,1]$ und alle $n \in \mathbb{N}_0$ gilt. Beides ist für $n = 0$ offensichtlich.

$$\begin{aligned} x - p_{n+1}^2(x) &= x - p_n^2(x) - p_n(x)\left(x - p_n^2(x)\right) - \frac{1}{4}\left(x - p_n^2(x)\right)^2 \\ &= (x - p_n^2(x))\left(1 - p_n(x) - \frac{1}{4}(x - p_n^2(x))\right) \\ &= (x - p_n^2(x))\left(\left(1 - \frac{p_n(x)}{2}\right)^2 - \frac{x}{4}\right) \end{aligned}$$

Aus der Induktionsvoraussetzung folgt $0 \le p_n(x) \le 1$, also $\frac{1}{2} \le 1 - \frac{p_n(x)}{2} \le 1$ und $(1 - \frac{p_n(x)}{2})^2 \ge \frac{1}{4} \ge \frac{x}{4}$. Das zeigt die Induktion. Also ist $(p_n(x))_{n\in\mathbb{N}_0}$ für $x \in [0,1]$ monoton wachsend mit $0 \le p_n^2(x) \le x$. Dann gilt $(1 - \frac{p_n(x)}{2})^2 - \frac{x}{4} \le 1 - \frac{x}{4}$ und deshalb $0 \le x - p_n^2(x) \le x \cdot \left(1 - \frac{x}{4}\right)^n$. Wegen $\frac{1}{1-\frac{x}{4}} = \sum_{n=0}^{\infty}(\frac{x}{4})^n \ge 1 + \frac{x}{4}$ folgt dann aus der Bernoulli Ungleichung $\frac{1}{(1-\frac{x}{4})^n} \ge 1 + \frac{nx}{4}$ und $0 \le x - p_n^2(x) \le \frac{x}{1+\frac{nx}{4}} < \frac{4}{n}$. Also konvergiert $(p_n^2(x))_{n\in\mathbb{N}}$ auf $x \in [0,1]$ gleichmäßig gegen x. Die Funktion $[0,1] \to [0,1], \quad x \mapsto \sqrt{x}$ ist die Umkehrfunktion von $[0,1] \to [0,1], \quad x \mapsto x^2$. Weil die zweite Funktion stetig ist, ist wegen Korollar 5.18 die erste stetig und wegen Satz 5.22 sogar gleichmäßig stetig. Dann konvergiert die Folge $(p_n(x) = \sqrt{p_n^2(x)})_{n\in\mathbb{N}}$ gleichmäßig gegen $\sqrt{x}$. **q.e.d.**

Beweis des Satzes von Stone–Weierstraß: Wegen Lemma 9.49 konvergiert eine Folge $(p_n)_{n\in\mathbb{N}}$ von Polynomen auf $[0,1]$ gleichmäßig gegen $x \mapsto \sqrt{x}$. Für jedes $f \in A$ konvergiert dann $(p_n(\frac{f^2}{\|f\|_\infty^2}))_{n\in\mathbb{N}}$ in $C_b(X,\mathbb{R})$ gegen $\sqrt{(\frac{f}{\|f\|_\infty})^2}$. Also gehört $|f| = \|f\|_\infty\sqrt{(\frac{f}{\|f\|_\infty})^2}$ zum Abschluss $\bar{A}$ von A. Aus Korollar 2.20 folgt mit $\||f| - |g|\|_\infty \le \|f - g\|_\infty$ für alle $f, g \in C_b(X,\mathbb{R})$ die Stetigkeit von $f \mapsto |f|$. Wegen der Stetigkeit von $+$ und $\cdot$ auf $C_b(X,\mathbb{R})$ ist $\bar{A}$ eine Algebra mit $|f| \in \bar{A}$ für $f \in \bar{A}$. Für $f, g \in \bar{A}$ gehören

$$\max(f,g) = \frac{1}{2}(f + g + |f-g|) \quad \text{und} \quad \min(f,g) = \frac{1}{2}(f + g - |f-g|)$$

zu $\bar{A}$. Weil A die Punkte von X trennt, gibt es für alle $x \ne y \in X$ und alle $\alpha, \beta \in \mathbb{R}$ ein Element $f \in A$ mit $f(x) = \alpha$ und $f(y) = \beta$. Sei nämlich g eine Funktion mit $g(x) \ne g(y)$. Dann ist $f = \alpha + \frac{\beta-\alpha}{g(y)-g(x)}(g - g(x))$ eine solche Funktion.

Sei jetzt $f \in C_b(X,\mathbb{R})$ eine fest vorgegebene Funktion und $\epsilon > 0$. Dann gibt es für alle $x, y \in X$ eine Funktion $g_{x,y} \in \bar{A}$ die bei x und y mit f übereinstimmt. Dann gibt es ein $\delta_{x,y} > 0$, so dass $g_{x,y}(z) < f(z) + \epsilon$ für alle $z \in B(y, \delta_{x,y})$ gilt. Durch Übergang zu einer endlichen Teilüberdeckung von $\{B(y,\delta_{x,y}) \mid y \in X\}$ und dem Minimum der entsprechenden Funktionen $g_{x,y} \in \bar{A}$ gibt es eine Funktion $g_x \in \bar{A}$, die $g_x(x) = f(x)$

und $g_x < f + \epsilon$ erfüllt. Wegen der Stetigkeit von f und g_x gibt es für alle $x \in X$ ein $\delta_x > 0$, so dass $f(y) - \epsilon < g_x(y)$ für alle $y \in B(x, \delta_x)$ gilt. Durch Übergang zu einer endlichen Teilüberdeckung von $\{B(x, \delta_x) \mid x \in X\}$ und dem Maximum der entsprechenden Funktionen g_x finden wir schließlich eine Funktion g in $\bar{A}$, die $f - \epsilon < g < f + \epsilon$ auf X erfüllt. Weil ϵ beliebig ist folgt $f \in \bar{A}$. **q.e.d.**

Für jeden metrischen Raum X besitzt die Algebra $C_b(X, \mathbb{C})$ folgende komplexe Konjugation, die jedes $f \in C_b(X, \mathbb{C})$ auf $\bar{f} \in C_b(X, \mathbb{C}$, mit $\bar{f}(x) = \overline{f(x)}$ abbildet. Diese Abbildung ist ein Algebrahomomorphismus, d.h. sie ist linear und erhält das Produkt.

Korollar 9.50.* *Sei (X, d) ein kompakter metrischer Raum und $A \subset C_b(X, \mathbb{C})$ eine Unteralgebra, die $1 \in C_b(X, \mathbb{C})$ und für jedes $f \in A$ auch die komplex konjugierte Funktion $\bar{f} \in A$ enthält und die Punkte trennt, d.h. für alle $x \neq y \in X$ gibt es $f \in A$ mit $f(x) \neq f(y)$. Dann ist $\bar{A} = C_b(X, \mathbb{C})$.*

Beweis*: Jedes $f \in A$ ist die Summe einer reellen Funktion $\frac{1}{2}(f + \bar{f}) \in A$ und des Produktes von $\imath$ mit einer reellen Funktion $\frac{\imath}{2}(\bar{f} - f) \in A$. Also folgt die Aussage aus dem Satz von Stone-Weierstraß. **q.e.d.**

Satz 9.51.* *(Satz von Dini) Auf einem kompakten metrischen Raum (X, d) konvergiert eine monotone Funktionenfolge $(f_n)_{n \in \mathbb{N}}$ von stetigen reellen Funktionen gleichmäßig, wenn sie punktweise gegen eine stetige Funktion f konvergiert.*

Beweis*: Sei $(f_n)_{n \in \mathbb{N}}$ eine monoton wachsende Folge in $C_b(X, \mathbb{R})$, die punktweise gegen $f \in C_b(X, \mathbb{R})$ konvergiert. Dann gibt es zu jedem $\epsilon > 0$ und $x \in X$ ein $n(x) \in \mathbb{N}$, so dass $f(x) - f_{n(x)}(x) < \frac{\epsilon}{3}$ gilt. Da $f_{n(x)}$ und f stetig sind gibt es ein $\delta(x)$, so dass

$$|f_{n(x)}(x) - f_{n(x)}(y)| < \frac{\epsilon}{3} \quad \text{und} \quad |f(x) - f(y)| < \frac{\epsilon}{3} \quad \text{für alle } y \in B(x, \delta(x)) \text{ gilt.}$$

Dann gilt dort auch $f(y) - f_{n(x)}(y) < \epsilon$. Wähle eine endliche Teilüberdeckung von $\{B(x, \delta(x)) \mid x \in X\}$. Dann gilt für $m \geq$ Maximum der entsprechenden $n(x)$

$$f(y) - f_m(y) \leq f(y) - f_{n(x)}(y) < \epsilon$$

auf den Mengen der Teilüberdeckung. Das zeigt die gleichmäßige Konvergenz. **q.e.d.**

Definition 9.52.* *(relativkompakt) Eine Teilmenge eines metrischen Raumes heißt relativkompakt, wenn der Abschluss kompakt ist.*

Lemma 9.53.* *Eine Teilmenge A eines metrischen Raumes (X, d) ist genau dann relativkompakt, wenn jede Folge in A eine in X konvergente Teilfolge besitzt.*

Beweis*: Wenn A relativkompakt ist, dann besitzt wegen Satz 9.24 jede Folge in A eine konvergente Teilfolge, deren Grenzwert im Abschluss $\bar{A}$ liegt. Hat umgekehrt jede Folge in A eine konvergente Teilfolge, dann gibt es wegen Lemma 9.19 für jede Folge $(x_n)_{n \in \mathbb{N}}$

im Abschluss von A auch eine Folge $(a_n)_{n\in\mathbb{N}}$ in A mit $d(x_n, a_n) < \frac{1}{n}$. Dann konvergiert die jeder konvergenten Teilfolge von $(a_n)_{n\in\mathbb{N}}$ entsprechende Teilfolge von $(x_n)_{n\in\mathbb{N}}$ gegen den gleichen Grenzwert wie die entsprechende Teilfoge von $(a_n)_{n\in\mathbb{N}}$. Wegen Satz 9.24 ist dann der Abschluss von A kompakt. **q.e.d.**

Satz 9.54.* *(Arzela–Ascoli) Sei X ein kompakter und Y ein vollständiger metrischer Raum. Eine Teilmenge $\mathcal{F} \subset C_b(X, Y)$ ist genau dann relativkompakt, wenn*

(i) *für jedes $x \in X$ die Menge $\{f(x) \mid f \in \mathcal{F}\}$ relativkompakt ist, und*

(ii) *für jedes $x \in X$ die Menge $\mathcal{F}$ gleichgradig stetig ist in x, d.h. für jedes $x \in X$ und jedes $\epsilon > 0$ gibt es ein $\delta > 0$ mit $f(x') \in B(f(x), \epsilon)$ für alle $x' \in B(x, \delta), f \in \mathcal{F}$.*

Beweis*: Zunächst zeigen wir, dass wenn die Menge $\mathcal{F}$ die Bedingungen (i)-(ii) erfüllt, jede Folge $(f_n)_{n\in\mathbb{N}}$ in $\mathcal{F}$ eine in $C_b(X, Y)$ konvergente Teilfolge $(g_n)_{n\in\mathbb{N}}$ besitzt. Dafür zeigen wir zuerst, dass $\mathcal{F}$ auf X sogar gleichmäßig gleichgradig stetig ist. Für jedes $\epsilon > 0$ und jedes $y \in X$ gibt es wegen (ii) ein $\delta_y > 0$, so dass $d(f(x), f(y)) < \frac{\epsilon}{2}$ für alle $f \in \mathcal{F}$ aus $d(x, y) < 2\delta_y$ folgt. Wegen der Kompaktheit von X hat die Überdeckung $\{B(y, \delta_y) \mid y \in X\}$ eine endliche Teilüberdeckung $X = B(y_1, \delta_1) \cup \ldots \cup B(y_N, \delta_N)$. Sei δ das Minimum von $\delta_1, \ldots, \delta_N$. Dann enthält für alle Paare $x, x' \in X$ mit $d(x, x') < \delta$ einer der Bälle $B(y_1, \delta_1), \ldots, B(y_N, \delta_N)$ den einen Punkt x. Damit sind beide in einem der Bälle $B(y_1, 2\delta_1), \ldots, B(y_N, 2\delta_N)$ enthalten. Daraus folgt $d(f(x), f(x')) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$ für alle $f \in \mathcal{F}$. Also ist $\mathcal{F}$ auf ganz X gleichmäßig gleichgradig stetig.

