Analysis IV: Einführung in die Funktionalanalysis und Lebesgueintegration

Ausarbeitung einer Vorlesung

vom Sommersemester 1992

JOACHIM WEIDMANN

Fachbereich Mathematik
der Universität Frankfurt

Überarbeitet im Sommersemester 2003
Stand: 21. Oktober 2003

Inhaltsverzeichnis

1 Metrische Räume und Topologie

1.1 Metrische Räume

Sei X eine Menge. Eine Abbildung $d : X \times X \to [0, \infty)$ heißt eine *Metrik* (Abstandsfunktion) auf X, wenn gilt

(M1) $d(x, y) = 0 \Leftrightarrow x = y$ (***Positivität***),

(M2) $d(x, y) = d(y, x)$ für alle $x, y \in X$ (***Symmetrie***),

(M3) $d(x, z) \leq d(x, y) + d(y, z)$ für alle $x, y, z \in X$ (***Dreiecksungleichung***).

(X, d), die Menge X versehen mit einer Metrik d, heißt ein *metrischer Raum*.

Beispiel 1.1 Sei $X = \mathbb{R}^m$ oder X eine Teilmenge von $\mathbb{R}^m$ $(m \in \mathbb{N})$.

$$d_p(x, y) = \begin{cases} \left\{ \sum\limits_{j=1}^{m} |x_j - y_j|^p \right\}^{1/p} & \text{für } 1 \leq p < \infty, \\ \max\{|x_j - y_j| : j = 1, \ldots, m\} & \text{für } p = \infty. \end{cases}$$

Es ist offensichtlich, daß die d_p die Eigenschaften (M1) und (M2) erfüllen. Für $p = 1$, 2 und ∞ ist auch (M3) leicht zu beweisen:

$\mathbf{p = 1}$: $d_1(x, z) = \sum_{j=1}^{m} |x_j - z_j| \leq \sum_{j=1}^{m} \{|x_j - y_j| + |y_j - z_j|\} = d_1(x, y) + d_1(y, z)$.

$\mathbf{p = 2}$: Es gilt $d_2(x, y) = |x - y|$, wobei $|\cdot|$ die euklidische Norm in $\mathbb{R}^m$ ist; mit der Dreiecksungleichung für $|\cdot|$ folgt

$$d_2(x, z) = |x - z| \leq |x - y| + |y - z| = d_2(x, y) + d_2(y, z).$$

$$\begin{aligned} \mathbf{p = \infty}: \ d_\infty(x, z) &= \max\{|x_j - z_j| : j = 1, \ldots, m\} \\ &\leq \max\{|x_j - y_j| + |y_j - z_j| : j = 1, \ldots, m\} \\ &\leq \max\{|x_j - y_j| : j = 1, \ldots, m\} + \max\{|y_j - z_j| : j = 1, \ldots, m\} \\ &= d_\infty(x, y) + d_\infty(y, z). \end{aligned}$$

Den Beweis von (M3) für $1 < p < 2$ und $2 < p < \infty$ wollen wir hier übergehen; er wird in § 4 (im Zusammenhang mit der Untersuchung der ℓ_p–Räume) nachgeholt, vgl. Aufgabe 4.9. $\square$

Beispiel 1.2 Sei X eine beliebige Menge. Durch

$$d(x,y) = \begin{cases} 0 & \text{für } x = y, \\ 1 & \text{für } x \neq y \end{cases}$$

ist eine Metrik auf X definiert, die *diskrete Metrik*. □

Satz 1.3 a) *Ist d eine Metrik auf X, $f : [0,\infty) \to [0,\infty)$ zweimal stetig differenzierbar mit*

$$f(0) = 0, f'(s) > 0 \quad \textit{und} \quad f''(s) \leq 0 \quad \textit{für} \quad s > 0$$

(es genügt auch: $f(0) = 0,\ f'(s) > 0)$ und f' fallend für alle $s > 0$), so ist

$$e(x,y) := f(d(x,y))$$

ebenfalls eine Metrik auf X.

b) *Ist d eine Metrik auf X, so ist auch*

$$e(x,y) := \frac{d(x,y)}{1 + d(x,y)}$$

eine Metrik auf X (man beachte, daß gilt $e(x.y) < 1$ für alle $x, y \in X$).

Beweis. a) (M1): Wegen $f(s) > 0$ für $s > 0$ gilt $e(x,y) = 0$ genau dann, wenn $d(x,y) = 0$ gilt, also genau dann, wenn $x = y$ gilt.

(M2) ist offensichtlich.

(M3) Mit Hilfe der Eigenschaften von d und f folgt

$$\begin{aligned} e(x,z) &= f(d(x,z)) \leq f(d(x,y) + d(y,z)) \\ &= f(d(x,y)) + \int\limits_{d(x,y)}^{d(x,y)+d(y,z)} f'(t)\,\mathrm{d}t \leq f(d(x,y)) + \int\limits_{0}^{d(y,z)} f'(t)\,\mathrm{d}t \\ &= f(d(x,y)) + f(d(y,z)) = e(x,y) + e(y,z). \end{aligned}$$

b) folgt offenbar aus Teil a. ■

1.2 Topologie

Sei X eine Menge. Eine *Topologie* $\mathcal{T}$ auf X ist eine Familie von Teilmengen von X mit den Eigenschaften

(T1) $\emptyset \in \mathcal{T}$ und $X \in \mathcal{T}$,

(T2) der Durchschnitt von *endlich vielen* Mengen aus $\mathcal{T}$ ist wieder aus $\mathcal{T}$,

(T3) die Vereinigung von *beliebig vielen* Mengen aus $\mathcal{T}$ ist wieder aus $\mathcal{T}$.

Die Mengen aus $\mathcal{T}$ werden als die, bezüglich der Topologie $\mathcal{T}$, *offenen Mengen* bezeichnet. $(X;\mathcal{T})$, die Menge X versehen mit der Topologie $\mathcal{T}$, heißt ein *topologischer Raum*.

Besonders einfach ist es, mit Hilfe einer Metrik eine Topologie zu erzeugen. Dazu definieren wir in einem metrischen Raum (X, d) zunächst die Mengen

$$K(x,\varepsilon) = K_d(x,\varepsilon) := \{y \in X : d(x,y) < \varepsilon\},$$

$$\overline{K}(x,\varepsilon) = \overline{K}_d(x,\varepsilon) := \{y \in X : d(x,y) \leq \varepsilon\};$$

sie heißen die „*offenen*“ bzw. „*abgeschlossenen*“ *ε–Kugeln* um x (wir werden sehen, daß diese Kugeln bezüglich der durch die Metrik d erzeugten Topologie tatsächlich offen bzw. abgeschlossen sind).

Die *durch die Metrik d erzeugte Topologie* $\mathcal{T}_d$ wird definiert durch:

$$A \in \mathcal{T}_d \Leftrightarrow \text{ zu jedem } x \in A \text{ existiert ein } \varepsilon > 0 \text{ mit } K(x,\varepsilon) \subset A.$$

Wegen $\overline{K}(x,\delta) \subset K(x,\varepsilon)$ für $\delta < \varepsilon$ kann in dieser Definition $K(x,\varepsilon)$ durch $\overline{K}(x,\varepsilon)$ ersetzt werden.

Sei $\varepsilon > 0$. Für jedes $y \in K(x,\varepsilon)$ und jedes $\eta \leq \varepsilon - d(x,y)$ gilt $K(y,\eta) \subset K(x,\varepsilon)$. Also ist $K(x,\varepsilon)$ aus $\mathcal{T}_d$, d. h. $K(x,\varepsilon)$ ist tatsächlich offen bezüglich $\mathcal{T}_d$. Dies gilt auch für $\varepsilon = 0$, da $K(x,0) = \emptyset$ gilt.

Zwei Metriken d und e auf X heißen *äquivalent*, wenn sie die gleiche Topologie erzeugen. Sie heißen *gleichmäßig äquivalent* (bzw. *uniform äquivalent*), wenn ein $c > 0$ existiert mit

$$\frac{1}{c}\, d(x,y) \leq e(x,y) \leq c\, d(x,y) \text{ für alle } x, y \in X.$$

Satz 1.4 a) *Sei d eine Metrik auf X, e eine gemäß Satz 1.3 erzeugte Metrik (d. h. $e = f(d)$, z. B. $e = d/(1+d)$), so sind d und e* äquivalent,

b) *Die Metriken d_p $(1 \leq p \leq \infty)$ auf $\mathbb{R}^m$ (vgl. Beispiel 1.1) sind äquivalent.*

Beweis. a) Dies folgt unmittelbar aus der Definition von $\mathcal{T}_d$ und $\mathcal{T}_e$, wenn man beachtet

$$K_d(x,\delta) = K_e(x,\varepsilon) \text{ für } \varepsilon = f(\delta) \text{ bzw. } \delta = f^{-1}(\varepsilon).$$

b) Es gilt $d_\infty \leq d_p \leq m^{1/p} d_\infty$ für $1 \leq p < \infty$ und somit

$$K_{d_\infty}(x, m^{-1/p}\delta) \subset K_{d_p}(x,\delta) \subset K_{d_\infty}(x,\delta).$$

■

Sei $(X, \mathcal{T})$ ein topologischer Raum, $x \in X$. Eine Teilmenge U von X heißt eine *Umgebung* von x, wenn eine offene Menge A (d. h. $A \in \mathcal{T}$) existiert mit $x \in A \subset U$.

Speziell ist in einem metrischen Raum (X, d) die Menge U genau dann eine Umgebung von x, wenn ein $\varepsilon > 0$ existiert mit $K_d(x, \varepsilon) \subset U$.

Satz 1.5 *Sei $(X, \mathcal{T})$ ein topologischer Raum. Eine Teilmenge A von X ist genau dann offen (d. h. $A \in \mathcal{T}$) wenn A Umgebung jedes Punktes von A ist.*

Beweis. Nach obiger Definition der Umgebungen eines Punktes ist klar, daß jede offene Menge Umgebung jedes ihrer Punkte ist. — Sei umgekehrt A Umgebung jedes Punktes von A, B die Vereinigung aller offenen Teilmengen von A; B ist also selbst offen und $B \subset A$. Da zu jedem $x \in A$ eine offene Menge A_x existiert mit $x \in A_x \subset A$, folgt andererseits $A \subset B$. Also ist $A = B$ offen. ■

Ein topologischer Raum $(X, \mathcal{T})$ heißt *separiert* oder *hausdorffsch*, wenn zu je zwei Punkten $x \neq y$ aus X disjunkte Umgebungen U_x und U_y von x bzw. y existieren. (Man sagt dann auch: $(X, \mathcal{T})$ erfüllt das *zweite Trennungsaxiom*. Das *erste Trennungaxiom* fordert: Zu je zwei Punkten $x \neq y$ aus X existiert eine Umgebung U_x von x, die y nicht enthält. In dem hier besonders interessierenden Fall der topologischen Vektorräume sind diese beiden Trennungsaxiome äquivalent, vgl. z. B. Beweis von Satz 2.5).

Offensichtlich ist jeder metrische Raum (X, d) separiert: Mit $\varepsilon := \frac{1}{2}d(x, y)$ gilt $K(x, \varepsilon) \cap K(y, \varepsilon) = \emptyset$.

Sei $(X, \mathcal{T})$ ein topologischer Raum, $Y \subset X$. Dann ist $(Y, \mathcal{T}_Y)$ mit

$$\mathcal{T}_Y = \Big\{ B \subset Y : \exists A \in \mathcal{T} \quad \text{mit} \quad B = A \cap Y \Big\}$$

ebenfalls ein topologischer Raum. Die Topologie $\mathcal{T}_Y$ heißt die durch $\mathcal{T}$ auf Y *induzierte Topologie*.

Sind $\mathcal{T}_1$ und $\mathcal{T}_2$ Topologien auf X, so heißt $\mathcal{T}_1$ *feiner* als $\mathcal{T}_2$ (oder $\mathcal{T}_2$ *gröber* als $\mathcal{T}_1$), wenn gilt $\mathcal{T}_1 \supset \mathcal{T}_2$.

Auf jeder Menge X gibt es eine *feinste* und eine *gröbste* Topologie:

(i) die feinste Topologie $\mathcal{T}_f$ besteht aus allen Teilmengen von X, d. h. jede Teilmenge von X ist offen, jede Menge $U \subset X$ mit $x \in U$ ist eine Umgebung von x,

(ii) die gröbste Topologie $\mathcal{T}_g$ besteht nur aus $\emptyset$ und X, d. h. nur $\emptyset$ und X sind offen, die einzige Umgebung von x ist X.

Es ist leicht zu sehen, daß die feinste Topologie durch die diskrete Metrik (Beispiel 1.2) erzeugt wird; sie heißt auch die *diskrete Topologie*.

1.3 Stetige Abbildungen

Seien $(X, \mathcal{T}_X)$ und $(Y, \mathcal{T}_Y)$ topologische Räume. Eine Abbildung $f : X \to Y$ heißt *stetig*, wenn das Urbild jeder offenen Menge aus Y eine offene Menge in X ist, d. h.

$$B \text{ offen in } Y \implies f^{-1}(B) := \{x \in X : f(x) \in B\} \text{ offen in } X.$$

Nach dieser Definition ist die Stetigkeit eine globale Eigenschaft; der folgende Satz erlaubt es, die Stetigkeit lokal zu untersuchen:

Satz 1.6 *Seien $(X, \mathcal{T}_X)$ und $(Y, \mathcal{T}_Y)$ topologische Räume. Eine Abbildung $f : X \to Y$ ist genau dann stetig, wenn für jedes $x \in X$ gilt: zu jeder Umgebung V von $f(x)$ existiert eine Umgebung U von x mit $f(U) \subset V$. (Diese letztere Eigenschaft nennt man* Stetigkeit im Punkt *x).*

Beweis. $\Rightarrow$: Sei V eine Umgebung von $f(x)$. Dann existiert eine offene Menge B in Y mit $f(x) \in B \subset V$. Da f stetig ist, ist auch $f^{-1}(B)$ offen in X und $x \in f^{-1}(B)$. Also ist $U := f^{-1}(B)$ eine Umgebung von x mit $f(U) \subset V$.