Für jedes $n \in \mathbb{N}$ wird X durch endlich viele Bälle mit Radius $\frac{1}{n}$ überdeckt. Sei $(x_l)_{l\in\mathbb{N}}$ eine Abzählung der Zentren aller dieser Bälle, die als Folge in X dicht liegt. Wegen (i) ist $A_l = \overline{\{f_n(x_l) \mid n \in \mathbb{N}\}}$ für alle $l \in \mathbb{N}$ eine kompakte Teilmenge von Y. Wir definieren induktiv eine Teilfolge von $(g_n)_{n\in\mathbb{N}}$ von $(f_n)_{n\in\mathbb{N}}$ und eine Folge $(a_l)_{l\in\mathbb{N}}$ in Y, so dass $d(g_n(x_l), a_l) < \frac{1}{n}$ für alle $l \in \mathbb{N}$ und alle $n \geq l$ gilt. Dafür wählen wir zunächst einen Häufungspunkt a_1 von $(f_n(x_1))_{n\in\mathbb{N}}$ und eine Teilfolge $(g_n)_{n\in\mathbb{N}}$ von $(f_n)_{n\in\mathbb{N}}$, so dass $d(g_n(x_1), a_1) \leq \frac{1}{n}$ für alle $n \in \mathbb{N}$ gilt. Induktiv wählen wir für jedes $L \in \mathbb{N} \setminus \{1\}$ einen Häufungspunkt a_L von $(g_n(x_L))_{n\in\mathbb{N}}$ und ersetzen alle Folgenglieder von $(g_n)_{n\in\mathbb{N}}$ mit Indizes $\geq L$ durch eine Teilfolge von $(g_n)_{n\geq L}$, so dass $d(g_n(x_L), a_L) < \frac{1}{n}$ für alle $n \geq L$ gilt. Dann gilt $d(g_n(x_l), a_l) < \frac{1}{n}$ für alle $l = 1, \ldots, L$ und $n \geq l$.

Für ein $\epsilon > 0$ wählen wir ein $\delta > 0$, so dass $d(g_n(x), g_n(x')) < \frac{\epsilon}{3}$ für alle $n \in \mathbb{N}$ und alle $x, x' \in X$ mit $d(x, x') < \delta$ gilt. Die Überdeckung $(B(x_l, \delta))_{l\in\mathbb{N}}$ von X besitzt eine endliche Teilüberdeckung. Also gibt es ein $M \in \mathbb{N}$, so dass an den Zentren der Bälle der Teilüberdeckung $d(g_m(x_l), g_n(x_l)) < \frac{\epsilon}{3}$ für alle $m, n \geq M$ gilt. Dann folgt

$$d(g_m(x), g_n(x)) \leq d(g_m(x), g_m(x_l)) + d(g_m(x_l), g_n(x_l)) + d(g_n(x_l), g_n(x)) < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon$$

für alle $x \in X$ und alle $m, n \geq M$. Also ist $(g_n)_{n\in\mathbb{N}}$ in $C_b(X, Y)$ eine Cauchyfolge und kovergiert wegen Satz 9.46 (iii) in $C_b(X, Y)$. Wegen Lemma 9.53 ist $\mathcal{F}$ relativkompakt.

Wenn umgekehrt $\mathcal{F}$ relativkompakt ist, dann besitzt wegen Lemma 9.53 mit jeder Folge in $\mathcal{F}$ für jedes $x \in X$ auch die Folge der entsprechenden Funktionswerte eine konvergente Teilfolge. Also erfüllt $\mathcal{F}$ die Bedingung (i).

Außerdem gibt es für jedes $x \in X$ und $\epsilon > 0$ endlich viele $f_1, \ldots, f_k$ im Abschluss von $\mathcal{F}$, so dass $B(f_1, \frac{\epsilon}{3}) \cup \ldots \cup B(f_k, \frac{\epsilon}{3})$ den Abschluss von $\mathcal{F}$ überdeckt. Weil $f_1, \ldots, f_k$ stetig sind, gibt es $\delta_1, \ldots, \delta_k > 0$, so dass $f_i(x') \in B(f_i(x), \frac{\epsilon}{3})$ aus $x' \in B(x, \delta_i)$ für $i = 1, \ldots, k$ folgt. Für alle $x' \in B(x, \min\{\delta_1, \ldots, \delta_k\})$ und $f \in \mathcal{F}$ gilt für ein f_i

$$d(f(x'), f(x)) \leq d(f(x'), f_i(x')) + d(f_i(x'), f_i(x)) + d(f_i(x), f(x)) < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3}.$$ **q.e.d.**

9.5 Lineare Operatoren

Die Ableitung einer Funktion von mehreren Veränderlichen ist eine lineare Abbildung. Zur Vorbereitung der Differentialrechnung von Funktionen mehrerer Veränderlicher behandeln wir in diesem Abschnitt solche linearen Abbildungen zwischen normierten Vektorräumen. Dabei betrachten wir wieder Vektorräume über dem Körper $\mathbb{K}$.

Definition 9.55. *Eine Abbildung $A : V \to W$ von einem Vektorraum V in einen Vektorraum W heißt linear, wenn für alle $v, w \in V$ und $\lambda \in \mathbb{K}$ gilt*

$$A(v + w) = Av + Aw \quad \textit{und} \quad A(\lambda v) = \lambda Av.$$

Satz 9.56. *Seien V und W normierte Vektorräume und $A : V \to W$ eine lineare Abbildung. Dann ist folgendes äquivalent:*

(i) *A ist stetig in 0.*

(ii) *A ist stetig.*

(iii) *A ist gleichmäßig stetig.*

(iv) *Es gibt ein $C > 0$, so dass für alle $v \in V$ gilt $\|Av\| \leq C\|v\|$.*

(v) *A ist auf $B(0, 1)$ beschränkt, d.h. $\|Av\| \leq C$ für alle $\|v\| < 1$ mit $0 < C < \infty$.*

Beweis:(i)⇒(v): Wenn A in 0 stetig ist, dann enthält das Urbild jedes Balles $B(0, \epsilon) \subset W$, also auch von $B(0, 1)$, einen Ball $B(0, \delta) \subset V$. Damit folgt $\|Av\| < 1$ aus $\|v\| < \delta$. Aus der Linearität folgt dann $\|Av\| = \frac{1}{\delta}\|A(\delta v)\| < \frac{1}{\delta}$ aus $\|v\| < 1$. Also ist (v) erfüllt.
(v)⇒(iv): Wegen der Linearität folgt aus (v), dass für alle $v \in V \setminus \{0\}$ gilt

$$\|Av\| = \left\| A\left(2\|v\| \cdot \frac{v}{2\|v\|}\right) \right\| = \left\| 2\|v\| \cdot A\left(\frac{v}{2\|v\|}\right) \right\| = 2\|v\| \cdot \left\| A\left(\frac{v}{2\|v\|}\right) \right\| \leq 2\|v\|C.$$

(iv)⇒(iii): Für $v, w \in V$ folgt $\|Av - Aw\| = \|A(v - w)\| \leq C\|v - w\|$ aus (iv). Also ist A sogar lipschitzstetig mit Lipschitzkonstante C. Dann gilt auch (iii).
(iii)⇒(ii) und (ii)⇒(i): Sind offensichtlich. **q.e.d.**

Satz 9.57. *Jede lineare Abbildung A von $\mathbb{K}^n$ in einen normierten Vektorraum ist stetig.*

Beweis: Wir benutzen wieder die Basis $e_1, \dots, e_n$ von $\mathbb{K}^n$. Dann gilt für alle $v \in \mathbb{K}^n$

$$\|Av\| \leq |v_1| \cdot \|Ae_1\| + \ldots + |v_n| \cdot \|Ae_n\| \qquad \leq \|v\|_1 \max\{\|Ae_1\|, \dots, \|Ae_n\|\}.$$

Also folgt die Aussage aus Satz 9.37 und Satz 9.56. **q.e.d.**

Definition 9.58. *Seien V, W normierte Vektorräume. Dann sei $\mathcal{L}(V,W)$ die Menge aller linearen stetigen Abbildungen von V nach W zusammen mit den Abbildungen:*

$$\begin{aligned} + : \quad & \mathcal{L}(V,W) \times \mathcal{L}(V,W) \to \mathcal{L}(V,W), \quad (A,B) \mapsto A + B : \quad V \to W, \quad v \mapsto Av + Bv \\ \cdot : \quad & \mathbb{K} \times \mathcal{L}(V,W) \to \mathcal{L}(V,W), \quad (\lambda, A) \mapsto \lambda A : \quad V \to W, \quad v \mapsto \lambda Av \end{aligned}$$
$$\|\cdot\| : \mathcal{L}(V,W) \to \mathbb{R}, \; A \mapsto \|A\| = \sup\{\|Av\| \mid v \in B(0,1)\} = \sup\{\|Av\| \mid v \in \overline{B(0,1)}\}.$$

Für $w \in \overline{B(0,1)}$ gilt nämlich $\|Aw\| = \lim\limits_{\lambda \uparrow 1} \lambda \|Aw\| = \lim\limits_{\lambda \uparrow 1} \|A \lambda w\| \leq \sup\{\|Av\| \mid v \in B(0,1)\}$.

Satz 9.59. *$\mathcal{L}(V,W)$ ist ein normierter Untervektorraum von $C_b(\overline{B(0,1)}, W)$.*

Beweis: Wegen Satz 9.56 $\mathcal{L}(V,W)$ ist die Menge aller linearen Abbildungen $A : V \to W$, deren Einschränkungen $A|_{\overline{B(0,1)}}$ in $C_b(\overline{B(0,1)}, W)$ liegen. Aus der Linearität zweier solcher Abbildungen A und B folgt die Linearität von $A + B$ und $\lambda \cdot A$ für $\lambda \in \mathbb{K}$. Für jeden linearen Operator $A \in \mathcal{L}(V,W)$ und $v \in V \setminus \{0\}$ gilt $Av = \|v\| \cdot A(\frac{v}{\|v\|})$. Also ist A durch seine Werte auf $\overline{B(0,1)}$ eindeutig bestimmt und die Norm von $\mathcal{L}(V,W)$ ist einfach die Supremumsnorm der stetigen Abbildung von $\overline{B(0,1)}$ nach W. Also ist $\mathcal{L}(V,W)$ ein normierter Untervektorraum von $C_b(\overline{B(0,1)}, W)$. **q.e.d.**

Offenbar gilt $\|A\| = \sup\{\frac{\|Av\|}{\|v\|} \mid v \in V \setminus \{0\}\}$ für $A \in \mathcal{L}(V,W)$ und $\|Av\| \leq \|A\| \cdot \|v\|$ für alle $v \in V$. Aus der Konvergenz einer Folge $(A_n)_{n \in \mathbb{N}}$ in $\mathcal{L}(V,W)$ folgt also die gleichmäßige Konvergenz auf $\overline{B(0,1)}$ und die punktweise Konvergenz auf V. Für $V = \mathbb{K}^n$ ist der Abschluss der Einheitskugel $\overline{B(0,1)} = \{v \in \mathbb{K}^n \mid \|v\| \leq 1\}$ kompakt. Deshalb gibt es also für jedes $A \in \mathcal{L}(\mathbb{K}^n, W)$ ein $v \in \mathbb{K}^n \setminus \{0\}$ mit $\|A\| = \|A \frac{v}{\|v\|}\| = \frac{\|Av\|}{\|v\|}$.

Satz 9.60. *Seien V ein normierter Vektorraum und W ein Banachraum. Dann ist $\mathcal{L}(V,W)$ ein Banachraum.*

Beweis: Wegen Satz 9.44 (ii)-(iii) und Satz 9.59 müssen wir nur zeigen, dass die Grenzwerte $A \in C_b(\overline{B(0,1)}, W)$ von Cauchyfolgen $(A_n)_{n \in \mathbb{N}}$ in $\mathcal{L}(V,W)$ linear sind. Für jedes $v \in V$ ist $(A_n v)_{n \in \mathbb{N}}$ wegen $\|(A_n - A_m)v\| \leq \|A_n - A_m\| \cdot \|v\|$ eine Cauchyfolge in W. Also setzt sich A als punktweiser Grenzwert zu folgender Abbildung fort:

$$A : V \to W, \quad v \mapsto Av = \lim_{n \to \infty} A_n v \quad \text{für alle } v \in V.$$

Wir müssen nur noch zeigen, dass A linear ist. Aus der Linearität von A_n folgt

$$\begin{aligned}
&\|A(v+w)-(Av+Aw)\| \leq \\
&\leq \|(A-A_n)(v+w)-(A-A_n)v-(A-A_n)w))\| + \|A_n(v+w)-(A_nv+A_nw)\| \\
&\qquad \leq \|(A-A_n)(v+w)\| + \|(A-A_n)v\| + \|(A-A_n)w\|, \text{ und} \\
&\|\lambda Av - A(\lambda v)\| \leq \|\lambda(A-A_n)v-(A-A_n)(\lambda v)\| + \|\lambda A_nv - A_n(\lambda v)\| \\
&\qquad\qquad \leq |\lambda|\|(A-A_n)v\| + \|(A-A_n)(\lambda v)\|.
\end{aligned}$$

für alle $n \in \mathbb{N}$ und alle $v, w \in V$. Im Grenzwert $n \to \infty$ konvergieren die rechten Seiten für alle $v, w \in V$ und $\lambda \in \mathbb{K}$ punktweise gegen Null, so dass A linear ist. **q.e.d.**

Satz 9.61. *Seien U, V und W normierte Vektorräume und $A \in \mathcal{L}(U, V)$ und $B \in \mathcal{L}(V, W)$, dann ist $B \circ A \in \mathcal{L}(U, W)$ und es gilt $\|B \circ A\| \leq \|B\| \cdot \|A\|$. Insbesondere ist die Abbildung $\circ : \mathcal{L}(U, V) \times \mathcal{L}(V, W) \to \mathcal{L}(U, W), \quad (A, B) \mapsto B \circ A$ stetig.*

Beweis: Für alle $u \in U$ gilt $\|(B \circ A)u\| = \|B(A(u))\| \leq \|B\| \cdot \|Au\| \leq \|B\| \cdot \|A\| \cdot \|u\|$. Also folgt die Ungleichung $\|B \circ A\| \leq \|B\| \cdot \|A\|$ aus der Definition 9.58. Für zwei normierte Vektorräume V, W mit Normen $\|\cdot\|_V$ und $\|\cdot\|_W$ ist

$$\|\cdot\|_{V\times W} : V \times W \to \mathbb{R}, \quad (v, w) \mapsto \|v\|_V + \|w\|_W$$

eine Norm auf $V \times W$ und induziert die Metrik des kartesischen Produktes der metrischen Räume V und W. Für $(A, B), (A', B') \in \mathcal{L}(U, V) \times \mathcal{L}(V, W)$ gilt dann

$$\begin{aligned}
\|B \circ A - B' \circ A'\| &= \|B \circ A - B \circ A' + B \circ A' - B' \circ A'\| \\
&= \|B \circ (A - A') + (B - B') \circ A'\| \\
&\leq \|B\| \cdot \|A - A'\| + \|B - B'\| \cdot \|A'\| \\
&\leq (\|A - A'\| + \|B - B'\|)(\|B\| + \|A'\|) \\
&\leq (\|A - A'\| + \|B - B'\|)(\|B\| + \|A\| + \|A' - A\|) \\
&\leq (\|(A, B) - (A', B')\|(\|B\| + \|A\| + \|(A, B) - (A', B')\|).
\end{aligned}$$

Also ist diese Abbildung im Punkt $(A, B) \in \mathcal{L}(U, V) \times \mathcal{L}(V, W)$ stetig. **q.e.d.**

Wir bezeichnen die Komposition $B \circ A$ von linearen Operatoren auch mit BA.

Definition 9.62. *Auf einem normierten Vektorraum V ist $\mathcal{L}(V) = \mathcal{L}(V, V)$ mit*

$$\circ : \mathcal{L}(V) \times \mathcal{L}(V) \to \mathcal{L}(V), \quad (A, B) \mapsto AB \quad \textit{und} \quad \|\cdot\| : \mathcal{L}(V) \to \mathbb{R}, \quad A \mapsto \|A\|$$

eine normierte Algebra mit Eins $\mathbb{1}_V$ und eine Banachalgebra für einen Banachraum V.

Satz 9.63. *(Neumannsche Reihe) Sei $\mathcal{A}$ eine Banachalgebra mit Eins $\mathbb{1}$ und $A \in \mathcal{A}$ ein Operator mit $\|A\| < 1$. Dann ist $\mathbb{1} - A$ invertierbar und es gilt $(\mathbb{1} - A)^{-1} = \sum\limits_{n=0}^{\infty} A^n$.*

Beweis: Wegen $\|A^n\| \leq \|A\|^n$ ist $\left(\sum A^n\right)_{n\in\mathbb{N}_0}$ für $\|A\| < 1$ eine Cauchyfolge mit

$$\left\|\sum_{n=M}^{N} A^n\right\| \leq \sum_{n=M}^{N} \|A\|^n = \frac{\|A\|^M - \|A\|^{N+1}}{1 - \|A\|} \leq \frac{\|A\|^M}{1 - \|A\|} \text{ für } 0 \leq M \leq N.$$

Also konvergiert diese Reihe gegen ein $B \in \mathcal{A}$. Wie im Satz 9.61 ist wegen $\|B \cdot A\| \leq \|B\| \cdot \|A\|$ die Multiplikation $\cdot : \mathcal{A} \times \mathcal{A} \to \mathcal{A}$ stetig und es gilt

$$(\mathbb{1} - A)B = \sum_{n=0}^{\infty} A^n - \sum_{n=1}^{\infty} A^n = \mathbb{1} \quad \text{und} \quad B(\mathbb{1} - A) = \sum_{n=0}^{\infty} A^n - \sum_{n=1}^{\infty} A^n = \mathbb{1}.$$

Also ist $\sum_{n=0}^{\infty} A^n$ das Inverse von $\mathbb{1} - A$ mit $\left\|(\mathbb{1} - A)^{-1}\right\| \leq \frac{1}{1 - \|A\|}$. **q.e.d.**

Jede Potenzreihenfunktion $f(x) = \sum_{n\in\mathbb{N}_0} a_n x^n$ mit Konvergenzradius $R > 0$ definiert also auf jeder Banachalgenra $\mathcal{A}$ eine Abbildung

$$f : \{A \in \mathcal{A} \mid \|A\| < R\} \to \mathcal{A},\ A \mapsto f(A) = \sum_{n=0}^{\infty} a_n A^n \text{ mit } \|f(A)\| \leq \sum_{n=0}^{\infty} |a_n| \|A\|^n < \infty.$$

Viele der Aussagen, die wir für Potenzreihenfunktionen auf $\mathbb{K}$ gezeigt haben, lassen sich auf Potenzreihenfunktionen auf Banachalgebren $\mathcal{A}$ ausdehnen. Aber weil im allgemeinen $AB \neq BA$ für $A, B \in \mathcal{A}$, gilt im allgemeinen auch $\exp(A)\exp(B) \neq \exp(A + B)$.

Definition 9.64. *Eine Derivation einer Algebra $\mathcal{A}$ ist eine lineare Abbildung $D : \mathcal{A} \to \mathcal{A}$, die $D(A \cdot B) = D(A) \cdot B + A \cdot D(B)$ für alle $A, B \in \mathcal{A}$ erfüllt.*

Übungsaufgabe 9.65. **(i)** *Zeige, dass jedes Element A einer Algebra $\mathcal{A}$ folgende Derivation definiert:*

$$D_A : \mathcal{A} \to \mathcal{A}, \quad B \mapsto AB - BA$$

(ii) *Sei $\mathcal{A}$ eine Banachalgebra und $D \in \mathcal{L}(\mathcal{A})$ eine Derivation von $\mathcal{A}$. Zeige dass* $\exp(D)$ *ein Algebraisomorphismus von $\mathcal{A}$ ist, d.h. ein invertierbares Element von*

$$\{C \in \mathcal{L}(\mathcal{A}) \mid C(A \cdot B) = C(A) \cdot C(B) \text{ für alle } A, B \in \mathcal{A}\}.$$

(iii) *Zeige* $\exp(D_A)B = \exp(A) \cdot B \cdot \exp(-A)$ *für alle A, B in einer Banachalgebra $\mathcal{A}$.*

In der Vorlesung Analysis III wird gezeigt, dass die Derivationen der Algebra $C^\infty(\mathbb{R})$ von der Form $f \mapsto gf'$ für ein $g \in C^\infty(\mathbb{R})$ sind. Außerdem werden die entsprechenden Algebraisomorphismen bestimmt, soweit sie existieren.

In dem Buch L. Gillman, M. Jerison:"Rings of continuous functions" wird gezeigt, dass für jeden kompakten metrischen Raum X die einzigen Algebraisomorphismen von $C(X, \mathbb{R})$ von der Form $f \mapsto f \circ \Phi$ für einen Homöomorphismus $\Phi : X \to X$ sind. Daraus lässt sich folgern, dass $\mathcal{L}(C(X, \mathbb{R}))$ nur die triviale Derivation von $C(X, \mathbb{R})$ enthält.