$\Leftarrow$: Sei B offen in Y. Ist $x \in f^{-1}(B)$, so ist B eine Umgebung von $f(x)$. Also existiert eine Umgebung U von x mit $f(U) \subset B$, und somit $x \in U \subset f^{-1}(B)$; d. h. $f^{-1}(B)$ ist Umgebung von x. Da dies für alle $x \in f^{-1}(B)$ gilt, ist $f^{-1}(B)$ offen. ■

Offensichtlich ist eine Topologie $\mathcal{T}_1$ auf X genau dann *feiner* als eine Topologie $\mathcal{T}_2$ auf X, wenn id : $(X, \mathcal{T}_1) \to (X, \mathcal{T}_2)$ stetig ist. Die induzierte Topologie auf einer Teilmenge A von $(X, \mathcal{T})$ ist die *gröbste* Topologie, für die die Einbettung $A \to (X, \mathcal{T})$ stetig ist.

Sei $(X, \mathcal{T})$ ein topologischer Raum. Die Folge (x_n) aus X *konvergiert gegen* $x \in X, x_n \to x$, wenn für jede Umgebung U von x ein $N \in \mathbb{N}$ existiert mit $x_n \in U$ für $n \geq N$. (Nur wenn dies nötig erscheint, sagt man „konvergiert bezüglich $\mathcal{T}$“ und schreibt „$x_n \xrightarrow{\mathcal{T}} x$“.)

Sei $(X, \mathcal{T})$ ein topologischer Raum, $x \in X$. Die *Familie aller Umgebungen* von x bezeichnen wir mit $\mathcal{U}_x$. Eine Familie $\mathcal{B}_x$ von Umgebungen von x heißt eine *Umgebungsbasis* von x, wenn zu jedem $U \in \mathcal{U}_x$ ein $B \in \mathcal{B}_x$ existiert mit $B \subset U$. Natürlich ist $\mathcal{U}_x$ selbst eine Umgebungsbasis.

Der Raum $(X, \mathcal{T})$ erfüllt das *erste Abzählbarkeitsaxiom*, wenn jeder Punkt eine abzählbare Umgebungsbasis besitzt.

Jeder metrische Raum (X, d) erfüllt das erste Abzählbarkeitsaxiom: Die Familie $\{K_d(x, 1/n) : n \in \mathbb{N}\}$ ist eine abzählbare Umgebungsbasis von x.

Eine Abbildung $f : (X, \mathcal{T}_X) \to (X, \mathcal{T}_Y)$ heißt *folgenstetig*, wenn gilt

$$x_n \to x \quad \text{in} \quad (X, \mathcal{T}_X) \Longrightarrow f(x_n) \to f(x) \quad \text{in} \quad (Y, \mathcal{T}_{\mathcal{Y}}).$$

Satz 1.7 *Seien $(X, \mathcal{T}_X)$ und $(Y, \mathcal{T}_Y)$ topologische Räume, $f : X \to Y$.*

a) *Ist f stetig, so ist f folgenstetig.*

b) *Erfüllt $(x, \mathcal{T}_X)$ das erste Abzählbarkeitsaxiom, so folgt umgekehrt die Stetigkeit von f aus der Folgenstetigkeit.*

Beweis. a) Sei $x_n \to x$ in X, V eine Umgebung von $f(x)$. Dann existiert eine Umgebung U von x mit $f(U) \subset V$. Wegen $x_n \to x$ existiert ein $N \in \mathbb{N}$ mit $x_n \in U$ für $n \geq N$, also $f(x_n) \in V$ für $n \geq N$, d. h. $f(x_n) \to f(x)$.

b) Wir nehmen an, daß f *nicht* stetig ist. Dann existiert ein $x \in X$ und eine Umgebung V von $f(x)$ so, daß für jede Umgebung U von x gilt $f(U) \not\subset V$. Sei $\{U_n : n \in \mathbb{N}\}$ eine abzählbare Umgebungsbasis von x,

$$\tilde{U}_n := \bigcap_{j=1}^{n} U_j.$$

Dann ist auch $\{\tilde{U}_n :\in \mathbb{N}\}$ eine Umgebungsbasis von x, und es gilt $\tilde{U}_n \supset \tilde{U}_{n+1}$. Wegen $f(\tilde{U}_n) \not\subset V$ existiert für jedes $n \in \mathbb{N}$ ein $x_n \in \tilde{U}_n$ mit $f(x_n) \notin V$, d. h. $f(x_n) \not\to f(x)$. Andererseits existiert zu jeder Umgebung U von x ein $N \in \mathbb{N}$ mit

$$\tilde{U}_n \subset \tilde{U}_N \subset U \qquad \text{für alle} \quad n \geq N,$$

also $x_n \in U$ für $n \geq N, x_n \to x$. Somit ist f *nicht* folgenstetig. ■

In einem metrischen Raum (X, d) gilt offenbar $x_n \to x$ genau dann, wenn $d(x_n, x) \to 0$ gilt. Damit folgt das

Korollar 1.8 *Sind (X, d) und (Y, e) metrische Räume, so sind folgende Aussagen äquivalent:*

(i) $f : X \to Y$ *ist stetig,*

(ii) $\forall x \in X \quad \forall \varepsilon > 0 \quad \exists \delta = \delta(x, \varepsilon) > 0$ *so, daß gilt:* $d(x, y) < \delta \Rightarrow e(f(x), f(y)) < \varepsilon$,

(iii) $d(x_n, x) \to 0 \ \Rightarrow \ e(f(x_n), f(x)) \to 0.$

Satz 1.9 *Sei (X, d) ein metrischer Raum, $y \in X$, $A \subset X$, $A \neq \emptyset$. Dann sind die Funktionen*

$$\begin{aligned} f_y(x) &:= d(x, y), \\ f_A(x) &:= d(x, A) := \inf\{d(x, y) : y \in A\} \end{aligned}$$

stetig.

Beweis. Wir zeigen tatsächlich

$$|f_y(x) - f_y(z)| \leq d(x, z), \quad |f_A(x) - f_A(z)| \leq d(x, z).$$

Die erste Behauptung ist ein Spezialfall der zweiten (kann aber mit Hilfe der Dreiecksungleichung leicht direkt bewiesen werden, Übung).

Für jedes $\varepsilon > 0$ existiert ein $y_\varepsilon \in A$ mit

$$d(x, y_\varepsilon) \leq d(x, A) + \varepsilon.$$

Also gilt

$$d(z, A) \leq d(z, y_\varepsilon) \leq d(z, x) + d(x, y_\varepsilon) \leq (d(z, x) + d(x, A)) + \varepsilon.$$

Da dies für jedes $\varepsilon > 0$ gilt, folgt

$$d(z, A) \leq d(x, A) + d(z, x) \quad \text{also} \quad d(z, A) - d(x, A) \leq d(z, x).$$

Entsprechend gilt natürlich

$$d(x, A) \leq d(z, A) + d(z, x) \quad \text{also} \quad d(x, A) - d(z, A) \leq d(z, x),$$

und somit

$$|d(z, A) - d(x, A)| \leq d(z, x).$$

■

1.4 Abgeschlossene Mengen, abgeschlossene Hülle

Sei $(X, \mathcal{T})$ ein topologischer Raum. Eine Teilmenge A von X heißt *abgeschlossen*, wenn ihr Komplement $\complement A = X \backslash A$ offen ist. Die Mengen $\emptyset$ und X sind also (unabhängig von der Topologie) stets offen *und* abgeschlossen.

Die *abgeschlossene Hülle* (auch *Abschließung*) $\overline{A}$ einer Teilmenge A von X ist der Durchschnitt aller abgeschlossenen Teilmengen B von X, die A enthalten. $\overline{A}$ ist abgeschlossen, denn

$$\begin{aligned}
\complement \overline{A} &= \complement \Big\{ \cap \{B : A \subset B,\ B \text{ abgeschlossen}\} \Big\} \\
&= \cup \Big\{ \complement \{B : A \subset B,\ B \text{ abgeschlossen}\} \Big\} \\
&= \cup \Big\{ C : \complement A \supset C,\ C \text{ offen} \Big\}
\end{aligned}$$

ist offen (als Vereinigung von offenen Mengen). Insbesondere ist die abgeschlossene Hülle von A die kleinste abgeschlossene Menge, die A enthält.

Satz 1.10 *Sei (X, d) ein metrischer Raum.*

a) *$A \subset X$ ist genau dann abgeschlossen, wenn A die Grenzwerte aller in X konvergenten Folgen aus A enthält.*

b) *$\overline{A}$ ist die Menge der Grenzwerte der in X konvergenten Folgen aus A.*

Beweis. a) $\Rightarrow$: Sei A abgeschlossen, (x_n) aus A, $x_n \to x$. Wir nehmen an: $x \notin A$. Da $X \backslash A$ offen ist, existiert eine Umgebung U von x mit $U \subset X \backslash A$. Wegen $x_n \to x$ gilt dann $x_n \in U \subset X \backslash A$ für große n; das ist ein Widerspruch zu $x_n \in A$.

$\Leftarrow$: A enthalte die Grenzwerte aller in X konvergenten Folgen aus A. Wir nehmen an, daß A *nicht* abgeschlossen ist, also $X \backslash A$ *nicht* offen. Dann existiert ein $x \in X \backslash A$ so, daß für jede Umgebung U von x gilt $U \not\subset X \backslash A$, d. h. $U \cap A \neq \emptyset$. Also existiert für jedes $n \in \mathbb{N}$ ein $x_n \in K(x, 1/n) \cap A$. Für diese Folge gilt $x_n \in A$ und $x_n \to x \in X \backslash A$ im Widerspruch zur Voraussetzung.

b) Da $\overline{A}$ eine abgeschlossene Menge ist mit $\overline{A} \supset A$, enthält $\overline{A}$ nach Teil a) die Grenzwerte aller in X konvergenten Folgen aus A. Sei nun $x \in \overline{A}$. Wäre x nicht Grenzwert einer Folge (x_n) aus A, so gäbe es ein $\varepsilon > 0$ mit $K(x, \varepsilon) \cap A = \emptyset$. Also wäre A enthalten in der abgeschlossenen Menge $X \backslash K(x, \varepsilon)$; damit gilt dann auch $\overline{A} \subset X \backslash K(x, \varepsilon)$, im Widerspruch zur $x \in \overline{A}$. ■

Korollar 1.11 *In einem metrischen Raum (X, d) ist $\overline{K}(x, \varepsilon)$ abgeschlossen.*

Beweis. Ist (x_n) eine Folge aus $\overline{K}(x, \varepsilon)$ mit $x_n \to y \in X$. Dann gilt (vgl. Satz 1.9)

$$d(x, y) = \lim_{n\to\infty} d(x, x_n) \leq \varepsilon,$$

also $y \in \overline{K}(x, \varepsilon)$. Nach Satz 1.10 a ist also $\overline{K}(x, \varepsilon)$ abgeschlossen. (Wir überlassen es dem Leser, alternativ hierzu die Offenheit von $\complement\,\overline{K}(x, \varepsilon)$ zu beweisen.) ■

Bemerkung 1.12 *In den meisten metrischen Räumen gilt zwar $\overline{K(x, \varepsilon)} = \overline{K}(x, \varepsilon)$ für jedes $\varepsilon > 0$. Im allgemeinen ist dies jedoch falsch. Betrachten wir z. B. eine aus mindestens zwei Punkten bestehende Menge X mit der diskreten Metrik d (vgl. Beispiel 1.2), so gilt für jedes $x \in X$*

$$\overline{K(x, 1)} = K(x, 1) = \{x\} \neq X = \overline{K}(x, 1).$$

1.5 Kompakte Mengen

Ein separierter topologischer Raum $(X, \mathcal{T})$ heißt *kompakt*, wenn jede offene Überdeckung $\mathcal{U}$ von X eine endliche Überdeckung enthält. Eine Teilmenge K von X heißt *kompakt*, wenn K als topologischer Raum mit der durch $\mathcal{T}$ induzierten Topologie kompakt ist, d. h. wenn jede (in X) offene Überdeckung $\mathcal{U}$ von K eine endliche Überdeckung enthält.

Jede endliche Menge ist offensichtlich kompakt.

Satz 1.13 *Seien $(X, \mathcal{T}_X)$ und $(Y, \mathcal{T}_Y)$ separierte Topologische Räume, $(X, \mathcal{T}_X)$ kompakt und $f : X \to Y$ stetig. Dann ist $f(X)$ kompakt in $(Y, \mathcal{T}_Y)$.*

Beweis. Sei $\mathcal{U}$ eine offene Überdeckung von $f(X)$ in Y. Da f stetig ist, ist $f^{-1}(U)$ offen für jedes $U \in \mathcal{U}$, d. h. $\{f^{-1}(U) : U \in \mathcal{U}\}$ ist eine offene Überdeckung von X. Da X kompakt ist, existieren $U_1, \ldots, U_n \in \mathcal{U}$ mit

$$X \subset \bigcup_{j=1}^{N} f^{-1}(U_j), \quad \text{also} \quad f(X) \subset \bigcup_{j=1}^{n} U_j.$$

Also enthält $\mathcal{U}$ eine endliche Überdeckung, d. h. $f(X)$ ist kompakt. ■

Ein topologischer Raum $(X, \mathcal{T})$ heißt *folgenkompakt*, wenn jede Folge (x_n) aus X eine in X konvergente Teilfolge enthält. Ein metrischer Raum (X, d) heißt *total beschränkt*, wenn zu jedem $\varepsilon > 0$ endlich viele $x_1, \ldots, x_n$ existieren mit $X = \cup_{j=1}^{n} K(x_j, \varepsilon)$.

Satz 1.14 *Sei (X, d) ein metrischer Raum.*

a) *X ist genau dann kompakt, wenn X folgenkompakt ist.*

b) *Ist X kompakt, so ist X total beschränkt.*

Beweis. a) $\Rightarrow$: Sei X *nicht* folgenkompakt, d. h. es gibt eine Folge (x_n) aus X, die keine konvergente Teilfolge enthält, d. h. für jedes $x \in X$ existiert ein $\varepsilon(x) > 0$ mit

$$K(x, \varepsilon(x)) \cap \{x_n : n \in \mathbb{N}\} \quad \text{ist endlich.}$$

$\{K(x, \varepsilon(x)) : x \in X\}$ ist eine offene Überdeckung von X, die aber keine endliche Überdeckung enthalten kann, da sonst $\{x_n : N \in \mathbb{N}\}$ endlich wäre. Also ist X *nicht* kompakt.

$\Leftarrow$: Sei X folgenkompakt, $\mathcal{U}$ eine offene Überdeckung von X.

(α) Es existiert ein $\delta > 0$ so, daß für jedes $x \in X$ ein $U_x \in \mathcal{U}$ existiert mit $K(x, \delta) \subset U_x$.

Beweis von (α): Wir nehmen an, daß kein solches δ existiert. Dann gibt es für jedes $n \in \mathbb{N}$ $(\delta = 1/n)$ ein $x_n \in X$ mit

$$K\Big(x_n, \frac{1}{n}\Big) \not\subset U \quad \text{für } \textit{jedes} \quad U \in \mathcal{U}.$$

Da X folgenkompakt ist, existiert eine Teilfolge (x_{n_k}) mit $x_{n_k} \to x \in X$. Es gilt dann $x \in U$ für ein geeignetes $U \in \mathcal{U}$. Da U offen ist existiert ein $\varepsilon > 0$ mit $K(x, \varepsilon) \subset U$. Wegen $x_{n_k} \to x$ ist $x_{n_k} \in K(x, \frac{\varepsilon}{2})$ für $k \geq k_1$ und somit $K(x_{n_k}, n_k^{-1}) \subset U$ für $k \geq k_2$. Das ist ein Widerspruch zu unserer Annahme.

(β) Zu jedem $\delta > 0$ existieren endlich viele $x_1, \ldots, x_n$ mit $\bigcup_{j=1}^{n} K(x_j, \delta) = X$.

Beweis von (β): Wir nehmen an, daß dies *nicht* gilt. Dann kann man induktiv eine Folge (x_n) aus X definieren mit

$$d(x_j, x_{n+1}) > \delta \quad \text{für} \quad j = 1, \ldots, n,$$

für alle $n \in \mathbb{N}$. Diese Folge enthält keine konvergente Teilfolge, im Widerspruch zur Folgenkompaktheit von X.

Wählen wir nun $\delta > 0$ nach (α) und zu diesem δ die $x_1, \ldots, x_n$ nach (β). Zu jedem x_j gibt es nach (α) ein $U_j \in \mathcal{U}$ mit $K(x_j, \delta) \subset U_j$. Damit folgt

$$X = \bigcup_{j=1}^{n} K(x_j, \delta) \subset \bigcup_{j=1}^{n} U_j,$$

d. h. $\mathcal{U}$ enhält eine endliche Überdeckung; X ist also kompakt.

b) Für jedes $\varepsilon > 0$ ist $\{K(x, \varepsilon) : x \in X\}$ eine offene Überdeckung von X. Da X kompakt ist, gibt es also endlich viele $x_1, \ldots, x_n$ mit $X = \bigcup_{j=1}^{n} K(x_j, \varepsilon)$. ■

Satz 1.15 *Jede abgeschlossene Teilmenge einer kompakten Menge ist kompakt.*

Beweis. Sei X kompakt, $A \subset X$ abgeschlossen, $\mathcal{U}$ eine offene Überdeckung von A. Dann ist $\mathcal{U} \cup \{X \backslash A\}$ eine offene Überdeckung von X; diese enthält eine endliche Überdeckung. Also reichen endlich viele Mengen aus $\mathcal{U}$ zur Überdeckung von A. ■

1.6 Vollständige metrische Räume, der Satz von Baire

Eine Folge (x_n) in einem metrischen Raum (X, d) heißt eine *Cauchyfolge*, wenn zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert mit $d(x_n, x_m) < \varepsilon$ für $n, m \geq N$.

Ein metrischer Raum (X, d) heißt *vollständig*, wenn jede Cauchyfolge aus X in X konvergent ist. Z. B. ist $\mathbb{R}^m$ bezüglich der in Beispiel 1.1 angegebenen Metriken vollständig.

Der folgende Satz ist für viele Anwendungen wichtig.

Satz 1.16 (Banachscher Fixpunktsatz) *Sei (X,d) ein metrischer Raum, $f : X \to X$ eine* Kontraktion, *d. h. es existiert ein $\gamma < 1$ mit $d(f(x), f(y)) \leq \gamma d(x,y)$ für alle $x, y \in X$. Dann hat f genau einen Fixpunkt.*

Beweis. *Existenz*: Für ein beliebiges $x_0 \in X$ sei $x_{n+1} := f(x_n)$ für $n \in \mathbb{N}_0$. Dann gilt

$$d(x_{n+1}, x_n) \leq \gamma d(x_n, x_{n-1}) \leq \ldots \leq \gamma^n d(x_1, x_0).$$

Also gilt für $p, q \in \mathbb{N}$ mit $p < q$

$$d(x_p, x_q) \leq d(x_p, x_{p+1}) + \ldots + d(x_{q-1}, x_q) \leq d(x_0, x_1)\Big\{\gamma^p + \ldots + \gamma^{q-1}\Big\}.$$

Dies wird beliebig klein für hinreichend große p und q; also ist (x_n) eine Cauchyfolge, d. h. es existiert ein $x \in X$ mit $x_n \to x$. Wegen $f(x) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} x_{n+1} = x$ ist dies ein Fixpunkt.

Eindeutigkeit: Sind a, b Fixpunkte, so gilt

$$d(a,b) = d(f(a), f(b)) \leq \gamma d(a.b).$$

Wegen $\gamma < 1$ ist dies nur möglich, wenn $d(a,b) = 0$ gilt, also $a = b$. ∎

Bemerkung 1.17 Es genügt in obigem Satz nicht $d(f(x), f(y)) < d(x,y)$ vorauszusetzen. In $X = [-\infty, 0]$ hat $f(x) = x - \exp x$ keinen Fixpunkt, obwohl wegen $0 < f'(x) < 1$ gilt $|f(x) - f(y)| < |x - y|$. (Der Eindeutigkeitsteil des Satzes gilt auch in dieser Situation unverändert.)

Beispiel 1.18 Ist $f(x,y)$ stetig und erfüllt eine Lipschitz–Bedingung bezüglich y ($|f(x,y) - f(x,y')| \leq L|y - y'|$), so ist das AWP

$$y' = f(x,y), \quad y(x_0) = y_0$$

auf $[x_0 - 1/L, x_0 + 1/L]$ eindeutig lösbar. Durch Umschreiben des Problems in

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))\,\mathrm{d}t$$

erkennt man, daß gerade die Fixpunkte der Abbildung

$$F : F(y)(x) := y_0 + \int_{x_0}^{x} f(t, y(t)\,\mathrm{d}t$$

gesucht sind. Sei nun $0 < a < 1/L$ .Für $x \in [x_0 - a, x_0 + a]$ und $y, z \in C([x_0 - a, x_0 + a])$ gilt

$$\begin{aligned} |F(y(x)) - F(z(x))| &= \Big| \int_{x_0}^{x} \Big\{ f(t, y(t)) - f(t, z(t))\,\mathrm{d}t \Big| \\ \leq \int_{x_0}^{x} L|y(t) - z(t)|\,\mathrm{d}t &\leq \max \Big\{ |y(x) - z(x)| : x \in [x_0 - a, x_0 + a] \Big\}. \end{aligned}$$

F ist also eine Kontraktion in dem vollständigen metrischen Raum $C([x_0 - a, x_0 + a])$.
□

Satz 1.19 (Satz von Baire) *Sei (X, d) ein vollständiger metrischer Raum, (A_j) eine Folge von abgeschlossenen Teilmengen aus X mit $X = \cup_{j=1}^{\infty} A_j$. Dann enthält mindestens ein A_j eine offene Kugel $K(x, \delta)$. (Man sagt auch, daß ein vollständiger metrischer Raum von* zweiter Kategorie *ist, weshalb dieser Satz auch als* Bairescher Kategoriensatz *bezeichnet wird.)*

Beweis. Wir nehmen an, daß kein A_j eine Kugel enthält. $X \backslash A_1$ ist offen und nicht leer. Also existieren

$$r_1 < \frac{1}{2} \quad \text{und} \quad x_1 \in X \backslash A_1 \quad \text{mit} \quad K(x_1, r_1) \subset X \backslash A_1.$$

Auf Grund unserer Annahme ist $K(x_1, \frac{r_1}{2}) \not\subset A_2$, d. h.

$$K(x_1, \frac{r_1}{2}) \cap (X \backslash A_2) \quad \text{ist offen und nicht leer.}$$

Also existieren

$$r_2 < 2^{-2} \quad \text{und} \quad x_2 \quad \text{mit} \quad K(x_2, r_2) \subset K(x_1, \frac{r_1}{2}) \cap (X \backslash A_2).$$

Auf Grund unserer Annahme ist $K(x_2, \frac{r_2}{2}) \not\subset A_3$, d. h.

$$K(x_2, \frac{r_2}{2}) \cap (X \backslash A_3) \quad \text{ist offen und nicht leer.}$$

Also existieren

$$r_3 < 2^{-3} \quad \text{und} \quad x_3 \quad \text{mit} \quad K(x_3, r_3) \subset K(x_2, \frac{r_2}{2}) \cap (X \backslash A_3),$$

usw. ...

Dies liefert Folgen (x_n) aus X und (r_n) aus $(0, \infty)$ mit

$$r_n < 2^{-n} \quad \text{und} \quad K(x_n, r_n) \subset K(x_{n-1}, \frac{r_{n-1}}{2}) \cap (X \backslash A_n).$$

Für jedes $n \in \mathbb{N}$ gilt also

$$x_n \in K(x_N, r_n) \quad \text{für alle} \quad n \geq N,$$

d. h. (x_n) ist eine Cauchyfolge. Da X vollständig ist, existiert ein $x \in X$ mit $x_n \to x$. Es gilt für alle $n \in \mathbb{N}$

$$x \in \overline{K}(x_{n+1}, r_{n+1}) \subset \overline{K}(x_n, \frac{r_n}{2}) \subset K(x_n, r_n),$$

also $x \in X \backslash A_n$ für alle $n \in \mathbb{N}$, im Widerspruch zu $X = \cup A_n$. ■

Dieser Satz wird bei der späteren Untersuchung linearer Operatoren zwischen Banachräumen eine entscheidende Rolle spielen (vgl. Kap. 7).

Satz 1.20 a) *Jeder kompakte metrische Raum ist vollständig.*

b) *Sei (X, d) ein vollständiger metrischer Raum, $A \subset X$. A ist genau dann abgeschlossen, wenn der metrische Raum (A, d) vollständig ist.*

c) *Sei (X, d) ein vollständiger metrischer Raum. X ist genau dann kompakt, wenn X total beschränkt ist.*

Beweis. a) und b) können dem Leser überlassen werden.
c) $\Rightarrow$: Wurde bereits in Satz 1.14 bewiesen.
$\Leftarrow$: Nach Satz 1.14 a) genügt es, die Folgenkompaktheit von X zu beweisen. Sei (x_n) eine Folge aus X. Wir definieren induktiv Teilfolgen $(x_n^{(j+1)})$ von (x_n) mit

$$(x_n^{(j)})_{n\in\mathbb{N}} \supset (x_n^{(j+1)})_{n\in\mathbb{N}} \quad \text{und} \quad d(x_\ell^{(j)}, x_m^{(j)}) \le \frac{2}{j} \quad \text{für} \quad j, \ell, m \in \mathbb{N}.$$

Die Folgen $(x_n^{(0)} := (x_n), (x_n^{(1)}), \ldots, (x_n^{(j)})$ seien bestimmt. Wegen der totalen Beschränktheit von X gibt es

$$z_1, \ldots, z_k \in X \quad \text{mit} \quad X = \bigcup_{\ell=1}^{k} K(z_\ell, \frac{1}{j+1}).$$

Also enthält $(x_n^{(j)})$ eine (unendliche) Teilfolge $(x_n^{(j+1)})$, die ganz in *einem* $K(z_\ell, \frac{1}{j+1})$ liegt. Offensichtlich gilt $d(x_\ell^{(j+1)}, x_m^{(j+1)}) \le \frac{2}{j+1}$ für alle $\ell, m \in \mathbb{N}$.

Die Diagonalfolge $(x_n^{(n)})$ ist offensichtlich eine Cauchyfolge, also eine konvergente Teilfolge der vorgegebenen Folge (x_n). ■

1.7 Erzeugung von Topologien aus Umgebungen, Produkttopologie

Wir haben in Abschnitt 1.3, ausgehend von einer vorgegebenen Topologie $\mathcal{T}$ auf X (d. h. den sogenannten offenen Mengen), die Umgebungen der Punkte $x \in X$ definiert. Wir wollen nun umgekehrt aus den Umgebungen (bzw. Umgebungsbasen) der Punkte von X die Topologie konstruieren. Dazu ist es wichtig zu wissen, welche Eigenschaften eine Familie von Mengen erfüllen muß, damit sie ein Umgebungssystem (bzw. eine Umgebungsbasis) bildet.