Kapitel 10

Differentialrechnung von Funktionen mehrerer Veränderlicher

10.1 Ableitungen von $f : X \to Y$

Definition 10.1. *(Ableitung) Seien X und Y normierte Vektorräume. Eine Abbildung f von einer offenen Menge $U \subset X$ nach Y heißt im Punkt $x_0 \in U$ differenzierbar, wenn es ein $A \in \mathcal{L}(X, Y)$ gibt, so dass die folgende Abbildung in x_0 stetig ist:*

$$U \to \mathbb{R}, \quad x \mapsto \begin{cases} \frac{\|f(x)-f(x_0)-A(x-x_0)\|}{\|x-x_0\|} & \text{für } x \neq x_0 \\ 0 & \text{für } x = x_0. \end{cases}$$

A heißt Ableitung von f bei x_0 und wird mit $f'(x_0)$ oder $\frac{df}{dx}(x_0)$ bezeichnet. Wenn A und B beide diese Bedingung erfüllen, dann gibt es für alle $\epsilon > 0$ ein $\delta > 0$ mit

$$\frac{\|(A-B)(x-x_0)\|}{\|x-x_0\|} \leq \frac{\|A(x-x_0)-f(x)+f(x_0)\|}{\|x-x_0\|} + \frac{\|f(x)-f(x_0)-B(x-x_0)\|}{\|x-x_0\|} < 2\epsilon$$

für alle $x \in B(x_0, \delta) \setminus \{x_0\}$. Mit $x - x_0 = \delta y \in X$ für $0 < \|y\| < 1$ folgt $\|(A-B)y\| = \frac{\|(A-B)\delta y\|}{\delta} < 2\epsilon\|y\| < 2\epsilon$ und $\|A - B\| \leq 2\epsilon$. Weil das für alle $\epsilon > 0$ gilt, ist $A = B$.

Satz 10.2. *Sei $f : U \subset X \to Y$ in $x_0 \in U$ differenzierbar. Dann ist f in x_0 stetig.*

Beweis: Wegen $f(x) - f(x_0) = A(x - x_0) + (f(x) - f(x_0) - A(x - x_0))$ folgt aus der Differenzierbarkeit von f in x_0, dass es ein $\delta > 0$ gibt, so dass $\|f(x) - f(x_0)\| \leq (\|A\| + 1)\|x - x_0\|$ gilt für alle $\|x - x_0\| < \delta$. Dann ist f in x_0 auch stetig. **q.e.d.**

Beispiel 10.3. **(i)** *Sei f konstant. Dann ist* $\frac{\|f(x)-f(x_0)\|}{\|x-x_0\|} = 0$ *für $x \neq x_0$. Also ist f differenzierbar mit $f'(x_0) = 0$.*

(ii) *Für $f \in \mathcal{L}(X,Y)$ und $x \neq x_0 \in X$ gilt* $\quad \dfrac{\|f(x) - f(x_0) - f(x - x_0)\|}{\|x - x_0\|} = 0.$
Also ist f in x_0 differenzierbar und $f'(x_0) = f$.
(iii) *Die Abbildung $Y \to \mathcal{L}(\mathbb{R}, Y), \quad y \mapsto$ Multiplikation mit y besitzt offenbar die Umkehrabbildung $\mathcal{L}(\mathbb{R}, Y) \to Y, \quad A \mapsto A(1)$ und ist ein isometrischer Isomorphismus von normierten Vektorräumen. Deshalb können wir die Ableitungen von differenzierbaren $f : (a,b) \to Y$ auf $(a,b) \subset \mathbb{R}$ auch durch $f' : (a,b) \to Y$ beschreiben. Für $x \neq x_0$ gilt*

$$\frac{\|f(x) - f(x_0) - f'(x_0)(x - x_0)\|}{\|x - x_0\|} = \left\| \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right\|.$$

Dann sind für $f : (a,b) \to \mathbb{K}$ und $x_0 \in (a,b)$ die Definitionen 7.1 und 10.1 äquivalent.

Satz 10.4. **(i)** *Sei $f, g : U \subset X \to Y$ in $x_0 \in U$ differenzierbar. Dann sind $f + g$ und $\lambda \cdot f$ für alle $\lambda \in \mathbb{K}$ in x_0 differenzierbar und es gilt*

$$(f+g)'(x_0) = f'(x_0) + g'(x_0) \quad \text{bzw.} \quad (\lambda f)'(x_0) = \lambda f'(x_0).$$

(ii) *Seien $f, g : U \subset X \to Y$ in x_0 differenzierbar und Y eine normierte Algebra. Dann ist $f \cdot g$ in x_0 differenzierbar und es gilt (Leibnizregel)*

$$(f \cdot g)'(x_0) = f'(x_0) \cdot g(x_0) + f(x_0) \cdot g'(x_0) \quad \text{mit}$$
$$(f'(x_0) \cdot g(x_0))(x) = f'(x_0)(x) \cdot g(x_0) \text{ und } (f(x_0) \cdot g'(x_0))(x) = f(x_0) \cdot g'(x_0)(x).$$

(iii) *Seien $f : U \subset X \to Y$ in $x_0 \in U$ und $g : V \subset Y \to Z$ in $f(x_0) \in V$ differenzierbar mit $f[U] \subset V$. Dann ist $g \circ f$ im Punkt x_0 differenzierbar mit*

$$(g \circ f)'(x_0) = g'(f(x_0)) \circ f'(x_0).$$

Beweis:(i) Mit den drei Eigenschaften von Normen folgt (i) für $x \neq x_0$ aus

$$\frac{\|f(x) + g(x) - f(x_0) - g(x_0) - f'(x_0)(x - x_0) - g'(x_0)(x - x_0)\|}{\|x - x_0\|}$$
$$\leq \frac{\|f(x) - f(x_0) - f'(x_0)(x - x_0)\|}{\|x - x_0\|} + \frac{\|g(x) - g(x_0) - g'(x_0)(x - x_0)\|}{\|x - x_0\|}.$$
$$\frac{\|\lambda f(x) - \lambda f(x_0) - \lambda f'(x_0)(x - x_0)\|}{\|x - x_0\|} = |\lambda| \frac{\|f(x) - f(x_0) - f'(x_0)(x - x_0)\|}{\|x - x_0\|}.$$

(ii) Weil in einer normierten Algebra $\|AB\| \leq \|A\| \cdot \|B\|$ gilt, folgt für $x \neq x_0$

$$\frac{\|f(x) \cdot g(x) - f(x_0) \cdot g(x_0) - f'(x_0)(x - x_0)g(x_0) - f(x_0)g'(x_0)(x - x_0)\|}{\|x - x_0\|}$$
$$\leq \frac{\|f(x) - f(x_0) - f'(x_0)(x - x_0)\|}{\|x - x_0\|} \|g(x)\| + \|f'(x_0)\| \cdot \|g(x) - g(x_0)\| +$$
$$+ \|f(x_0)\| \frac{\|g(x) - g(x_0) - g'(x_0)(x - x_0)\|}{\|x - x_0\|}.$$

Weil g in x_0 stetig ist, gibt es für jedes $\epsilon > 0$ ein $\delta > 0$, so dass aus $\|x - x_0\| < \delta$ folgt $\|g(x) - g(x_0)\| < \epsilon$ bzw. $\|g(x)\| \leq \|g(x_0)\| + \epsilon$. Dann folgt (ii).

(iii) Aus Satz 9.61 folgt für $x \neq x_0$ $\dfrac{\|(g \circ f)(x) - (g \circ f)(x_0) - g'(f(x_0))f'(x_0)(x - x_0)\|}{\|x - x_0\|}$

$$\leq \frac{\|g(f(x)) - g(f(x_0)) - g'(f(x_0))(f(x) - f(x_0))\|}{\|f(x) - f(x_0)\|} \cdot \frac{\|f(x) - f(x_0)\|}{\|x - x_0\|} + \frac{\|g'(f(x_0))(f(x) - f(x_0) - f'(x_0)(x - x_0))\|}{\|x - x_0\|}$$

$$\leq \frac{\|g(f(x)) - g(f(x_0)) - g'(f(x_0))(f(x) - f(x_0))\|}{\|f(x) - f(x_0)\|} \cdot \left(\|f'(x_0)\| + \frac{\|f(x) - f(x_0) - f'(x_0)(x - x_0)\|}{\|x - x_0\|}\right) + \|g'(f(x_0))\| \frac{\|f(x) - f(x_0) - f'(x_0)(x - x_0))\|}{\|x - x_0\|}.$$

Ersetze hierbei für $f(x) = f(x_0)$ die beiden Differenzenquotienten von g wie in der Definition von $g'(f(x_0))$ durch Null. Aus Satz 10.2 und Korollar 9.32 folgt (iii). **q.e.d.**

10.2 Schrankensatz

Wir verallgemeinern in diesem Abschnitt den Schrankensatz auf differenzierbare Abbildungen zwischen normierten Vektorräumen.

Lemma 10.5. *Seien f eine stetige Abbildung von einem kompakten Intervall $[a, b]$ mit $a < b$ in einen normierten Vektorraum Y und ϕ eine stetige reelle Funktion auf $[a, b]$. Wenn im Komplement einer abzählbaren Teilmenge von (a, b) sowohl f als auch ϕ differenzierbar sind und dort gilt $\|f'\| \leq \phi'$, dann gilt auch $\|f(b) - f(a)\| \leq \phi(b) - \phi(a)$.*

Beweis: Sei $(x_n)_{n\in\mathbb{N}}$ eine Abzählung der Punkte in $[a, b]$, an denen entweder f oder ϕ nicht differenzierbar ist oder $\|f'\| > \phi'$. Mit kontinuierlicher Induktion zeigen wir

$$\|f(x) - f(a)\| \leq \phi(x) - \phi(a) + \epsilon(x - a) + \epsilon \sum\nolimits_{x_n < x} 2^{-n} \quad \text{für alle } \epsilon > 0 \text{ und } x \in [a, b].$$

Sei also $\epsilon > 0$ und A_ϵ die Menge aller $y \in [a, b]$, so dass diese Ungleichung für alle $x \in [a, y)$ gilt. Wegen der Stetigkeit von f und ϕ gilt auch für $y = \sup A_\epsilon$

$$\begin{aligned} \|f(y) - f(a)\| &\leq \phi(y) - \phi(a) + \epsilon(y - a) + \epsilon \sup\left\{\sum\nolimits_{x_n < x} 2^{-n} \mid x \in (a, y)\right\} \\ &= \phi(y) - \phi(a) + \epsilon(y - a) + \epsilon \sum\nolimits_{x_n < y} 2^{-n}. \end{aligned}$$

Deshalb ist A_ϵ ein Intervall von der Form $A_\epsilon = [a, y]$. Wenn $y \in (a, b)$ und f und ϕ in y differenzierbar sind und $\|f'(y)\| \le \phi'(y)$ gilt, dann gibt es ein $\delta > 0$, so dass

$$\phi'(y) - \frac{\epsilon}{2} < \frac{\phi(x) - \phi(y)}{x - y} < \phi'(y) + \frac{\epsilon}{2} \quad \text{und} \quad \frac{\|f(x) - f(y) - f'(y)(x - y)\|}{|x - y|} < \frac{\epsilon}{2}$$

für alle $x \in (y - \delta, y) \cup (y, y + \delta)$ gilt. Dann folgt für $x \in (y, y + \delta)$

$$\begin{aligned}
\|f(x) - f(y)\| &\le \|f'(y)(x - y)\| + \|f(x) - f(y) - f'(y)(x - y)\| \\
&< \left(\|f'(y)\| + \frac{\epsilon}{2}\right)|x - y| \le \left(\phi'(y) - \frac{\epsilon}{2} + \epsilon\right)(x - y) \\
&< \left(\frac{\phi(x) - \phi(y)}{x - y} + \epsilon\right)(x - y) = \phi(x) - \phi(y) + \epsilon(x - y) \quad \text{und} \\
\|f(x) - f(a)\| &\le \|f(x) - f(y)\| + \|f(y) - f(a)\| \\
&< \phi(x) - \phi(y) + \epsilon(x - y) + \phi(y) - \phi(a) + \epsilon(y - a) + \epsilon \sum\nolimits_{x_n < y} 2^{-n} \\
&\le \phi(x) - \phi(a) + \epsilon(x - a) + \epsilon \sum\nolimits_{x_n < x} 2^{-n}.
\end{aligned}$$