Sind $\mathcal{U}_x$ die Familien aller Umgebungen von $x \in X$, so gilt:

(i) $x \in U$ für jedes $U \in \mathcal{U}_x$,

(ii) sind $U, V \in \mathcal{U}_x$, so ist $U \cap V \in \mathcal{U}_x$ (d. h. $\mathcal{U}_x$ ist ein „*Filter*“),

(iii) ist $U \in \mathcal{U}_x$ und $V \supset U$, so ist $V \in \mathcal{U}_x$,

(iv) ist $U \in \mathcal{U}_x$, so gibt es ein $V \in \mathcal{U}_x$ mit $U \in \mathcal{U}_y$ für jedes $y \in V$.

Für die Umgebungsbasen $\mathcal{B}_x$ von $x \in X$ gilt entsprechend:

(i') $x \in U$ für jedes $U \in \mathcal{B}_x$,

(ii') sind $U, V \in \mathcal{B}_x$, so existiert ein $W \in \mathcal{B}_x$ mit $W \subset U \cap V$ ($\mathcal{B}_x$ ist „*Filterbasis*“),

(iv') zu jedem $U \in \mathcal{B}_x$ gibt es ein $V \in \mathcal{B}_x$ so, daß zu jedem $y \in V$ ein $W \in \mathcal{B}_y$ existiert mit $W \subset U$.

(Die Eigenschaften (i), (ii), (iii), (i') und (ii') sind klar; in (iv) wählt man für V z. B. eine offene Menge mit $x \in V \subset U$; in (iv') wählt man ein in dieser offenen Umgebung enthaltenes $V \in \mathcal{B}_x$.)

Satz 1.21 *Sei X eine Menge.*

a) *Für jedes $x \in X$ sei $\mathcal{U}_x$ eine Familie von Teilmengen von X mit den Eigenschaften (i), (ii), (iii) und (iv). Dann gibt es genau eine Topologie $\mathcal{T}$ auf X, für die $\mathcal{U}_x$ die Umgebungssysteme von $x \in X$ sind. $\mathcal{T}$ ist gegeben durch*

$$A \in \mathcal{T} \Leftrightarrow \forall x \in A \ \exists U \in \mathcal{U}_x \textit{ mit } U \subset A.$$

b) *Für jedes $x \in X$ sei $\mathcal{B}$ eine Familie von Teilmengen von X mit den Eigenschfaten* (i'), (ii') *und* (iv'). *Dann gibt es genau eine Topologie $\mathcal{T}$ auf X, für die $\mathcal{B}_x$ Umgebungsbasen von $x \in X$ sind. $\mathcal{T}$ ist gegeben durch*

$$A \in \mathcal{T} \Leftrightarrow \forall x \in A \ \exists U \in \mathcal{B}_x \textit{ mit } U \subset A.$$

Beweis. a) Falls es überhaupt eine Topologie gibt, für die die $\mathcal{U}_x$ Umgebungssysteme sind, muß $\mathcal{T}$ nach Satz 1.5 gegeben sein durch (Eindeutigkeit)

$$A \in \mathcal{T} \Leftrightarrow \forall x \in A \ \exists U \in \mathcal{U}_x \quad \text{mit} \quad U \subset A.$$

Dieses $\mathcal{T}$ erfüllt offenbar die Eigenschaften einer Topologie. Jedes $U \in \mathcal{U}_x$ ist auf Grund der Konstruktion Umgebung von x in der Topologie $\mathcal{T}$: Ist W die Menge aller y mit

$U \in \mathcal{U}_y$, so ist $W \subset U$; zu jedem $y \in W$ existiert nach (iv) ein $V_y \in \mathcal{U}_y$ mit $U \in \mathcal{U}_z$ für jedes $z \in V_y$; also ist $V_y \subset W$ und somit $W \in \mathcal{U}_y$ für jedes $y \in W$, d. h. $W \in \mathcal{T}$ und somit U eine $\mathcal{T}$–Umgebung von x.

Sei umgekehrt U eine $\mathcal{T}$–Umgebung von x, d. h. es existiert eine Menge $A \in \mathcal{T}$ mit $x \in A \subset U$. Nach Definition von $\mathcal{T}$ enthält A ein $V \in \mathcal{U}_x$. Wegen $V \subset A \subset U$ und Eigenschaft (iii) ist also $U \in \mathcal{U}_x$.

b) Nimmt man zu $\mathcal{B}_x$ alle Obermengen von Mengen aus $\mathcal{B}_x$ hinzu, so erhält man ein System $\mathcal{U}_x$ mit den Eigenschaften (i), (ii), (iii) und (iv). Hierzu gibt es genau eine Topologie $\mathcal{T}$, die man offenbar auch durch

$$A \in \mathcal{T} \Leftrightarrow \forall x \in A\ \exists U \in \mathcal{B}_x \quad \text{mit} \quad U \subset A$$

definieren kann, so, daß $\mathcal{U}_x$ die Umgebungssysteme sind. Für jedes x ist dann offenbar $\mathcal{B}_x$ eine Umgebungsbasis. ■

Seien $(X_j, \mathcal{T}_j)$ topologische Räume $(j = 1, \ldots, n)$. Für jedes $x = (x_1, \ldots, x_n) \in \prod_{j=1}^{n} X_j$ sei

$$\mathcal{B}_x = \Big\{ \prod_{j=1}^{n} U_{x_j} : U_{x_j} \in \mathcal{U}_{x_j} \Big\},$$

wobei $\mathcal{U}_{x_j}$ das Umgebungssystem von $x_j \in X_j$ ist. $\mathcal{B}_x$ erfüllt die Eigenschaft (i'), (ii') und (iv'). Nach Satz 1.17 gibt es also genau eine Topologie $\mathcal{T}$ auf $\prod_{j=1}^{n} X_j$, für die die $\mathcal{B}_x$ Umgebungsbasen von $x = (x_1, \ldots, x_n)$ sind:

$$A \in \mathcal{T} \Leftrightarrow \forall x = (x_1, \ldots, x_n) \in A \quad \exists U_{x_j} \in \mathcal{U}_{x_j} \quad \text{mit} \quad \prod_{j=1}^{n} U_{x_j} \subset A.$$

Diese Topologie auf $\prod x_j$ wird als die *Produkttopologie* der $\mathcal{T}_1, \ldots, \mathcal{T}_n$ bezeichnet, kurz

$$\mathcal{T} = \mathcal{T}_1 \times \ldots \times \mathcal{T}_n.$$

Offensichtlich ist die Produkttopologie gerade so konstruiert, daß sie die schwächste Topologie auf dem Produkt ist, für die die *Projektionen*

$$\Big(\prod_{j=1}^{n} X_j, \mathcal{T}\Big) \to (X_k, \mathcal{T}_k), \quad (x_1, \ldots, x_n) \mapsto x_k \quad (k = 1, \ldots, n)$$

stetig sind.

Sind (X_j, d_j) $(j = 1, \ldots, n)$ metrische Räume, so wird die Produkttopologie auf $\prod X_j$ erklärt durch die Metrik

$$e_1(x, y) := \sum_{j=1}^{n} d_j(x_j, y_j)$$

für $x = (x_1, \ldots, x_n) \in \prod X_l$, $y = (y_1, \ldots, y_n) \in \prod Y_j$, denn es gilt

$$\begin{aligned} &\prod_{j=1}^{n} K_{d_j}(x_j, \frac{\varepsilon}{n}) \subset K_{e_1}(x, \varepsilon), , \\ &K_{e_1}(x, \max_j \varepsilon_j) \subset \prod_{j=1}^{n} K_{d_j}(x, \varepsilon). \end{aligned}$$

Offensichtlich kann die Produkttopologie auch erzeugt werden durch die Metriken

$$e_p(x, y) = \begin{cases} \Big\{ \sum_{j=1}^{n} d_j(x_j, y_j)^p \Big\}^{1/p} & \text{für } 1 \le p < \infty, \\ \max_j \{d_j(x_j, y_j)\} & \text{für } p = \infty. \end{cases}$$

1.8 Übungsaufgaben

1.1 a) Seien d_m $(m \in \mathbb{N})$ Metriken auf X. Dann ist auch $d := \sum_m 2^{-m} d_m/(1+d_m)$ eine Metrik auf X.

b) Seien d_m und d wie in a). Dann gilt: (x_n) konvergiert gegen x bezüglich d genau dann, wenn dies bezüglich jeder der Metriken d_m gilt.

c) Sei $X = C(\mathbb{R})$ der Raum der auf $\mathbb{R}$ stetigen Funktionen,

$$d_m(f, g) = \max\{|f(t) - g(t)| : |t| \le m\} \qquad (m \in \mathbb{N}, f, g \in C(\mathbb{R}));$$

d sei wie in a) erklärt. Ist $f(t) = \sum a_n t^n$ eine Potenzreihe mit Konvergenzradius ∞, so gilt $f_k \to f$ bezüglich d für $f_k(t) = \sum_{n \le k} a_n t^n$.

1.2 Sei $\mathcal{T}$ eine Topologie auf X.

a) Ist $\mathcal{T}$ separiert, so ist jede endliche Menge abgeschlossen.

b) $\mathcal{T}$ ist genau dann separiert, wenn jeder Punkt gleich dem Durchschnitt aller seiner abgeschlossenen Umgebungen ist.

1.3 In $\mathbb{R}$ sei $\mathcal{T}$ die Familie der Teilmengen A für die gilt: Zu jedem $x \in A$ existiert ein $\varepsilon > 0$ mit $[x, x+\varepsilon) \subset A$.

a) $\mathcal{T}$ ist eine Topologie auf $\mathbb{R}$.

b) Man charakterisiere

α) die konvergenten Folgen in $(\mathbb{R}, \mathcal{T})$,

β) die stetigen Funktionen $f : (\mathbb{R}, \mathcal{T}) \to \mathbb{R}$.

1.4 a) Auf $[0, 2\pi)$ ist durch

$$d(x,y) = \min\Big\{|x-y|, |x-y+2\pi|, |x-y-2\pi|\Big\}$$

eine Metrik erklärt (geometrische Deutung am Einheitskreis).

b) Man charakterisiere die stetigen Funktionen $f : \big([0, 2\pi), d\big) \to \mathbb{R}$.

1.5 Seien X, Y topologische Räume, $f : X \to Y$.

a) f ist genau dann stetig, wenn für jede abgeschlossene Teilmenge $B \subset Y$ das Urbild $f^{-1}(B)$ in X abgeschlossen ist.

b) Sei f stetig.

α) Für jede Teilmenge $A \subset X$ gilt $f\overline{(A)} \subset \overline{f(A)}$.

β) Im allgemeinen gilt **nicht**:

(i) $f\overline{(A)} = \overline{f(A)}$,

(ii) für jede abgeschlossene (offene) Menge $A \subset X$ ist $f(A)$ abgeschlossen (offen).

1.6 Seien $(X, \mathcal{T}_X)$ und $(Y, \mathcal{T}_Y)$ topologische Räume.

(a) Ist $\mathcal{T}_X$ die feinste Topologie auf X (die diskrete) *oder* $\mathcal{T}_Y$ die gröbste Topologie auf Y, so ist jede Abbildung $f : X \to Y$ stetig

(b) Ist $\mathcal{T}_X$ die gröbste Topologie und $\mathcal{T}_Y$ separiert, so sind nur die konstanten Abbildungen $f : X \to Y$ stetig.

1.7 (a) Sei (X, d) ein kompakter metrischer Raum, $f : X \to X$ eine *isometrische* Abbildung (d. h. $d(f(x), f(y)) = d(x, y)$). Dann ist f bijektiv.
Anleitung: Man zeige: Gibt es ein x_0, das nicht im Bild von f liegt, so enthält die Folge $x_1 = f(x_0), \ldots, x_{n+1} = f(x_n), \ldots$ keine konvergente Teilfolge.

(b) Die Aussage von Teil a) gilt nicht, wenn (X, d) nicht kompakt ist

1.8 (a) Seien d, e *uniform äquivalente* Metriken auf X, d. h. es gilt

$$d(x, y) \leq c_1 e(x, y), \quad e(x, y) \leq c_2 d(x, y) \quad \text{für alle} \quad x, y \in X.$$

Dann sind die durch d und e erzeugten Topologien gleich und (X, d) ist genau dann vollständig, wenn (X, e) vollständig ist.

(b) Sei d eine Metrik auf X, e eine Metrik auf Y. Dann sind

$$d_p((x_1, y_1), (x_2, y_2)) = \begin{cases} d(x_1, x_2) + e(y_1, y_2) & \text{für } p = 1 \\ \{d(x_1, x_2)^2 + e(y_1, y_2)^2\}^{1/2} & \text{für } p = 2 \\ \max\{d(x_1, x_2), e(y_1, y_2)\} & \text{für } p = \infty \end{cases}$$

Metriken auf $X \times Y$ und je zwei davon sind uniform äquivalent.

1.9 Auf $\mathbb{R}$ ist durch

$$d(x, y) = |\arctan x - \arctan y|$$

eine Metrik definiert. Die hierdurch erzeugte Topologie stimmt mit der üblichen Topologie auf $\mathbb{R}$ überein. $(\mathbb{R}, d)$ ist aber nicht vollständig.

1.10 Zwei Metriken sind genau dann äquivalent, wenn bezüglich beider Metriken die gleichen Folgen konvergent sind.

1.11 Sind $(X, \mathcal{T}_X)$ und $(Y, \mathcal{T}_Y)$ kompakte topologische Räume, so ist auch das topologische Produkt $(X \times Y, \mathcal{T}_X \times \mathcal{T}_Y)$ kompakt. (Das gilt dann auch für Produkte von endlich vielen kompakten Räumen. Nach dem Satz von Tychonoff (vgl. z. B.Schubert: Topologie) gilt dies auch für Produkte aus beliebig vielen kompakten Räumen.)