Woraus $y + \delta \in A_\epsilon$ folgt, im Widerspruch zu $y = \sup A_\epsilon$. Wenn es andererseits ein $N \in \mathbb{N}$ gibt mit $x_N = y < b$, dann gibt es ein $\delta > 0$, so dass aus $x \in (y, y + \delta)$ folgt

$$\|f(x) - f(y)\| - (\phi(x) - \phi(y)) < \epsilon 2^{-N}.$$

Dann folgt für dieselben x wieder $\quad \|f(x) - f(a)\| \le \|f(x) - f(y)\| + \|f(y) - f(a)\|$

$$< \epsilon 2^{-N} + \phi(x) - \phi(a) + \epsilon(y - a) + \epsilon \sum_{x_n < y} 2^{-n} \le \phi(x) - \phi(a) + \epsilon(x - a) + \epsilon \sum_{x_n < x} 2^{-n}.$$

Also gilt wieder $y + \delta \in A_\epsilon$, was $y = \sup A_\epsilon$ widerspricht. Dann muß aber $\sup A_\epsilon = b$ gelten. Weil das für alle $\epsilon > 0$ gilt, folgt auch $\quad \|f(b) - f(a)\| \le \phi(b) - \phi(a)$. **q.e.d.**

Korollar 10.6. *(Schrankensatz) Sei f eine stetige Abbildung von einer offenen Teilmenge U des normierten Vektorraumes X in den normierten Vektorraum Y. Wenn f im Komplement einer abzählbaren Teilmenge S von $D = \{(1-t)a + tb \mid t \in [0, 1]\} \subset U$ differenzierbar ist mit $a, b \in U$, und die Ableitung auf $D \setminus S$ beschränkt ist, dann gilt*

$$\begin{aligned}
\|f(b) - f(a)\| &\le \|b - a\| \sup\{\|f'(x)\| \mid x \in D \setminus S\} && \textit{und} \\
\|f(b) - f(a) - A(b - a)\| &\le \|b - a\| \sup\{\|f'(x) - A\| \mid x \in D \setminus S\} && \textit{für } A \in \mathcal{L}(X, Y).
\end{aligned}$$

Auf konvexen Mengen ist jede obere Schranke an $\|f'\|$ eine Lipschitzkonstante von f.

Beweis: Die Abbildung $x : \mathbb{R} \to X$, $t \mapsto x(t) = (t - 1)a + tb = a + t(b - a)$ ist wegen Beispiel 10.3 (i)-(ii) und Satz 10.4 (i) diffferenzierbar mit $x'(t) : s \mapsto s(b - a)$

als Element von $\mathcal{L}(\mathbb{R}, X)$ bzw. mithilfe von Beispiel 10.3 (iii) mit $x'(t) = (b - a) \in X$. Wegen Satz 10.4 (iii) ist dann auch die Abbildung $t \mapsto f(x(t))$ bei den $t \in [0, 1]$ differenzierbar, für die f bei $x(t) \in U$ differenzierbar ist, mit der Ableitung $(f \circ x)'(t) : s \mapsto sf'(x(t))(b-a)$ bzw. mithilfe von Beispiel 10.3 (iii) $(f \circ x)'(t) = f'(x(t))(b-a) \in Y$. Dann folgt die erste Behauptung aus Lemma 10.5 mit den beiden Funktionen

$$f \circ x : [0, 1] \to Y, t \mapsto f(x(t)), \quad \phi : [0, 1] \to \mathbb{R}, t \mapsto t\|b - a\| \sup\{\|f'(x)\| \mid x \in D \setminus S\}.$$

Die zweite Behauptung folgt aus diesem Lemma mit den beiden Funktionen

$$t \mapsto f(x(t)) - tA(b - a) \in Y, \quad t \mapsto t\|b - a\| \sup\{\|f'(x) - A\| \mid x \in D \setminus S\} \in \mathbb{R}.$$ **q.e.d.**

10.3 Partielle Ableitungen

Definition 10.7. *Eine Abbildung f von einer offenen Teilmenge U eines normierten Vektorraumes X in einen normierten Vektorraum Y heißt stetig differenzierbar, wenn*

(i) *f in allen $x_0 \in U$ differenzierbar ist, und*

(ii) *die Abbildung $f' : U \to \mathcal{L}(X, Y), \quad x \mapsto f'(x)$ stetig ist.*

Wegen Beispiel 10.3 (iii) stimmt im Falle von $X = Y = \mathbb{R}$ diese Definition mit der Definition von stetig differenzierbaren reellen Funktionen auf Intervallen in $\mathbb{R}$ überein.

Definition 10.8. *(partielle Ableitung) Sei f eine Funktion von einer offenen Teilmenge $U \subset X_1 \times X_2$ des kartesischen Produktes der normierten Vektorräume X_1 und X_2 in den normierten Vektorraum Y. Dann heißt f im Punkt $(x_1, x_2) \in U \subset X_1 \times X_2$ partiell differenzierbar, falls die Abbildung $x \mapsto f(x, x_2)$ im Punkt $x = x_1$, und die Abbildung $x \mapsto f(x_1, x)$ im Punkt $x = x_2$ differenzierbar ist. Die Ableitungen heißen partielle Ableitungen an der Stelle (x_1, x_2) und werden mit $\frac{\partial f}{\partial x_1}(x_1, x_2)$ und $\frac{\partial f}{\partial x_2}(x_1, x_2)$ bezeichnet. Allgemeiner heißt eine Abbildung von einer offenen Menge $U \subset X_1 \times \ldots \times X_n$ eines n-fachen kartesischen Produktes von normierten Vektorräumen in einen normierten Vektorraum Y im Punkt $(x_1, \ldots, x_n) \in U$ partiell differenzierbar, wenn für $i = 1, \ldots, n$ die Abbildungen $x \mapsto f(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n)$ bei $x = x_i$ differenzierbar sind.*

Die wichtigsten Beispiele sind reelle Funktionen auf offenen Teilmengen des $\mathbb{R}^n$. Für jedes $A \in \mathcal{L}(X_1 \times X_2, Y)$ sind die beiden folgenden Abbildungen stetig und linear

$$A_1 : X_1 \to Y, \quad x_1 \mapsto A((x_1, 0)) \quad \text{und} \quad A_2 : X_2 \to Y, \quad x_2 \mapsto A((0, x_2)).$$

Weil für $x \in X_1 \setminus \{x_1\}$ mit $(x, x_2) \in U$ bzw. $x \in X_2 \setminus \{x_2\}$ mit $(x_1, x) \in U$

$$\frac{\|f(x, x_2) - f(x_1, x_2) - A((x, x_2) - (x_1, x_2))\|}{\|(x, x_2) - (x_1, x_2)\|} = \frac{\|f(x, x_2) - f(x_1, x_2) - A_1(x - x_1)\|}{\|x - x_1\|}$$
$$\frac{\|f(x_1, x) - f(x_1, x_2) - A((x_1, x) - (x_1, x_2))\|}{\|(x_1, x) - (x_1, x_2)\|} = \frac{\|f(x_1, x) - f(x_1, x_2) - A_2(x - x_2)\|}{\|x - x_2\|}$$

gilt, folgt aus den Definitionen der folgende Satz:

Satz 10.9. *Eine im Punkt $(x_1, x_2) \in U$ differenzierbar Funktion f von einer offenen Teilmenge $U \subset X_1 \times X_2$ des kartesischen Produktes zweier normierter Vektorräume in einen normierten Vektorraum Y ist in (x_1, x_2) auch partiell differenzierbar.* **q.e.d.**

Beispiel 10.10. *Die Umkehrung dieses Satzes gilt nicht:*

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad (x, y) \mapsto \begin{cases} \frac{2xy}{x^2+y^2} & \text{für } (x, y) \neq 0 \\ 0 & \text{für } (x, y) = 0 \end{cases} \quad \textit{besitzt die partiellen Ableitungen}$$

$$\frac{\partial f}{\partial x} = \begin{cases} \frac{2y}{x^2+y^2} - \frac{4x^2y}{(x^2+y^2)^2} & \text{für } (x, y) \neq 0 \\ 0 & \text{für } (x, y) = 0 \end{cases} \quad \frac{\partial f}{\partial y} = \begin{cases} \frac{2x}{x^2+y^2} - \frac{4y^2x}{(x^2+y^2)^2} & \text{für } (x, y) \neq 0 \\ 0 & \text{für } (x, y) = 0 \end{cases}$$

Für alle $r \in (0, \infty)$ und alle $\phi \in \mathbb{R}$ gilt $f(r\cos\phi, r\sin\phi) = 2\sin\phi\cos\phi = \sin(2\phi)$, und deshalb $\lim_{r\to 0+} f(r\cos\phi, r\sin\phi) = \sin(2\phi)$. Also ist f im Punkt $(x, y) = 0$ nicht stetig.

Aber es gilt folgende Umkehrung.

Satz 10.11. *Seien X_1, X_2, Y normierte Vektorräume, $U \subset X_1 \times X_2$ offen und $f : U \to Y$ bei $x = (x_1, x_2) \in U$ partiell differenzierbar. Wenn eine der partiellen Ableitungen*

$$\frac{\partial f}{\partial x_1} : U \to \mathcal{L}(X_1, Y) \quad \textit{oder} \quad \frac{\partial f}{\partial x_2} : U \to \mathcal{L}(X_2, Y)$$

auf U existiert und bei x stetig ist, dann ist f bei x differenzierbar. Also ist f genau dann auf U stetig differenzierbar, wenn f auf U stetig partiell differenzierbar ist.

Beweis: Wenn $A_1 \in \mathcal{L}(X_1, Y)$ und $A_2 \in \mathcal{L}(X_2, Y)$, dann sind auch die Abbildungen

$$X_1 \times X_2 \to Y, \quad (x_1, x_2) \to A_1(x_1) \text{ bzw.} \quad X_1 \times X_2 \to Y, \quad (x_1, x_2) \to A_2(x_2)$$

lineare und stetige Abbildungen in $\mathcal{L}(X_1 \times X_2, Y)$. Also ist

$$A_1 \times A_2 : X_1 \times X_2 \to Y, \quad (x_1, x_2) \mapsto A_1(x_1) + A_2(x_2)$$

eine stetige lineare Abbildung in $\mathcal{L}(X_1 \times X_2, Y)$. Umgekehrt sind für jede stetige lineare Abbildung $A \in \mathcal{L}(X_1 \times X_2, Y)$ die Abbildungen $A_1 : X_1 \to Y$, $x_1 \mapsto A((x_1, 0))$ und $A_2 : X_2 \to Y$, $x_2 \mapsto A((0, x_2))$ stetig und linear. Und es gilt

$$\|A((x_1, x_2))\| = \|A((x_1, 0)) + A((0, x_2))\| \leq \|A((x_1, 0))\| + \|A((0, x_2))\|.$$

Aufgrund der Definition der Norm von $\mathcal{L}(X_1 \times X_2, Y)$ folgt dann

$$\|A_1\| \leq \|A\| \qquad \|A_2\| \leq \|A\| \qquad \|A\| \leq \|A_1\| + \|A_2\| \leq 2\|A\|.$$

Also ist die Abbildung $\quad \mathcal{L}(X_1, Y) \times \mathcal{L}(X_2, Y) \to \mathcal{L}(X_1 \times X_2, Y), \quad (A_1, A_2) \mapsto A$ eine bijektive Abbildung von normierten Vektorräumen und die beiden Normen von $\mathcal{L}(X_1, Y) \times \mathcal{L}(X_2, Y)$ und $\mathcal{L}(X_1 \times X_2, Y)$ sind bezüglich dieser Identifikation äquivalent. Daraus folgt, dass für jede stetig differenzierbare Funktion $f : U \to Y$, die beiden partiellen Ableitungen $\frac{\partial f}{\partial x_1} : U \to \mathcal{L}(X_1, Y)$ und $\frac{\partial f}{\partial x_2} : U \to \mathcal{L}(X_2, Y)$ stetig sind.

Wenn umgekehrt f in (x_1, x_2) partiell differenzierbar ist, dann folgt

$$\left\| f(y_1, y_2) - f(x_1, x_2) - \frac{\partial f(x_1, x_2)}{\partial x_1}(y_1 - x_1) - \frac{\partial f(x_1, x_2)}{\partial x_2}(y_2 - x_2) \right\| \leq$$
$$\leq \left\| f(y_1, y_2) - f(x_1, y_2) - \frac{\partial f(x_1, x_2)}{\partial x_1}(y_1 - x_1) \right\| +$$
$$+ \left\| f(x_1, y_2) - f(x_1, x_2) - \frac{\partial f(x_1, x_2)}{\partial x_2}(y_2 - x_2) \right\|.$$

Wenn $\frac{\partial f}{\partial x_1}$ auf U existiert und bei (x_1, x_2) stetig ist, dann gibt es für jedes $\epsilon > 0$ ein $\delta > 0$, so dass

$$\left\| \frac{\partial f(z_1, z_2)}{\partial x_1} - \frac{\partial f(x_1, x_2)}{\partial x_1} \right\| < \epsilon$$

für $(z_1, z_2) \in B((x_1, x_2), \delta)$ gilt. Aus Korollar 10.6 folgt für $(y_1, y_2) \in B((x_1, x_2), \delta)$

$$\left\| f(y_1, y_2) - f(x_1, y_2) - \frac{\partial f(x_1, x_2)}{\partial x_1}(y_1 - x_1) \right\| < \epsilon \|y_1 - x_1\|.$$

Weil $\frac{\partial f}{\partial x_2}$ in (x_1, x_2) existiert, gibt es auch ein $\delta' > 0$, so dass für $y_2 \in B(x_2, \delta')$ folgt

$$\left\| f(x_1, y_2) - f(x_1, x_2) - \frac{\partial f(x_1, x_2)}{\partial x_2}(y_2 - x_2) \right\| < \epsilon \|y_2 - x_2\|.$$

Dann folgt für $(y_1, y_2) \in B((x_1, x_2), \min\{\delta, \delta'\})$ auch

$$\|f(y_1, y_2) - f(x_1, x_2) - f'(x_1, x_2)((y_1, y_2) - (x_1, x_2))\| < \epsilon(\|y_1 - x_1\| + \|y_2 - x_2\|),$$

wobei $f'(x_1, x_2) = \frac{\partial f(x_1,x_2)}{\partial x_1} \times \frac{\partial f(x_1,x_2)}{\partial x_2}$ durch die partiellen Ableitungen gegeben ist. Also ist f differenzierbar, und mit den partiellen Ableitungen stetig differenzierbar. **q.e.d.**

Durch mehrmaliges Anwenden erhalten wir dann auch die entsprechende Aussage für Abbildungen von offenen Teilmengen U des n-fachen kartesischen Produktes von normierten Vektorräumen in einen normierten Vektorraum. Unsere wichtigsten Beispiele sind wieder reelle Funktionen auf offenen Teilmengen des $\mathbb{R}^n$.

Beispiel 10.12. (i)
$$f : \mathbb{R}^2 \to \mathbb{R}, \quad (x, y) \mapsto \begin{cases} \frac{2xy}{x^2+y^2} & \textit{für } (x, y) \neq 0 \\ 0 & \textit{für } (x, y) = 0. \end{cases}$$

$$\frac{\partial f(x, y)}{\partial x} = \begin{cases} \frac{2y(y^2-x^2)}{(x^2+y^2)^2} & \textit{für } (x, y) \neq 0 \\ 0 & \textit{für } (x, y) = 0 \end{cases} \quad \frac{\partial f(x, y)}{\partial y} = \begin{cases} \frac{2x(x^2-y^2)}{(x^2+y^2)^2} & \textit{für } (x, y) \neq 0 \\ 0 & \textit{für } (x, y) = 0 \end{cases}$$

Für $y = 0$ ist $\frac{\partial f}{\partial x} = 0$ und für $x = 0$ ist $\frac{\partial f}{\partial y} = 0$, so dass diese partiellen Ableitungen für alle $(x, y) \in \mathbb{R}^2$ existieren. Allerdings sind sie in keiner Umgbung von $(x, y) = (0, 0)$ beschränkt, und deshalb auch nicht stetig. Wir hatten schon im Beispiel 10.10 gesehen, dass f bei $(0, 0)$ nicht stetig und deshalb auch nicht differenzierbar ist.

(ii)
$$f : \mathbb{R}^2 \to \mathbb{R}, \quad (x, y) \mapsto \begin{cases} \frac{x^3 - y^3}{x^2 + y^2} & \text{für } (x, y) \neq 0 \\ 0 & \text{für } (x, y) = 0. \end{cases}$$

Offenbar gilt $|f(x, y)| \le \frac{|x|^3}{x^2+y^2} + \frac{|y|^3}{x^2+y^2} \le |x| + |y|$. Also ist f stetig.

$$\frac{\partial f}{\partial x} = \begin{cases} \frac{3x^2(x^2+y^2) - 2x(x^3-y^3)}{(x^2+y^2)^2} = \frac{x(x^3+3xy^2+2y^3)}{(x^2+y^2)^2} & \text{für } (x, y) \neq 0 \\ 1 & \text{für } (x, y) = 0 \end{cases}$$

$$\frac{\partial f}{\partial y} = \begin{cases} -\frac{3y^2(x^2+y^2) + 2y(x^3-y^3)}{(x^2+y^2)^2} = -\frac{y(y^3+3yx^2+2x^3)}{(x^2+y^2)^2} & \text{für } (x, y) \neq 0 \\ -1 & \text{für } (x, y) = 0. \end{cases}$$

Wegen $\frac{\partial f(x,0)}{\partial x} = 1$ und $\frac{\partial f(0,y)}{\partial y} = -1$ ist f partiell differenzierbar. In Beispiel 10.14 (iii) werden wir sehen, dass f in $(0, 0)$ nicht differenzierbar ist.
(iii) *Alle Polynome in endlich vielen Variablen sind partiell unendlich oft stetig differenzierbar, und deshalb differenzierbar.*

Definition 10.13. *(Richtungsableitung) Für eine Funktion $f : U \to Y$ von einer offenen Teilmenge U eines normierten Vektorraumes X in einem normierten Vektorraum Y ist die Richtungsableitung in $x_0 \in U$ in Richtung $x_1 \in X$ die Ableitung bei $t = 0$ von*

$$(-\epsilon, \epsilon) \to Y, \quad t \mapsto f(x_0 + tx_1).$$

Wie in Beispiel 10.3 (iii) identifizieren wir dabei $\mathcal{L}(\mathbb{R}, Y)$ durch $A \mapsto A(1)$ mit Y.

Beispiel 10.14. (i) *Sei $U \subset X$ offen und $f : U \to Y$ in $x_0 \in U$ differenzierbar. Für $x_1 \in X$ ist die Abbildung $x : \mathbb{R} \to X$, $t \mapsto x(t) = x_0 + tx_1$ wegen Beispiel 10.3 (i)-(ii) und Satz 10.4 differenzierbar mit $x'(t) = x_1$ im Sinne von Beispiel 10.3 (iii). Dann gibt es ein Intervall $(-\epsilon, \epsilon)$ im Urbild $x^{-1}[U]$ von U unter x, und wegen Satz 10.4 (iii) ist $(-\epsilon, \epsilon) \to Y$, $t \mapsto f(x(t))$ bei $t = 0$ differenzierbar mit der Ableitung $f'(x_0)(x_1)$ im Sinne von Beispiel 10.3 (iii). Also existiert die Richtungsableitung und es gilt*

$$\frac{d}{dt} f(x_0 + tx_1)|_{t=0} = f'(x_0)(x_1).$$

(ii)
$$f : \mathbb{R}^2 \to \mathbb{R}, \quad (x, y) \mapsto \begin{cases} \frac{2xy}{x^2+y^2} & \text{für } (x, y) \neq 0 \\ 0 & \text{für } (x, y) = 0. \end{cases}$$

Für $t \neq 0$ ist dann $f(t\cos(\phi), t\sin(\phi)) = \sin(2\phi)$ und $f(t\cos(\phi), t\sin(\phi)) = 0$ für $t = 0$. Also ist f in $t = 0$ für $\phi \notin \frac{\pi}{2}\mathbb{Z}$ nicht stetig und auch nicht differenzierbar. Für $\phi \in \frac{\pi}{2}\mathbb{Z}$ existieren die Richtungsableitungen und verschwinden.

(iii) $\quad f:\mathbb{R}^2\to\mathbb{R},\quad (x,y)\mapsto\begin{cases}\frac{x^3-y^3}{x^2+y^2} & \text{für } (x,y)\neq 0\\ 0 & \text{für } (x,y)=0.\end{cases}$

Dann gilt $f(t(x,y))=f(tx,ty)=tf(x,y))$ für $(x,y)\in\mathbb{R}^2$ und $x_0=(0,0)$. Also existieren alle Richtungsableitungen und setzen sich im Punkt $(0,0)$ zu f zusammen. Wegen $2=f(2,0)\neq f(1,1)+f(1,-1)=0+1$ ist f im Punkt $(0,0)$ nicht differenzierbar.

(iv) $\quad f:\mathbb{R}^2\to\mathbb{R},\quad (x,y)\mapsto\begin{cases}\frac{2x^3y}{x^4+y^2} & \text{für } (x,y)\neq 0\\ 0 & \text{für } (x,y)=0.\end{cases}$

Die Richtungsableitungen von f in $(0,0)$ in Richtung von (x,y) verschwinden alle:

$$\frac{d}{dt}f(tx,ty)\bigg|_{t=0}=\begin{cases}\frac{d}{dt}\frac{2t^2x^3y}{t^2x^4+y^2}\Big|_{t=0}=0 & \text{für } y\neq 0\\ 0 & \text{für } y=0,\end{cases}\quad |f(x,y)|=2|x|\sqrt{\frac{x^4}{x^4+y^2}\frac{y^2}{x^4+y^2}}\leq 2|x|.$$

Wegen $\frac{d}{dt}(t,t^2)=(1,2t)$ und $\frac{d}{dt}f(t,t^2)=\frac{d}{dt}t=1\neq 0$ ist f in $(0,0)$ nicht differenzierbar.

Wir wollen den wichtigsten Fall von Funktionen $\mathbb{R}^n\to\mathbb{R}^m$ genauer betrachten.