1.12 Seien (X, d) und (Y, e) metrische Räume, X sei kompakt, $f : X \to Y$ bijektiv und stetig (nach Satz 1.13 ist dann auch Y kompakt). Man zeige, daß auch f^{-1} stetig ist.

2 Topologische Vektorräume, lokalkonvexe Räume

2.1 Topologische Vektorräume

Eine Topologie $\mathcal{T}$ auf einem Vektorraum X über $\mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$) heißt *mit den linearen Operationen verträglich*, oder kurz eine *Vektorraumtopologie*, wenn die beiden Abbildungen

$$\begin{array}{rrcll} A: & X \times X & \to & X, & (x, y) \mapsto x + y, \\ M: & \mathbb{K} \times X & \to & X, & (a, x) \mapsto ax \end{array}$$

stetig sind. Dabei ist $\mathbb{K}$ mit der üblichen Topologie versehen, $X \times X$ und $\mathbb{K} \times X$ mit der jeweiligen Produkttopologie.

$(X, \mathcal{T})$, den Vektorraum X versehen mit einer Vektorraumtopologie $\mathcal{T}$, nennt man dann einen *topologischen Vektorraum.*

Satz 2.1 *Sei $(X, \mathcal{T})$ ein topologischer Vektorraum, $v \in X, a \in \mathbb{K}\backslash\{0\}$. Dann sind die Abbildungen*

$$A_v : X \to X,\ x \mapsto x + v, \quad \textit{und} \quad M_a : X \to X,\ x \mapsto ax$$

Homöomorphismen *(d. h. sie sind stetig, bijektiv und haben eine stetige Inverse).*

Beweis. $\mathbf{A_v}$: Wegen der Stetigkeit von A gibt es zu jeder Umgebung U von $x + v$ Umgebungen U_x von x und U_v von v mit

$$y + w \in U \quad \text{für} \quad y \in U_x \quad \text{und} \quad w \in U_v,$$

also insbesondere

$$y + v \in U \quad \text{für alle} \quad y \in U_x.$$

Das ist die Stetigkeit von A_v. Die Abbildung A_v ist offensichtlich bijektiv, $(A_v)^{-1} = a_{-v}$, und A_{-v} ist ebenfalls stetig (Beweis wie für A_v).

$\mathbf{M_a}$: Wegen der Stetigkeit von M gibt es zu jeder Umgebung U von ax Umgebungen U_x von x (in X) und U_a von a (in $\mathbb{K}$) mit

$$by \in U \quad \text{für} \quad y \in U_x, b \in U_a,$$

also insbesondere

$$ay \in U \quad \text{für alle} \quad y \in U_x.$$

Das ist die Stetigkeit von M_a. Auch M_a ist bijektiv, $(M_a)^{-1} = M_{a^{-1}}$ ist stetig. ■

Damit erhält man leicht

Satz 2.2 *Sei $(X, \mathcal{T})$ ein topologischer Vektorraum, $v \in X, a \in \mathbb{K} \backslash \{0\}$.*

(a) Die folgenden Aussagen für $B \subset X$ sind äquivalent:

- *i. B ist offen.*
- *ii. $B + v = \{x + v : x \in B\}$ ist offen.*
- *iii. $aB = \{ax : x \in B\}$ ist offen.*

(b) Die folgenden Aussagen für $U \subset X$ sind äquivalent:

- *i. U ist eine Umgebung von $0 \in X$ (*Nullumgebung*),*
- *ii. $U + v$ ist eine Umgebung von v,*
- *iii. aU ist eine Nullumgebung.*

Beweis. a) Dies folgt aus der Stetigkeit von A_v und M_a mit Hilfe der Definition der Stetigkeit.

b) Dies folgt aus der Stetigkeit von A_v und M_a mit Hilfe der Charakterisierung der Stetigkeit durch Umgebungen. ■

Dieser Satz besagt: Wenn wir von einer Vektorraumtopologie die Nullumgebungen kennen, dann kennen wir alle Umgebungen und damit die Topologie insgesamt; natürlich genügt es dafür, eine Nullumgebungsbasis zu kennen. Aber, welche über Satz 1.21 hinausgehenden Eigenschaften muß eine Familie von Teilmengen von X erfüllen, damit sie eine mit den linearen Operationen verträgliche Topologie erzeugt? Diese Frage wird durch den folgenden Satz beantwortet.

Satz 2.3 *Eine Familie $\mathcal{B}$ von Teilmengen eines Vektorraumes X ist genau dann eine Nullumgebungsbasis einer Vektorraumtopologie auf X, wenn gilt:*
(TL 0) *$0 \in U$ für jedes $U \in \mathcal{B}$; zu $U, V \in \mathcal{B}$ existiert ein $W \in \mathcal{B}$ mit $W \subset U \cap V$,*
(TL 1) *für jedes $U \in \mathcal{B}$ existiert ein $V \in \mathcal{B}$ mit $V + V \subset U$,*

(TL 2) *für jedes $U \in \mathcal{B}$ existiert ein $V \in \mathcal{B}$ mit $aV \subset U$ für alle $a \in \mathbb{K}$ mit $|a| \leq 1$,*
(TL 3) *jedes $U \in \mathcal{B}$ ist* absorbierend *(auch* ausgeglichen*), d. h. für jedes $x \in X$ existiert ein $c > 0$ mit $ax \in U$ für jedes $a \in \mathbb{K}$ mit $|a| \leq c$.*

Beweis. $\Longrightarrow$: (TL 0) gilt für jede Umgebungsbasis des Punktes 0 (Filterbasis).
(TL 1) Dies ist die Stetigkeit von A im Punkt $(0,0) \in X \times X$.
(TL 2) Aus der Stetigkeit von M im Punkt $(0,0) \in \mathbb{K} \times X$ folgt: Für jedes $U \in \mathcal{B}$ existiert ein $\varepsilon > 0$ und ein $W \in \mathcal{B}$ mit $bW \subset U$ für alle $b \in \mathbb{K}, |b| \leq \varepsilon$. Für $V := \varepsilon W$ folgt hieraus: $aV = a\varepsilon W \subset U$ für alle $a \in \mathbb{K}$ mit $|a| \leq 1$.
(TL 3) Wir nehmen an, daß ein $U \in \mathcal{B}$ nicht absorbierend ist. Dann existiert ein $x \in X$ und eine Nullfolge (a_n) aus $\mathbb{K}$ mit $a_n x \notin U$ für alle $n \in \mathbb{N}$. Wegen $a_n x \to 0$ (Folgenstetigkeit von M) ist dies ein Widerspruch.

$\Longleftarrow$: Für jedes $x \in X$ sei $\mathcal{B}_x = \{U + x : U \in \mathcal{B}\}$ als Umgebungsbasis von x gewählt. Die Eigenschaften (i′), (ii′) und (iv′) aus Abschnitt 1.7 sind offensichtlich erfüllt (insbesondere folgt (iv′) aus (TL 0)).

Es ist zu zeigen, daß A und M bezüglich der hierdurch erzeugten Topologie $\mathcal{T}$ auf X stetig sind.

Stetigkeit von A: Sei U eine Umgebung von $x + y$. Dann existiert ein $U' \in \mathcal{B}$ mit $x + y + U' \subset U$. Nach (TL 1) existiert ein $V \in \mathcal{B}$ mit $V + V \subset U'$, also

$$(x + V) + (y + V) = (x + y) + (V + V) \subset (x + y) + U' \subset U.$$

Stetigkeit von M: Sei U eine Umgebung von ax. Dann existiert ein $U' \in \mathcal{B}$ mit $ax + U' \subset U$. Sei $n \in \mathbb{N}$ so, daß $|a| \leq n$ gilt. Durch iterative Anwendung von (TL 1) findet man ein $V' \in \mathcal{B}$ mit

$$V' + \ldots + V' \subset U' \qquad (n + 2 \quad \text{Terme}).$$

Nach (TL 2) existiert ein $V'' \in \mathcal{B}$ mit

$$cV'' \subset V' \qquad \text{für alle} \quad c \in \mathbb{K} \quad \text{mit} \quad |c| \leq 1.$$

Nach (TL 3) existiert ein $m \in \mathbb{N}$ mit

$$\frac{1}{m} x \in V''.$$

Damit erhalten wir für $|a-b| \leq \frac{1}{m}, y \in x + V''$:

$$\begin{aligned} by &= ax + (b-a)x + (b-a)(y-x) + a(y-x) \\ &= ax + [(b-a)m]\frac{1}{m}x + (b-a)(y-x) + n[\frac{a}{n}(y-x)], \end{aligned}$$

$$\begin{aligned} &(b-a)m \leq 1, \frac{1}{m}x \in V'' &&\Rightarrow [(b-a)m]\frac{1}{m}x \in V', \\ &|b-a| \leq 1, y-x \in V'' &&\Rightarrow b-a)(y-x) \in V', \\ &\frac{|a|}{n} \leq 1, y-x \in V'' &&\Rightarrow \frac{a}{n}(y-x) \in V' \quad \Rightarrow n\frac{a}{n}(y-x) \in nV', \end{aligned}$$

also

$$\begin{aligned} by &\in ax + V' + V' + nV' \subset ax + V' + \ldots + V' \quad (n+2 \text{ Terme}) \\ &\subset ax + U' \subset U. \end{aligned}$$

Damit ist auch die Stetigkeit der Multiplikation bewiesen. ■

Eine Teilmenge U des Vektorraumes X heißt *kreisförmig*, wenn für jedes $x \in U$ und $a \in \mathbb{K}$ mit $|a| \leq 1$ gilt $ax \in U$.

Satz 2.4 *Sei $(X, \mathcal{T})$ ein topologischer Vektorraum. Dann besitzt $(X, \mathcal{T})$ eine Nullumgebungsbasis aus kreisförmigen Mengen. (Wir werden deshalb im folgenden meist voraussetzen, daß die verwendeten Nullumgebungsbasen aus kreisförmigen Mengen bestehen.)*

Beweis. Nach (TL 2) bilden die kreisförmigen Mengen

$$\tilde{U} := \bigcup_{|\alpha| \leq 1} \alpha U \qquad (U \in \mathcal{B})$$

eine Nullumgebungsbasis. ■

2.2 Erzeugung von Topologien durch Halbnormen

Die gängigste Methode, auf einem Vektorraum eine mit linearen Operationen verträgliche Topologie zu erklären, ist die mit Hilfe von Normen oder Halbnormen. Eine Funktion

$$p : X \to [0, \infty)$$

heißt eine *Halbnorm* auf X, wenn gilt
(N 1) $p(ax) = |a|p(x)$ für $x \in X, a \in \mathbb{K}$ (*positiv homogen*),
(N 2) $p(x + y) \leq p(x) + p(y)$ für $x, y \in X$ (*Dreiecksungleichung*),
Eine Halbnorm p heißt eine *Norm*, wenn zusätzlich gilt
(N 3) $p(x) = 0 \Rightarrow x = 0$ (*Positivität*).
Wegen (N 1) gilt für jede Halbnorm: $p(0) = 0$; für eine Norm gilt also $p(x) = 0 \Leftrightarrow x = 0$. Normen werden im folgenden meist mit $\|\cdot\|$ bezeichnet; falls eine Unterscheidung verschiedener Normen nötig ist, werden Indices verwendet.

Auf einem Vektorraum X sei eine Familie

$$\{p_\alpha : \alpha \in A\}$$

von Halbnormen gegeben. Mit Hilfe der *Kugeln*

$$K(\alpha, \varepsilon) := \{x \in X : p_\alpha(x) < \varepsilon\}$$

wird eine *Nullumgebungsbasis* wie folgt definiert:

$$U(\alpha_1, \ldots, \alpha_n; \varepsilon_1, \ldots, \varepsilon_n) = \bigcap_{j=1}^{n} K(\alpha_j, \varepsilon_j) \quad \text{für} \quad n \in \mathbb{N}, \alpha_1, \ldots, \alpha_n \in A, \varepsilon_1, \ldots, \varepsilon_n > 0.$$

Es ist sehr leicht zu sehen, daß dieses Mengensystem die Eigenschaften (TL 0) – (TL 3) einer Nullumgebungsbasis einer Vektorraumtopologie erfüllt. Es wird also hierdurch eine mit den linearen Operationen verträgliche Topologie auf X erzeugt. Die Nullumgebungen dieser Basis sind offensichtlich kreisförmig.

Satz 2.5 *Die durch die Familie $\{p_\alpha : \alpha \in A\}$ von Halbnormen erzeugte Topologie ist genau dann separiert, wenn zu jedem $x \in X \backslash \{0\}$ ein $\alpha \in A$ existiert mit $p_\alpha(x) \neq 0$. Insbesondere ist jede durch eine (oder mehrere)* Normen *erzeugte Topologie separiert. — Ein topologischer Vektorraum X, dessen Topologie durch eine Norm $\|\cdot\|$ erzeugt wird, heißt ein* normierter Raum, *$(X, \|\cdot\|)$.*

Beweis. $\Longrightarrow$: Ist die durch $\{p_\alpha : \alpha \in A\}$ erzeugte Topologie separiert, so existiert zu jedem $x \neq 0$ eine Nullumgebung $U(\alpha_1, \ldots, \alpha_n; \varepsilon_1, \ldots, \varepsilon_n)$ mit

$$x \notin U(\alpha_1, \ldots, \alpha_n; \varepsilon_1, \ldots, \varepsilon_n).$$

Also gilt

$$p_{\alpha_j}(x) \geq \varepsilon_j \quad \text{für mindestens ein} \quad j \in \{1, \ldots, n\}.$$

Insbesondere ist $p_{\alpha_j}(x) \neq 0$ für dieses j.