Definition 10.15. *(Partielle Ableitungen in $\mathbb{R}^n$) Sei $f:U\to\mathbb{R}^m$ eine Funktion von einer offenen Teilmenge $U\subset\mathbb{R}^n$ in den $\mathbb{R}^m$. Dann sind die Komponenten $(f_1,\dots,f_m)$ von f offenbar reelle Funktionen auf U. Die Funktion f ist in $(x_1,\dots,x_n)\in U$ genau dann partiell differenzierbar, wenn für alle $i=1,\dots,n$ und $j=1,\dots,m$ die Funktionen*

$$x\mapsto f_j(x_1,\dots,x_{i-1},x,x_{i+1},\dots,x_n)$$

bei $x=x_i$ differenzierbar sind. Die entsprechenden Ableitungen heißen partielle Ableitungen von f und werden mit $\frac{\partial f_j}{\partial x_i}(x_1,\dots,x_n)$ bezeichnet. Wenn diese partiellen Ableitungen für alle $x\in U$ existieren, heißt f auf U partiell differenzierbar.

Definition 10.16. *(Vektorfeld, Gradient, Divergenz und Rotation) Eine Abbildung von einer offenen Menge $U\subset\mathbb{R}^n$ nach $\mathbb{R}^n$ wird Vektorfeld genannt. Das Vektorfeld*

$$\operatorname{grad} f=\nabla f=\left(\frac{\partial f}{\partial x_1},\dots,\frac{\partial f}{\partial x_n}\right)$$

der partiellen Ableitungen einer partiell differenzierbaren reellen Funktion f auf einer offenen Menge $U\subset\mathbb{R}^n$ heißt Gradient von f. Wenn f ein partiell differenzierbares Vektofeld ist, dann ist die Divergenz von f folgende reelle Funktion

$$\operatorname{div} f=\nabla\cdot f=\frac{\partial f_1}{\partial x_1}+\ldots+\frac{\partial f_n}{\partial x_n}.$$

Die lineare Abbildung $\quad\triangle:\quad f\mapsto\triangle f=\operatorname{div}\operatorname{grad} f=\dfrac{\partial^2 f}{\partial x_1^2}+\ldots+\dfrac{\partial^2 f}{\partial x_n^2}$

auf den zweimal partiell differenzierbaren reellen Funktionen heißt Laplaceoperator. Im Fall von $n=3$ ist die Rotation eines differenzierbaren Vektorfeldes f definiert durch

$$\operatorname{rot} f=\left(\frac{\partial f_3}{\partial x_2}-\frac{\partial f_2}{\partial x_3},\frac{\partial f_1}{\partial x_3}-\frac{\partial f_3}{\partial x_1},\frac{\partial f_2}{\partial x_1}-\frac{\partial f_1}{\partial x_2}\right).$$

Wenn die reelle Funktion f in $x_0 \in U$ differenzierbar ist, dann ist f auch in x_0 partiell differenzierbar und die partiellen Ableitungen sind die Richtungsableitungen in Richtung der kanonischen Einheitsvektoren $e_1, \ldots, e_n$ aus dem Beweis von Satz 9.37. Wegen der Linearität der Ableitung ist die Ableitung die lineare Abbildung:

$$f'(x_0) : \mathbb{R}^n \to \mathbb{R}, \quad (x_1, \ldots, x_n) \mapsto x_1 \frac{\partial f}{\partial x_1}(x_0) + \ldots + x_n \frac{\partial f}{\partial x_n}(x_0).$$

Wenn wir den $\mathbb{R}^n$ mit den Spaltenvektoren bezeichnen, können wir diese Abbildung durch das Matrixprodukt des Zeilenvektors ∇f mit dem Spaltenvektor x darstellen:

$$f'(x_0) : \mathbb{R}^n \to \mathbb{R}, \quad x \mapsto (\nabla f(x_0)) \cdot x.$$

Oder allgemeiner, für eine $\mathbb{R}^m$-wertige Funktion können wir die Ableitung $f'(x_0)$ von f an der Stelle x_0 als lineare Abbildung mit der Jacobimatrix identifizieren:

$$\frac{\partial f(x_0)}{\partial x} = \begin{pmatrix} \frac{\partial f_1(x_0)}{\partial x_1} & \cdots & \frac{\partial f_1(x_0)}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m(x_0)}{\partial x_1} & \cdots & \frac{\partial f_m(x_0)}{\partial x_n} \end{pmatrix} \qquad f'(x_0) : \mathbb{R}^n \to \mathbb{R}^m, \quad x \mapsto \frac{\partial f(x_0)}{\partial x} \cdot x.$$

Die lineare Abbildung ist einfach die Matrixmultiplikation der Jacobimatrix, einer $m \times n$-Matrix, mit dem Spaltenvektoren in $\mathbb{R}^n$. Insbesondere ist also die Richtungsableitung einer reellen Funktion f auf U an der Stelle $x_0 \in U$ in Richtung eines Vektors $x_1 \in \mathbb{R}^n$ das Skalarprodukt des Gradienten $\nabla f(x_0)$ von f an der Stelle x_0 mit dem Vektor x_1:

$$\frac{d}{dt} f(x_0 + tx_1)|_{t=0} = x_1 \cdot \nabla f(x_0).$$

Satz 10.9 und Satz 10.11 zeigen insbesondere, dass folgendes gilt:

Korollar 10.17. *Sei $f : U \to \mathbb{R}^m$ eine Funktion auf einer offenen Teilmenge $U \subset \mathbb{R}^n$.*

(i) *Wenn f in $x_0 \in U$ differenzierbar ist, dann existieren in x_0 alle partiellen Ableitungen $\frac{\partial f_j}{\partial x_i}(x_0)$ und setzen sich zu der Jacobimatrix zusammen.*

(ii) *Wenn f auf U stetig differenzierbar ist, dann existieren auf U die partiellen Ableitungen $\frac{\partial f_j}{\partial x_i}$ und setzen sich zusammen zu einer stetigen Funktion von U in die $m \times n$-Matrizen in $\mathbb{R}^{m \times n}$. Diese Matrizen heißen Jacobimatrizen von f.*

(iii) *Wenn f auf U partiell differenzierbar ist, und alle partiellen Ableitungen $\frac{\partial f_j}{\partial x_i}$ auf U stetig sind, dann ist f auf U stetig differenzierbar. Die Ableitung bei x_0 ist die Multiplikation der Jacobimatrix $\frac{\partial f(x_0)}{\partial x}$ mit Spaltenvektoren in $\mathbb{R}^n$.* **q.e.d.**

Definition 10.18. *Eine Nullstelle $x_0 \in U$ der Ableitung f' einer auf einer offenen Menge U reellen differenzierbaren Funktion f heißt kritischer Punkt.*

Satz 10.19. *Jedes lokale Maximum (bzw. Minimum) einer differenzierbaren reellen Funktion auf einer offenen Menge ist ein kritischer Punkt.*

Beweis: Sei x_0 ein solches lokales Maximum (bzw. Minimum). Dann ist für alle $x_1 \in X$ die entsprechende Abbildung $t \mapsto f(x_0 + tx_1)$ auf einer Umgebung von $t = 0$ differenzierbar und besitzt dort ein lokales Maximum (bzw. Minimum). Also verschwindet die entsprechende Richtungsableitung. Dann verschwindet auch $f'(x_0)$ auf allen x_1.**q.e.d.**

10.4 Höhere Ableitungen

Sei f eine auf einer offenen Teilmenge U eines Banachraumes X differenzierbare Funktion in den Banachraum Y.Wenn f zweimal differenzierbar ist, dann ist f' stetig. Die Ableitung f' ist dann eine stetige Abbildung von U nach $\mathcal{L}(X, Y)$. Die zweite Abbleitung $f''(x_0)$ ist an den Stellen $x_0 \in U$, wo sie existiert, ein Element von $\mathcal{L}(X, \mathcal{L}(X, Y))$.

Definition 10.20. *Eine Abbildung $A : V \times V \to W$ heißt bilinear, wenn für alle $v, v', v'' \in V$ und $\lambda \in \mathbb{K}$ gilt*

$$\begin{aligned} A(v + v'', v') &= A(v, v') + A(v'', v') \quad && \textit{und} \quad A(v, v' + v'') = A(v, v') + A(v, v'') \\ A(\lambda v, v') &= \lambda A(v, v') && \textit{und} \quad A(v, \lambda v') = \lambda A(v, v'). \end{aligned}$$

Das kartesische Produkt $V \times V$ von (normierten) Vektorräumen ist wieder ein normierter Vektorraum. Die bilinearen Abbildungen von $V \times V$ nach W unterscheiden sich von den linearen Abbildungen von $V \times V$ nach W. Es gibt einen anderen Vektorraum $V \otimes V$, den man das Tensorprodukt von V mit V nennt, so dass die lineare Abbildungen von $V \otimes V$ nach W genau die bilinearen Abbildungen von $V \times V$ nach W sind. Allerdings besitzt $V \otimes V$ keine natürliche Norm. Für die Dimensionen gilt

$$\dim(V \times V) = \dim(V) + \dim(V) \qquad \dim(V \otimes V) = \dim(V) \cdot \dim(V).$$

Die bilinearen Abbilungen von $V \times V$ nach W lassen sich mit den linearen Abbildungen von V in die linearen Abbildungen von V nach W identifizieren:

Lemma 10.21. *Eine Abbildung $A : V \times V \to W$ ist genau dann bilinear, wenn*

$$B : V \to \{\textit{Abbildungen } V \to W\}, \quad v \mapsto B(v), \quad B(v) : V \to W, \quad B(v)(v') = A(v, v')$$

eine lineare Abbildung von V in die linearen Abbildungen von V nach W ist. **q.e.d.**

Satz 10.22. *(Satz von Schwarz) Sei f eine differenzierbare Abbildung von einer offenen Teilmenge $U \subset X$ eines normierten Vektorraumes X in den normierten Vektorraum Y. Wenn f im Punkt x_0 zweimal differenzierbar ist, dann ist die der zweiten Ableitung entsprechende bilineare Abbildung $f''(x_0) : X \times X \to Y$ symmetrisch, d.h.*

$$(f''(x_0)x)y = (f''(x_0)y)x \quad \textit{für alle } x, y \in X.$$

Beweis: Für $t \in [0,1]$ und kleine $x, y \in X$ sei $g(t) = f(x_0 + tx + y) - f(x_0 + tx)$. Dann ist g differenzierbar mit der Ableitung mit Werten in $Y \simeq \mathcal{L}(\mathbb{R}, Y)$:

$$\begin{aligned} g'(t) &= f'(x_0 + tx + y)x - f'(x_0 + tx)x \\ &= ((f'(x_0 + tx + y) - f'(x_0)) - (f'(x_0 + tx) - f'(x_0)))x \end{aligned}$$

Weil f in x_0 zweimal differenzierbar ist, gibt es zu jedem $\epsilon > 0$ ein $\delta > 0$, so dass $B(x_0, 2\delta) \subset U$ und außerdem für $x, y \in B(0, \delta) \subset X$ und $t \in [0,1]$ die Ungleichungen

$$\begin{aligned} &\|f'(x_0 + tx + y) - f'(x_0) - f''(x_0)tx - f''(x_0)y\| \le \epsilon\|tx + y\| \le \epsilon(t\|x\| + \|y\|) \\ &\|f'(x_0 + tx) - f'(x_0) - f''(x_0)tx\| \le \epsilon\|tx\| = \epsilon t\|x\| \end{aligned}$$

gelten. Daraus folgt für $t \in [0,1]$ $\qquad \|g'(t) - (f''(x_0)y)x\| \le \epsilon\|x\|(2\|x\| + \|y\|)$.
Die Anwendung von Korollar 10.6 auf die Funktion $t \mapsto g(t) - t(f''(x_0)y)x$ ergibt dann

$$\|g(1) - g(0) - (f''(x_0)y)x\| \le \sup\{\|g'(t) - (f''(x_0)y)x\| \mid t \in [0,1]\} \le \epsilon\|x\|(2\|x\| + \|y\|).$$

Weil $g(1) - g(0) = f(x_0 + x + y) - f(x_0 + x) - f(x_0 + y) + f(x_0)$ in x und y symmetrisch ist gilt dann auch $\|g(1) - g(0) - (f''(x_0)x)y\| \le \epsilon\|y\|(2\|y\| + \|x\|)$. Daraus folgt

$$\|(f''(x_0)y)x - (f''(x_0)x)y\| \le 2\epsilon(\|x\|^2 + \|x\| \cdot \|y\| + \|y\|^2) \quad \text{für alle } x, y \in B(0, \delta).$$

Diese Ungleichung gilt wegen der Linearität nichtnur für $x, y \in B(0, \delta)$, sondern für $x, y \in X$. Im Grenzwert $\epsilon \to 0$ folgt $(f''(x_0)y)x = (f''(x_0)x)y$ für alle $x, y \in X$. **q.e.d.**

Zusammen mit Satz 10.11 erhalten wir

Korollar 10.23. *Sei $f : U \to \mathbb{R}^m$ eine zweimal stetig partiell differenzierbare Funktion von einer offenen Teilmenge $U \subset \mathbb{R}^n$ nach $\mathbb{R}^m$. Dann vertauschen die partiellen Ableitungen, d.h. für alle $i, j = 1, \ldots, n$ und $k = 1, \ldots, m$ gilt*

$$\partial_i \partial_j f_k = \frac{\partial^2 f_k}{\partial x_i \partial x_j} = \frac{\partial^2 f_k}{\partial x_j \partial x_i} = \partial_j \partial_i f_k \textit{ und } \operatorname{rot} \operatorname{grad} f = 0 \textit{ für } n = 3, m = 1. \qquad \textbf{q.e.d.}$$

Durch mehrfaches Anwenden und differenzieren erhalten wir dann auch

Korollar 10.24. *Sei f eine Abbildung von einer offenen Teilmenge U eines normierten Vektorraumes X in den normierten Vektorraum Y, die in $x_0 \in U$ n-mal differenzierbar ist. Dann ist $f^{(n)}(x_0)$ eine multilineare symmetrische Abbildung von $X \times X \times \ldots \times X$ nach Y. D.h. für jede Permutation $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ und $x_1, \ldots, x_n \in X$ gilt*

$$(\ldots((f^{(n)}(x_0)x_1)x_2)\ldots)x_n = (\ldots((f^{(n)}(x_0)x_{\sigma(1)})x_{\sigma(2)})\ldots)x_{\sigma(n)}. \qquad \textbf{q.e.d.}$$

Beispiel 10.25. *Sei $f : \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto f(x, y)$ mit $f(x, y) = xy\dfrac{x^2 - y^2}{x^2 + y^2}$ für $(x, y) \neq (0, 0)$ und $f(0, 0) = 0$. Dann ist f zweimal partiell differenzierbar.*

$$\frac{\partial f(x,y)}{\partial x} = y\frac{(x^2 - y^2)(x^2 + y^2) + 2x^2(x^2 + y^2) - 2x^2(x^2 - y^2)}{(x^2 + y^2)^2} = y\frac{x^4 + 4x^2y^2 - y^4}{(x^2 + y^2)^2}$$

$$\frac{\partial f(x,y)}{\partial y} = x\frac{(x^2 - y^2)(x^2 + y^2) - 2y^2(x^2 + y^2) - 2y^2(x^2 - y^2)}{(x^2 + y^2)^2} = x\frac{x^4 - 4x^2y^2 - y^4}{(x^2 + y^2)^2}.$$

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

Also existieren auf $\mathbb{R}^2$ alle zweiten partiellen Ableitungen, mit $\dfrac{\partial^2 f}{\partial y \partial x}(0, 0) \neq \dfrac{\partial^2 f}{\partial x \partial y}(0, 0)$.

Definition 10.26. *Sei $f : U \to \mathbb{R}$ eine reelle Funktion auf einer offenenTeilmenge U des normierten Vektorraumes X, die bei $x_0 \in U$ zweimal differenzierbar ist. Dann definiert die zweite Ableitung eine symmetrische Bilinearform auf X:*

$$f''(x_0) : X \times X \to \mathbb{R}, \qquad (x, y) \mapsto f''(x_0)(x, y).$$

Für $X = \mathbb{R}^n$ identifizieren wir die Elemente von X wieder mit den Spaltenvektoren. Dann ist diese Bilinearform durch die sogenannte Hessematrix gegeben:

$$f''(x_0)(x, y) = y^t \cdot \frac{\partial^2 f(x_0)}{\partial x^2} \cdot x \; = \sum_{i,j=1}^{n} x_j \frac{\partial^2 f(x_0)}{\partial x_j \partial x_i} y_i.$$

Satz 10.27. *Sei $f : U \to \mathbb{R}$ eine auf einer offenen Menge zweimal differenzierbare reelle Funktion f. Dann ist die zweite Ableitung bei allen lokalen Minima (Maxima) eine nicht negative (nicht positive) Bilinearform: $f''(x_0)(x, x) \geq 0$ bzw. ≤ 0 für alle $x \in X$. Gibt es umgekehrt einen kritischen Punkt $x_0 \in U$ und ein $\epsilon > 0$ mit*

$$f''(x_0)(x, x) \geq \epsilon\|x\|^2 \quad \text{bzw.} \quad f''(x_0)(x, x) \leq -\epsilon\|x\|^2 \quad \text{für alle } x \in X,$$

dann ist der kritische Punkt ein striktes lokales Minimum bzw. Maximum.

Beweis: Wenn x_0 ein lokales Maximum bzw. Minimum von f ist, dann für alle $x \in X$ auch $t = 0$ von $t \mapsto f(x_0 + tx)$. Deshalb folgt die erste Aussage aus Korollar 7.17.

Umgekehrt folgt aus der zweimaligen Differenzierbarkeit, dass für ein $\delta > 0$

$$-\tfrac{\epsilon}{2}\|x\|^2 < f'(x_0 + x)x - f'(x_0)x - f''(x_0)(x, x) < \tfrac{\epsilon}{2}\|x\|^2$$

für alle $x \in B(0, \delta)$ gilt. Daraus und den obigen Bedingungen folgt für die gleichen x

$$f'(x_0+x)x = f'(x_0+x)x - f'(x_0)x - f''(x_0)(x, x) + f''(x_0)(x, x) > \tfrac{\epsilon}{2}\|x\|^2 \text{ bzw. } < -\tfrac{\epsilon}{2}\|x\|^2.$$

Also ist $f(x_0+x)-f(x_0)=\int_0^1 f'(x_0+tx)xdt > \frac{\epsilon}{4}\|x\|^2$ bzw. $< -\frac{\epsilon}{4}\|x\|^2$. **q.e.d.**

Auf endlichdimensionalen Räumen zeigt Beweis von Satz 9.37 wie $f''(x_0)(x,x) \geq \epsilon\|x\|^2$ aus $f''(x_0)(x,x) > 0$ für alle $x \in X\setminus\{0\}$ folgt. In unendlichdimensionalen Räumen gilt das nicht, und diese Bedingung ist nicht hinreichend für ein lokales Minimum.

Beispiel 10.28. *Sei $f: C([0,1]) \to \mathbb{R},\ x \mapsto \int_0^1 x^2(t)(t-x(t))dt$. Dann ist $f''(0)(x,x) = 2\int_0^1 x^2(t)tdt > 0$ für alle $x \in C([0,1]) \setminus \{0\}$. Sei $x_\epsilon(t) = \begin{cases} \epsilon - t & \text{für } 0 \leq t < \epsilon \\ 0 & \text{für } \epsilon \leq t \leq 1 \end{cases}$ mit $\epsilon \in (0,1)$. Dann gilt $\|x_\epsilon\|_\infty = \epsilon$ und $f(sx_\epsilon) = s^2\int_0^\epsilon(\epsilon-t)^2tdt - s^3\int_0^\epsilon(\epsilon-t)^3dt = -\frac{s^2}{3}(\epsilon-t)^3t]_0^\epsilon - (\frac{s^2}{12}-\frac{s^3}{4})(\epsilon-t)^4]_0^\epsilon = (\frac{s^2}{12}-\frac{s^3}{4})\epsilon^4$. Also ist $x=0$ kein lokales Minimum.*

Zum Abschluss wollen wir das Taylorpolynom und die Taylorreihe einer Funktion $f: U \to Y$ auf einer offenen konvexen Teilmengen $U \subset X$ eines normierten Vektorraumes X in einem normierten Vektorraum Y betrachten. Für $x_0, x \in U$ sei

$$g: [0,1] \to \mathbb{R}, \quad t \mapsto f(x_0 + t(x-x_0))$$

Wenn f auf U n-mal differenzierbar ist, dann ist auch g n-mal differenzierbar. Wegen der Kettenregel Satz 10.4 (iii) ist die m-te Ableitung von g gleich

$$g^{(m)}(t) = (\ldots(f^{(m)}(x_0+t(x-x_0))(x-x_0))\ldots(x-x_0)),$$

also die m-lineare symmetrische Form zu $f^{(m)}(x_0+t(x-x_0))$ ausgewertet auf $((x-x_0),\ldots,(x-x_0)) \in X^{\times m}$. Dann erhalten wir für n-mal differenzierbare bzw. glatte Funktionen das Taylorpolynom der Ordning n von f bei x_0 bzw. die Taylorreihe:

$$x \mapsto \sum_{k=0}^{n} \frac{f^{(k)}(x_0)((x-x_0),\ldots,(x-x_0))}{k!} \text{ bzw. } x \mapsto \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)((x-x_0),\ldots,(x-x_0))}{k!}.$$

Eine unendlich oft differenzierbare Funktion f heißt wieder reell analytisch in x_0, wenn die entsprechende Taylorreihe auf einer Umgebung von x_0 gegen $f(x)$ konvergiert. Auf einem Banachraum X ist z.B. $\exp: \mathcal{L}(X) \to \mathcal{L}(X)$ eine analytische Funktion.

Für reelle Funktionen, also $Y = \mathbb{R}$ ergibt der Satz von Taylor:

Satz 10.29. *(von Taylor in höheren Dimensionen) Sei $f: U \to \mathbb{R}$ eine auf einer offenen konvexen Teilmenge eines normierten Vektorraumes definierte $(n+1)$-mal differenzierbare Funktion. Dann gibt es für jedes $x, x_0 \in U$ ein $\xi \in (0,1)$, so dass*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)((x-x_0),\ldots,(x-x_0))}{k!} + \frac{f^{(n+1)}(x_0+\xi(x-x_0))((x-x_0),\ldots,(x-x_0))}{(n+1)!}$$

gilt. Hierbei bezeichnen wir mit $f^{(k)}(x_0)$ bzw. $f^{(m+1)}(x_0+\xi(x-x_0))$ die entsprechende multilineare Abbildung von $X^{\times k}$ bzw. $X^{\times(n+1)}$ nach $\mathbb{R}$. Der erste Term heißt wieder Taylorpolynom von f in x_0 der Ordnung n und der zweite Term Restglied **q.e.d.**