$\Longleftarrow$: Sei $p_\alpha(x) \neq 0$. Dann hat die Nullumgebung

$$U := \left\{ y \in X : p_\alpha(y) < \frac{1}{2} p_\alpha(x) \right\}$$

offenbar die Eigenschaft

$$(0 + U) \cap (x + U) = \emptyset,$$

d. h. 0 und x haben disjunkte Umgebungen. ∎

Ein topologischer Raum heißt *metrisierbar*, wenn die Topologie durch eine Metrik beschrieben werden kann.

Satz 2.6 *Sei $(X, \mathcal{T})$ ein separierter topologischer Vektorraum, dessen Topologie durch eine höchstens abzählbare Familie $\{p_n\}$ von Halbnormen erzeugt wird. Dann ist $(X, \mathcal{T})$ metrisierbar; die Metrik kann* translationsinvariant *gewählt werden, d. h. es gilt $d(x, y) = d(x + z, y + z)$ für alle $x, y, z \in X$.*

Bemerkung 2.7 *Die gleiche Aussage gilt für jeden topologischen Vektorraum mit abzählbarer Nullumgebungsbasis (erstes Abzählbarkeitsaxiom), vgl. G.Köthe: Topologische lineare Räume, 15.11.*

Beweis von Satz 2.6. Offensichtlich erfüllen $d_n(x, y) := p_n(x - y)$ die Eigenschaften einer Metrik bis auf die Eigenschaft „$d_n(x, y) = 0 \Rightarrow x = y$“. Man zeigt leicht (vgl. Satz 1.3 und Aufgabe 1.1), daß dann auch

$$d(x, y) := \sum_n 2^{-n} \frac{d_n(x, y)}{1 + d_n(x, y)} = \sum_n 2^{-n} \frac{p_n(x - y)}{1 + p_n(x - y)}$$

diese Eigenschaften hat. Da $(X, \mathcal{T})$ separiert ist, existiert nach Satz 2.6 zu $x \neq y$ ein n mit $p_n(x-y) > 0$, also ist

$$d(x,y) > 0 \qquad \text{für} \quad x \neq y,$$

d. h. d ist eine Metrik.

Offenbar gelten die Inklusionen

$$K_d(0, 2^{-n} \min_j \frac{\varepsilon_j}{1+\varepsilon_j}) \subset \bigcap_{j=1}^n K(p_j, \varepsilon_j) \qquad \text{für alle} \quad n \in \mathbb{N}, \varepsilon_1, \ldots, \varepsilon_n > 0,$$

$$\bigcap_{j=1}^n K(p_j, \frac{\varepsilon}{2}) \subset K_d(0, \varepsilon) \qquad \text{für alle} \quad \varepsilon > 0, n \text{ mit } 2^{-n} < \frac{\varepsilon}{2}.$$

d. h. die Metrik d und die Familie von Halbnormen $\{p_n\}$ erzeugen die gleiche Topologie auf X. ■

Eine Teilmenge A eines Vektorraumes X heißt *absolut konvex*, wenn gilt

$$x, y \in A, |\alpha| + |\beta| \leq 1 \Longrightarrow \alpha x + \beta y \in A.$$

Eine Menge A ist offenbar genau dann absolut konvex, wenn sie konvex und kreisförmig ist.

Ein topologischer Vektorraum X heißt *lokalkonvex*, wenn er eine Nullumgebungsbasis aus absolut konvexen Mengen besitzt.

Da die angegebene Nullumgebungsbasis einer durch Halbnormen erzeugten Topologie offensichtlich aus absolut konvexen Mengen besteht, ist die durch Halbnormen erzeugte Topologie stets lokalkonvex. Es gilt auch die Umkehrung:

Satz 2.8 *Die Topologie auf einem lokalkonvexen Raum $(X, \mathcal{T})$ kann durch Halbnormen erzeugt werden.*

Beweis. Sei $\{U_\alpha : \alpha \in A\}$ eine Nullumgebungsbasis in $(X, \mathcal{T})$ aus absolut konvexen Mengen. Für jedes U_α definieren wir das zugehörige *Minkowski–Funktional*

$$p_\alpha(x) := \inf\{c \geq 0 : x \in cU_\alpha\}.$$

Die p_α sind Halbnormen:

$$\text{(N 1)} \qquad \begin{aligned} p_\alpha(ax) &:= \inf\{c \geq 0 : ax \in cU_\alpha\} \\ &= \inf\{|a|c \geq 0 : x \in cU_\alpha\} \\ &= |a| \cdot \inf\{c > 0 : x \in cU_\alpha\} = |a| p_\alpha(x). \end{aligned}$$

$$\text{(N 2)} \qquad \begin{aligned} p_\alpha(x+y) &= \inf\{c \geq 0 : x + y \in cU_\alpha\} \\ &\leq \inf\{d + e : d \geq 0, e \geq 0, x \in dU_\alpha, y \in eU_\alpha\} \\ &= \inf\{d \geq 0 : x \in dU_\alpha\} + \inf\{e \geq 0, y \in U_\alpha\} \\ &= p_\alpha(x) + p_\alpha(y). \end{aligned}$$

Wegen (vgl. Aufgabe 2.2)

$$U_\alpha \subset \{x \in X : p_\alpha(x) \leq 1\} \subset (1+\varepsilon)U_\alpha \qquad \text{für} \quad \varepsilon > 0,$$

erzeugen die Halbnormen p_α $(\alpha \in A)$ die gleiche Topologie wie die Umgebungsbasis $\{U_\alpha : \alpha \in A\}$. ■

2.3 Stetige lineare Abbildungen zwischen topologischen Vektorräumen

Wir wenden uns nun erstmals dem zentralen Gegenstand der Funktionalanalysis zu, den *stetigen linearen Abbildungen* von einem topologischen Vektorraum $(X, \mathcal{T}_X)$ in einem topologischen Vektorraum $Y, \mathcal{T}_Y$; dabei ist insbesondere auch der Fall $Y = \mathbb{K}$ (versehen mit der üblichen Topologie) interessant, man spricht dann von *stetigen linearen Funktionalen.*

Der folgende Satz besagt, daß es für den Nachweis der Stetigkeit einer linearen Abbildung genügt, die Stetigkeit im Nullpunkt allein zu untersuchen.

Satz 2.9 *Seien $(X, \mathcal{T}_X)$ und $(Y, \mathcal{T}_Y)$ topologische Vektorräume. Eine lineare Abbildung $L : X \to Y$ ist genau dann stetig, wenn sie im Nullpunkt stetig ist, d. h. wenn es zu jeder Nullumgebung V in Y eine Nullumgebung U in X gibt mit $LU \subset V$ (natürlich kann man sich dabei auf die Elemente entsprechender Nullumgebungsbasen beschränken).*

Beweis. $\Rightarrow$: Offensichtlich.

$\Leftarrow$: Sei $x \in X$; wir zeigen, daß L in x stetig ist. Dazu sei W eine Umgebung von Lx, d. h. $V := W - Lx$ ist eine Nullumgebung in Y. Da L in 0 stetig ist, existiert eine Nullumgebung U in X mit $LU \subset V$. Dann ist $\tilde{U} := U + x$ eine Umgebung von x mit

$$L\tilde{U} = LU + Lx \subset V + Lx = W.$$

Damit ist die Stetigkeit von L in x bewiesen. ■

Für den Fall, daß die Topologien auf X und Y durch Halbnormen definiert sind, bringen wir den vorhergehenden Satz in eine Form, die in der Regel leicht explizit nachprüfbar sein dürfte.

Satz 2.10 *Seien $(X, \mathcal{T}_X)$ und $(Y, \mathcal{T}_Y)$ topologische Vektorräume, deren Topologien durch Halbnormen $\{p_\alpha : \alpha \in A\}$ bzw. $\{q_\beta : \beta \in B\}$ erzeugt sind. Dann ist eine lineare Abbildung $L : X \to Y$ genau dann stetig, wenn zu jedem $\beta \in B$ eine endliche Menge $A_\beta \subset A$ und ein $C_\beta \geq 0$ existieren mit*

$$q_\beta(Lx) \leq C_\beta \max\{p_\alpha(x) : \alpha \in A_\beta\} \qquad \textit{für alle} \quad x \in X.$$

Korollar 2.11 *(a) Ist die Topologie auf Y durch eine Norm $\|\cdot\|$ erzeugt, so ist L genau dann stetig, wenn eine endliche Menge $A_0 \subset A$ und ein $C \geq 0$ existieren mit*

$$\|Lx\| \leq C \max\{p_\alpha(x) : \alpha \in A_0\} \qquad \textit{für alle} \quad x \in X.$$

(b) Sind die Topologien auf X und *Y durch Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ erzeugt, so ist L genau dann stetig, wenn ein $C \geq 0$ existiert mit*

$$\|Lx\|_2 \leq C\|x\|_1 \qquad \textit{für alle} \quad x \in X.$$

Beweis von Satz 2.10. $\Longrightarrow$: Sei L stetig, d. h. zu jeder Nullumgebung der Form

$$U_\beta := \{y \in Y : q_\beta(y) \leq 1\} \quad \text{in} \quad Y$$

gibt es eine Nullumgebung

$$U(\alpha_1, \ldots, \alpha_n; \varepsilon_1, \ldots, \varepsilon_n) := \{x \in X : p_{\alpha_j}(x) \leq \varepsilon_j \quad \text{für} \quad j = 1, \ldots, n\} \quad \text{in} \quad X$$

mit

$$LU(\alpha_1, \dots, \alpha_n; \varepsilon_1, \dots, \varepsilon_n) \subset U_\beta.$$

Für jedes $x \in X$ sei

$$K(x) := \max\{p_{\alpha_j}(x)\frac{1}{\varepsilon_j} : j = 1, \dots, n\}.$$

Dann gilt offenbar

$$x \in K(x)U(\alpha_1, \dots, \alpha_n; \varepsilon_1, \dots, \varepsilon_n) = U(\alpha_1, \dots, \alpha_n; K(x)\varepsilon_1, \dots, K(x)\varepsilon_n),$$

also

$$\begin{aligned} Lx \in L\Big(K(x)U(\alpha_1, \dots, \alpha_n; \varepsilon_1, \dots, \varepsilon_n)\Big) \subset K(x)U_\beta, \\ q_\beta(Lx) \quad &\leq \quad K(x) = \max\{p_{\alpha_j}(x)\frac{1}{\varepsilon_j} : j = 1, \dots, n\} \\ &\leq \quad \max\{p_{\alpha_j}(x) : j = 1, \dots, n\} \max\{\frac{1}{\varepsilon_j} : j = 1, \dots, n\}. \end{aligned}$$

Das ist die gewünschte Behauptung mit

$$A_\beta = \{\alpha_1, \dots, \alpha_n\} \quad \text{und} \quad C_\beta = \max\{\frac{1}{\varepsilon_j} : j = 1, \dots, n\}.$$

$\Longleftarrow$: Sei U eine Nullumgebung in Y, d. h.

$$U \supset U(\beta_1, \dots, \beta_n; \varepsilon_1, \dots, \varepsilon_n) = \bigcap_{i=1}^{n} K(\beta_i, \varepsilon_i).$$

Nach Voraussetzung gilt

$$q_{\beta_i}(Lx) \leq C_{\beta_i} \max \Big\{p_\alpha(x(: \alpha \in A_{\beta_i}\Big\},$$

also $Lx \in K(\beta_i, \varepsilon_i)$, für

$$Lx \in \bigcap_{\alpha \in A_{\beta_i}} K(\alpha_i, \varepsilon_i / C_{\beta_i})$$

und somit $x \in U$ sicher dann, wenn $x \in V$ gilt mit

$$V = \bigcap_{i=1}^{n} \bigcap_{\alpha \in A_{\beta_i}} K(\alpha_i, \varepsilon_i / C_{\beta_i}).$$

Da dieses V eine Nullumgebung in X ist, ist L stetig. ■

2.4 Vektorraumtopologien auf $\mathbb{K}^m$

Auf $\mathbb{K}^m$ wird üblicherweise die Topologie betrachtet, die durch die *euklidische Metrik*

$$d(x,y) = \Big\{ \sum_{j=1}^{m} |x_j - y_j|^2 \Big\}^{1/2} \qquad \text{für} \quad x = (x_1, \ldots, x_m), y = (y_1, \ldots, y_m),$$

bzw. die *euklidische Form*

$$\|x\| = \Big\{ \sum_{j=1}^{m} |x_j|^2 \Big\}^{1/2} \qquad \text{für} \quad (x_1, \ldots, x_m)$$

erzeugt wird. Wir bezeichnen sie kurz als die *euklidische Topologie* $\mathcal{E}$ auf $\mathbb{K}^m$.

Auch die anderen in Beispiel 1.1 definierten Metriken erzeugen offensichtlich diese Topologie. Da sie durch eine Norm (z. B. die euklidische) erzeugt wird, ist sie nach Abschnitt 2.2 mit den linearen Operationen verträglich. Es gilt sogar:

Satz 2.12 *Auf $\mathbb{K}^m$ gibt es nur eine mit den linearen Operationen verträgliche separierte Topologie, die euklidische Topologie $\mathcal{E}$.*

Beweis. Sei $\mathcal{T}$ eine mit den linearen Operationen verträgliche separierte Topologie auf $\mathbb{K}^m$. Es ist also $\mathcal{T} = \mathcal{E}$ zu beweisen.

a) $\mathcal{T} \subset \mathcal{E}$: Es ist zu zeigen, daß jede $\mathcal{T}$–Umgebung auch eine $\mathcal{E}$–Umgebung ist. Dafür genügt es zu zeigen, daß jede $\mathcal{T}$–Nullumgebung eine $\mathcal{E}$–Kugel um 0 enthält (denn diese bilden eine Nullumgebungsbasis der $\mathcal{E}$–Topologie).

Sei U eine $\mathcal{T}$–Nullumgebung. Da es eine Basis von kreisförmigen $\mathcal{T}$–Nullumgebungen gibt (Satz 2.4), gibt es eine kreisförmige $\mathcal{T}$–Nullumgebung V mit

$$V + \ldots + V \subset U \qquad (m \quad \text{Terme}).$$

Zu jedem Einheitsvektor $e_i = (\delta_{ij})_{j=1,\ldots,m}$ gibt es (da V absorbierend ist) ein $a_i > 0$ mit

$$ae_i \in V \qquad \text{für alle} \quad a \in \mathbb{K} \quad \text{mit} \quad |a| \leq a_i.$$

Für $b_1, \ldots, b_m$ mit $\sum |b_j|^2 \leq 1$ folgt also

$$\min\{a_i, : i = 1, \ldots, m\} \sum b_j e_j = \sum b_j \min\{a_i : i = 1, \ldots, m\} e_j \in V + \ldots + V \subset U,$$

d. h. U enthält die $\mathcal{E}$–Kugel um 0 mit dem Radius $\min\{a_i : i = 1, \dots, m\}$.

b) $\mathcal{E} \subset \mathcal{T}$: Wir zeigen zuerst:

(i) Es gibt eine kreisförmige $\mathcal{T}$–Nullumgebung, die keinen (nichttrivialen) Teilraum von $\mathbb{K}^m$ enthält.

Beweis von **(i)**: Hierfür genügt es offenbar zu zeigen: Ist U eine kreisförmige Nullumgebung, die höchstens k–dimensionale Teilräume enthält ($k > 0$), so gibt es eine kreisförmige Nullumgebung, die höchstens $(k-1)$–dimensionale Teilräume enthält. Sei also U eine kreisförmige Nullumgebung, die höchstens k–dimensionale Teilräume enthält. Dann existiert eine kreisförmige Nullumgebung V_1 mit $V_1 + V_1 \subset U$. Dieses V_1 enthält *höchstens einen* k–dimensionalen Teilraum: Wären H_1 und H_2 zwei verschiedene k–dimensionale Teilräume mit $H_j \subset V_1$, so wäre $H := H_1 + H_2$ ein mindestens $(k+1)$–dimensionaler Teilraum mit $H \subset V_1 + V_1 \subset U$.

Sei also H dieser k–dimensionale Teilraum mit $H \subset V_1, x \in H \backslash \{0\}$. Dann existiert eine kreisförmige Nullumgebung W mit $x \notin W$ (hier wird benutzt, daß $\mathcal{T}$ eine separierte Topologie ist). Die Nullumgebung $V := V_1 \cap W$ enthält höchstens $(k-1)$–dimensionale Teilräume.

(ii) Sei U eine kreisförmige $\mathcal{T}$–Nullumgebung, die keinen Teilraum enthält, V eine kreisförmige $\mathcal{T}$–Nullumgebung mit $V + V \subset U$. Dann existiert ein ϱ so, daß V in der $\mathcal{E}$–Kugel um 0 mit Radius ϱ enthalten ist. (Die $\mathcal{E}$–Kugel mit Radius r enthält also die $\mathcal{T}$–Umgebung $\frac{r}{\varrho} V$, womit die Behauptung bewiesen ist.)

Beweis von **(ii)**: Wir nehmen an, daß es kein solches ϱ gibt. Dann gibt es eine Folge (y_n) aus V mit $\|y_n\| \nearrow \infty$ ($\| \cdot \| =$ euklidische Norm). Die Folge (x_n) mit

$$x_n = \|y_n\|^{-1} y_n$$

ist beschränkt,enthält also eine $\mathcal{E}$–konvergente Teilfolge mit (x_{n_k})

$$x_{n_k} \xrightarrow{\mathcal{E}} x \qquad \text{mit} \quad x \neq 0 (\|x\| = 1).$$

Nach Teil a ist $\mathcal{T}$ die schwächere Topologie, d. h. es gilt dann auch

$$x_{n_k} \xrightarrow{\mathcal{T}} x.$$

Daraus folgt, , da $\|y_{n_k}\|$ eine Nullumgebung ist

$$x_{n_\ell} \in x + \|y_{n_k}\|^{-1} V \qquad \text{für} \quad \ell \geq \ell(k).$$

Da V kreisförmig ist, folgt hieraus für $\ell \geq \max(k, \ell(k))$

$$\begin{aligned} x \in x_{n_\ell} + \|y_{n_k}\|^{-1}V &= \|y_{n_\ell}\|^{-1}y_{n_\ell} + \|y_{n_k}\|^{-1}V \\ &\subset \|y_{n_\ell}\|^{-1}V + \|y_{n_k}\|^{-1}V \subset \|y_{n_k}\|^{-1}(V+V) \subset \|y_{n_k}\|^{-1}U, \end{aligned}$$

also

$$\|y_{n_k}\|x \in U \qquad \text{für} \quad k \in \mathbb{N}.$$

Wegen $\|y_{n_k}\| \to \infty$ ist dies ein Widerspruch zu der Voraussetzung, daß U keinen Teilraum enthält. ■

Korollar 2.13 *Auf jedem endlichdimensionalen Vektorraum gibt es nur eine Vektorraumtopologie.*

Beweis. Sei $\dim X = m$. Dann gibt es einen (algebraischen) Isomorphismus J von X auf $\mathbb{K}^m$. Sei $\mathcal{T}$ eine Vektorraumtopologie auf X. Wir definieren $\mathcal{T}'$ auf $\mathbb{K}^m$ durch

$$U' \in \mathcal{T}' \Longleftrightarrow U' = JU \quad \text{mit} \quad \mathcal{U} \in \mathcal{T}.$$

Man sieht leicht ein, daß $\mathcal{T}'$ eine Vektorraumtopologie auf $\mathbb{K}^m$ ist. Da es auf $\mathbb{K}^m$ nur eine Vektorraumtopologie gibt, und da $\mathcal{T}$ durch $\mathcal{T}'$ eindeutig bestimmt ist, ist damit die Behauptung bewiesen. ■

2.5 Übungsaufgaben

2.1 Das Minkowski–Funktional einer absolut–konvexen Menge U ist genau dann eine Norm, wenn U keinen nicht–trivialen Teilraum enthält.

2.2 Sei U eine absolut–konvexe Menge, p_α das zugehörige Minkowski–Funktional. Dann gilt $U \subset \{x \in X : p_U(x) \leq 1\} \subset (1+\varepsilon)U \quad$ für $\quad \varepsilon > 0$.

2.3 Sei $K = \{x \in \mathbb{R}^m : |x| < r\}$ mit einem $r > 0, \mathcal{D}(K)$ der Raum der beliebig oft differenzierbaren Funktionen auf $\mathbb{R}^m$, die außerhalb K verschwinden. Die Topologie auf $\mathcal{D}(K)$ werde durch die Normen

$$\|f\|_n := \sup\{|D^\alpha f(x)| : x \in K, |\alpha \leq n\} \qquad (n \in \mathbb{N}_0),$$

bzw. durch die Metrik

$$d(f,g) := \sum_n 2^{-n}\|f-g\|_n(1+\|f-g\|_n)^{-1}$$

beschrieben.

(a) Für jeden Multiindex α ist die Abbildung $D^\alpha : \mathcal{D}(K) \to \mathcal{D}(K)$ injektiv und stetig.

Im folgenden sei $m = 1$.

(a) Man bestimme den Wertebereich der Abbildung $D : \mathcal{D}(K) \to \mathcal{D}(K), f \mapsto f'$.

(b) Der Wertebereich von D ist abgeschlossen und hat Codimension 1 (Hyperebene).

2.4 Der Raum $\mathcal{D}(K)$ aus Aufgabe 2.3 ist vollständig.

2.5 Eine Teilmenge A eines topologischen Vektorraumes heißt *beschränkt*, wenn es zu jeder Nullumgebung U ein $\lambda \geq 0$ gibt mit $A \subset \lambda U$; sie heißt *relativ kompakt*, wenn $\overline{A}$ kompakt ist.

(a) Jede (relativ) kompakte Teilmenge eines topologischen Vektorraumes ist beschränkt.

(b) Im Raum $\mathcal{D}(K)$ aus Aufgabe 2.3 ist jede abgeschlossene beschränkte Menge kompakt (jede beschränkte Menge relativ kompakt); ein topologischer Vektorraum mit dieser Eigenschaft heißt ein *Montel–Raum*.

3 Lineare Funktionale, der Satz von Hahn-Banach, schwache Topologie

3.1 Der topologische Dualraum

Satz 3.1 *Seien* $(X, \mathcal{T}_X)$ *und* $(Y, \mathcal{T}_Y)$ *topologische Vektorräume über* $\mathbb{K}$. *Die Menge der stetigen linearen Abbildungen von* X *nach* Y *bildet, versehen mit den üblichen linearen Operationen, einen* $\mathbb{K}$*–Vektorraum. — Speziell gilt dies für die stetigen linearen Abbildungen von* X *nach* $\mathbb{K}$ *(die* stetigen linearen Funktionale auf X*); der Raum der stetigen Funktionale auf* $(X, \mathcal{T}_X)$ *wird als der* topologische Dualraum $X^* = (X, \mathcal{T}_X)^*$ *bezeichnet.*

Beweis. Für stetige lineare Abbildungen L, M von X nach Y und $a \in \mathbb{K}$ ist zu zeigen, daß auch aL und $L + M$ stetig sind.

aL: Ist $a = 0$, so ist aL die Nullabbildung, also sicher stetig. Sei also $a \neq 0$. Zu jeder Nullumgebung V in Y existiert eine Nullumgebung U in X mit $LU \subset V$, also $aL(a^{-1}U) \subset V$; da auch $a^{-1}U$ eine Nullumgebung in X ist, ist also aL stetig.

L + M: Zu jeder Nullumgebung V in Y gibt es eine Nullumgebung V' in Y mit $V' + V' \subset V$ und Nullumgebungen U_1 und U_2 in X mit

$$LU_1 \subset V' \text{ und } MU_2 \subset V',$$

also

$$(L + M)(U_1 \cap U_2) \subset V' + V' \subset V.$$

Da auch $U_1 \cap U_2$ eine Nullumgebung in X ist, ist also $L + M$ stetig. ■

In diesem Abschnitt interessieren wir uns insbesondere für *stetige Funktionale*. Das nächste Beispiel zeigt, daß es nicht auf jedem topologischen Vektorraum von Null verschiedene stetige lineare Funktionale gibt:

Beispiel 3.2 Auf $X := \tilde{C}[0,1]$, dem Raum der stückweise stetigen Funktionen auf $[0,1]$ sei

$$\|f\|_p := \left\{ \int_0^1 |f(x)|^p \, \mathrm{d}x \right\}^{1/p} \text{ mit } 0 < p < 1.$$

Wir benutzen diese Schreibweise, obwohl $\|\cdot\|_p$ für die hier betrachteten $p < 1$ keine Norm ist (dies ergibt sich aus dem Inhalt dieses Paragraphen, kann aber auch leicht direkt nachgewiesen werden; Aufgabe 3.4).

Durch

$$U_\varepsilon := \{f \in \tilde{C}[0,1] : \|f\|_p^p < \varepsilon\} \qquad (\varepsilon > 0)$$

wird eine Nullumgebungsbasis einer Topologie $\mathcal{T}_p$ gegeben (TL 0), (TL 2) und (TL 3) sind trivial.

(TL 1): Wir zeigen $U_{\varepsilon/2^{p+1}} + U_{\varepsilon/2^{p+1}} \subset U_\varepsilon$. Aus $f, g \in U_{\varepsilon/2^{p+1}}$ folgt

$$\begin{aligned} \|f+g\|_p^p &= \int |f+g|^p \,\mathrm{d}x \leq \int \{|2f|^p + |2g|^p\} \,\mathrm{d}x \\ &= 2^p \int \{|f|^p + |g|^p\} \,\mathrm{d}x \leq 2^{p+1} \frac{\varepsilon}{2^{p+1}} = \varepsilon, \end{aligned}$$

d. h. $f + g \in U_\varepsilon$.

Offenbar hat $\mathcal{T}_p$ eine abzählbare Nullumgebungsbasis, was aber hier nicht benötigt wird.

Wir zeigen, daß jedes von Null verschiedene lineare Funktional auf X bezüglich $\mathcal{T}_p$ unstetig ist. Sei also F ein lineares Funktional auf X, $f_0 \in X$, $F(f_0) = 1$. Für $s \in [0,1]$ definieren wir

$$\begin{aligned} f_s^{(1)}(x) &:= \begin{cases} f_0(x) & \text{für } 0 \leq x \leq s, \\ 0 & \text{für } s < x \leq 1, \end{cases} \\ f_s^{(2)}(x) &:= f_0(x) - f_s^{(1)}(x). \end{aligned}$$

Wenn s von 0 nach 1 läuft, wächst $\|f_s^{(1)}\|_p^p$ stetig von 0 bis $\|f_0\|_p^p$. Also existiert ein $t \in (0,1)$ mit

$$\|f_t^{(1)}\|_p^p = \|f_t^{(2)}\|_p^p = \frac{1}{2}\|f_0\|_p^p.$$

Es gilt

$$|F(f_t^{(j)}| \geq \frac{1}{2} \text{ für } j = 1 \text{ oder } 2.$$

Sei $f_1 = 2f_t^{(j)}$ mit diesem j. Dann gilt

$$|F(f_1)| \geq 1 \text{ und } \|f_1\|_p^p = 2^p \|f_t^{(j)}\|_p^p = 2^{p-1}\|f_0\|_p^p.$$

Verfährt man mit f_n entsprechend, usw., so erhält man eine Folge (f_n) aus X mit

$$|F(f_n)| \geq 1 \text{ und } \|f_n\|_p^p = 2^{n(p-1)}\|f_0\|_p^p \to 0 \text{ für } n \to \infty,$$

also $f_n \xrightarrow{\mathcal{T}} 0$, aber $|F(f_n)| \geq 1 \not\to 0$, d. h. F ist nicht stetig.

(Mit Hilfe des Lebesgue-Integrals [vgl. §5] kann analog zu $L^p(0,1)$ mit $p \geq 1$ [vgl. §7] auch $L^p(0,1)$ mit $p \leq 1$definiert werden; $\tilde{C}[0,1]$ ist dicht in $L^p(0,1)$) bezüglich dieser Topologie und die obige Konstruktion zeigt, daß auf $L^p(0,1)$ kein von 0 verschiedenes stetiges lineares Funktional existiert.) □

Satz 3.3 *Sei $(X, \mathcal{T})$ ein topologischer Vektorraum. Eine Halbnorm p auf X ist genau dann stetig, wenn sie im Nullpunkt stetig ist.*

Beweis. $\Longrightarrow$: Offensichtlich.

$\Longleftarrow$: Sei p stetig in 0, d. h. für jedes $\varepsilon > 0$ existiert eine kreisförmige Nullumgebung U in X mit $p(z) < \varepsilon$ für alle $z \in U$. Wir zeigen die Stetigkeit bei einem beliebigen $x \in X$: Für $y \in x + U$ sind $y - x \in U$ und (da U kreisförmig ist) $x - y \in U$, also

$$\begin{aligned}
&p(x) = p(x - y + y) \leq p(x-y) + p(y) \leq \varepsilon + p(y)\\
&p(y) = p(y - x + x) \leq p(y-x) + p(x) \leq \varepsilon + p(x),\\
&|p(x) - p(y)| \leq \varepsilon,
\end{aligned}$$

d. h. p ist stetig im Punkt x. ■

Der folgende Satz, der die topologischen Vektorräume mit nicht–trivialen stetigen linearen Funktionalen charakterisiert, kann erst mit Hilfe des Satzes von Hahn–Banach (Abschnitt 3.2) vollständig bewiesen werden (Implikation (ii) ⇒ (i)).

Satz 3.4 *Sei $(X, \mathcal{T})$ ein topologischer Vektorraum. Die folgenden Aussagen sind äquivalent:*

(i) *es existiert ein nicht–triviales stetiges lineares Funktional auf X (d. h. es gilt $(X, \mathcal{T})^* \neq \{0\}$),*

(ii) *es existiert eine nicht–triviale stetige Halbnorm auf X,*

(iii) *es existiert eine von X verschiedene absolut konvexe Nullumgebung in X.*

Beweis. **(i) $\Rightarrow$ (iii)**: Sei F ein nicht–triviales stetiges Funktional auf X. Dann ist

$$U := \{x \in X : |F(x) \leq 1\}$$

— *absolut konvex*: für $x, y \in U, a, b \in \mathbb{K}$ mit $|a| + |b| \leq 1$ gilt $|F(ax + by)| \leq a|F(x)| + b|F(y)| \leq 1$,

— eine *Nullumgebung*: $U = F^{-1}\{z \in \mathbb{K} : |z| \leq 1\}$,

— *von X verschieden*: Ist $F(x_0) \neq 0$, so ist $|F(2F(x_0)^{-1}x_0)| = 2$, also $2F(x_0)^{-1}x_0 \notin U$.

(iii) $\Rightarrow$ (ii): Sei $U \neq X$ eine absolut konvexe $\mathcal{T}$–Nullumgebung in X,

$$p_U(x) := \inf\{\lambda \in [0, \infty) : x \in \lambda U\}$$

das zugehörige Minkowski–Funktional (vgl. Beweis von Satz 2.8). Wegen

$$p_U(x) \leq \varepsilon \text{ für } x \in \varepsilon U$$

ist p_U stetig. Es existiert ein $x_0 \in X \backslash U$, also mit $p_U(x_0) \geq 1$, d. h. p_U ist nicht trivial.

(ii)$\Rightarrow$ (i): Diesen Teil beweisen wir im Anschluß an den Satz von Hahn–Banach im folgenden Abschnitt. ∎

3.2 Der Satz von Hahn–Banach

Bei der Konstruktion stetiger linearer Funktionale wird es außerordentlich nützlich sein, stetige lineare Funktionale auf einem Teilraum zu stetigen linearen Funktionalen auf dem gesamten Raum fortzusetzen (wobei man Eindeutigkeit in der Regel nicht erwarten darf, vgl. auch 7). Hierzu ist folgender Satz nützlich:

Satz 3.5 (Fortsetzungssatz von Hahn Banach) *Sei X ein Vektorraum, p eine Halbnorm auf X, Y ein Teilraum von X, F_0 ein lineares Funktional auf Y mit*

$$|F_0(x)| \leq p(x) \textit{ für alle } x \in Y,$$

Dann existiert ein lineares Funktional F auf X mit

$$\begin{aligned} F(x) = F_0(x) \quad & \textit{für alle } x \in X \textit{ (d. h. } F \textit{ ist } \text{Fortsetzung } \textit{von } F_0), \\ |F(x)| \leq p(x) \quad & \textit{für alle } x \in X. \end{aligned}$$

Beweis. a) Sei zuerst X ein *reeller* Vektorraum.

(α) Wir zeigen zunächst: Jedes auf einem echten Teilraum T von X definierte lineare Funktional f mit $|f(x)| \leq p(x)$ besitzt eine (echte) Fortsetzung f_1 auf einem Teilraum T_1 von X mit $T \subsetneq T_1$ und

$$|f_1(x)| \leq p(x) \text{ für alle } x \in T_1.$$

Sei $x_1 \in X \backslash T$, $T_1 := L\{T, x_1\}(L\{\ldots\} :=$ *Lineare Hülle* von $\{\ldots\})$, also $T \subsetneq T_1$. Eine Fortsetzung f_1 von f muß auf T_1 zwangsläufig die Form

$$f_1(x + ax_1) = f(x) + \lambda a \text{ für } x \in T, a \in \mathbb{R}$$

haben. Es stellt sich lediglich die Frage, ob $\lambda \in \mathbb{R}$ so gewählt werden kann, daß für alle Elemente aus T_1 die gewünschte Ungleichung gilt.

Aus der Voraussetzung über f folgt für $y, z \in T$:

$$\begin{aligned} f(z) - f(y) &= f(z-y) \leq p(z-y) \leq p(z+x_1) + p(y+x_1), \\ -f(y) - p(y+x_1) &\leq -f(z) + p(z+x_1), \end{aligned}$$

also

$$\lambda_1 := \sup_{y \in T}\{-f(y) - p(y+x_1)\} \leq \inf_{z \in T}\{-f(z) + p(z+x_1)\} =: \lambda_2.$$

Wählt man nun $\lambda \in [\lambda_1, \lambda_2] (\neq \emptyset)$, so gilt für

$$\begin{aligned}
\mathbf{a > 0}: \quad a\lambda &\leq a\lambda_2 \leq a\Big(-f(\frac{1}{a}x) + p(\frac{1}{a}x + x_1)\Big) = -f(x) + p(x + ax_1), \\
\mathbf{a < 0}: \quad a\lambda &\leq a\lambda_1 \leq a\Big(-f(\frac{1}{\alpha}x) - p(\frac{1}{a}x + x_1)\Big) = -f(x) + |a|p(\frac{1}{a}x + x_1) \\
&= -f(x) + p(x + ax_1), \\
\mathbf{a = 0}: \quad a\lambda &= 0 \leq -f(x) + p(x) = -f(x) + p(x + ax_1).
\end{aligned}$$

Also gilt in jedem Fall für $z = x + ax_1 \in T_1$

$$f_1(z) = f(x) + a\lambda \leq p(z).$$

Anwendung dieser Ergebnisse auf $-z$ liefert

$$-f_1(z) = f_1(-z) \leq p(-z) = p(z),$$

insgesamt also

$$|f_1(z)| \leq p(z) \text{ für alle } z \in T_1.$$

■/2

Um den Teil (β) des Beweises zu führen, benötigen wir einige neue Begriffe und das Lemma von Zorn. Eine Menge M heißt *halbgeordnet* bezüglich einer *Ordnungsrelation* $\prec$, wenn gilt:

(O 1) $a \prec a$ für alle $a \in M$,
(O 2) aus $a \prec b \prec c$ folgt $a \prec c$,
(O 3) aus $a \prec b$ und $b \prec a$ folgt a=b

(man beachte: *nicht* für jedes Paar $a, b \in M$ muß eine der Beziehungen $a \prec b$ oder $b \prec a$gelten), Eine Teilmenge N von M heißt *(total) geordnet*, wenn für jedes Paar $a, b \in N$ eine der Beziehungen $a \prec b$ oder $b \prec a$ gilt. Ein Element $a \in M$ heißt eine *obere Schranke* einer Teilmenge R von M, wenn für jedes $r \in R$ gilt $r \prec a$. Ein Element $m \in M$ heißt ein *maximales Element* in M, wenn aus $m \prec a$, $a \in M$ folgt $a = m$ (d. h., wenn es kein „echt größeres" Element in M gibt.

Satz von Zorn *(vgl. z. B. E. Kamke: Mengenlehre) Besitzt in einer halbgeordneten Menge jede geordnete Teilmenge eine obere Schranke, so existiert (mindestens) ein maximales Element.*

Damit können wir den ***Beweis*** von Satz 3.5 fortsetzen:

(β) Sei $\mathcal{F}$ die Menge aller Fortsetzungen f von F_0 mit

$$|f(x)| \le p(x) \text{ für alle } x \text{ aus dem Teilraum } D(f), \text{ auf dem } f \text{ definiert ist.}$$

Offenbar ist $\mathcal{F}$ bezüglich der Ordnungsrelation

$$„\subset\text{“} : f_1 \subset f_2 \Longleftrightarrow f_2 \text{ Fortsetzung von } f_1$$

halbgeordnet. Ist $\mathcal{F}_0$ eine total geordnete Teilmenge von $\mathcal{F}$, so ist offensichtlich

$$\hat{f} \text{ mit } D(\hat{f}) := \bigcup_{f \in \mathcal{F}_0} D(f) \text{ und } \hat{f}(x) = f(x) \text{ für } x \in D(f)$$

eine obere Schranke von $\mathcal{F}_0$. Nach dem Satz von Zorn existiert also ein maximales Element F in $\mathcal{F}$. Aus (α) folgt, daß F auf ganz X definiert ist (denn sonst wäre es nicht maximal).

b) Sei jetzt X ein komplexer Vektorraum. X kann auch als reeller Vektorraum aufgefaßt werden (gleiche Punktmenge, gleiche Addition; es ist aber nur die Multiplikation mit reellen Zahlen erlaubt). Dann ist

$$V_0(x) := \text{Re } F_0(x) \qquad (x \in T)$$

ein reell lineares Funktional auf T:

$$V_0(ax) = \text{Re } F_0(ax) = \text{Re}\,(aF_0(x)) = a\,\text{Re } F_0(x) = aV_0(x) \quad (a \in \mathbb{R}),$$
$$V_0(x+y) = \text{Re}\,(F_0(x) + F_0(y)) = V_0(x) + V_0(y);$$

und es gilt

$$|V_0(x)| \le |F_0(x)| \le p(x) \text{ für } x \in T.$$

Nach Teil a) existiert eine reell lineare Fortsetzung V von V_0 auf ganz X mit

$$|V(x)| \le p(x) \text{ für alle } x \in X.$$

Wir definieren

$$F(x) := V(x) - iV(ix) \text{ für } x \in X$$

und zeigen abschließend, daß F alle gewünschten Eigenschaften hat:

F ist Fortsetzung von F_0: Für $x \in T$ gilt

$$\begin{aligned} F(x) &= V(x) - iV(ix) = \text{Re } F_0(x) - i \,\text{Re } F_0(ix) = \text{Re } F_0(x) - i \,\text{Re}\,(iF_0(x)) \\ &= \text{Re } F_0(x) + i \,\text{Im } F_0(x) = F_0(x). \end{aligned}$$

F ist komplex linear: Für $x, y \in X$ gilt

$$\begin{aligned} F(x+y) &= V(x+y) - iV(ix+iy) = V(x) - iV(ix) + V(y) - iV(iy) \\ &= F(x) + F(y). \end{aligned}$$

Für $z = a + ib \in \mathbb{C}$ und $x \in X$ gilt

$$\begin{aligned} F(zx) &= F(ax + ibx) = F(ax) + F(ibx) \\ &= V(ax) - iV(iax) + V(ibx) - iV(-bx) \\ &= a(V(x) - iV(ix)) - ib(iV(ix) - V(x)) \\ &= aF(x) + ibF(x) = zF(x). \end{aligned}$$

Es gilt $|F(x)| \leq p(x)$ für alle $x \in X$: Für jedes $x \in X$ existiert ein $\varphi \in \mathbb{R}$ mit

$$\begin{aligned} |F(x)| &= e^{i\varphi} F(x) = F(e^{i\varphi} x) = \text{Re } F(e^{i\varphi} x) \\ &= V(e^{i\varphi} x) \leq p(e^{i\varphi} x) = p(x). \end{aligned}$$

■

Damit können wir nun auch den noch fehlenden Teil des ***Beweises von Satz 3.4*** nachtragen:

(ii) ⇒ (i): Sei p eine nicht–triviale stetige Halbnorm auf $(X, \mathcal{T})$, $x_0 \in X$ mit $p(x_0) \neq 0$. Wir definieren

$$F_0 : L\{x_0\} \to \mathbb{K}, F_0(ax_0) = ap(x_0) \quad (a \in \mathbb{K}).$$

Dann gilt offenbar

$$|F_0(x)| = p(x) \text{ für alle } x \in L\{x_0\}.$$

Nach dem Satz von Hahn–Banach existiert ein lineares Funktional F auf ganz X mit

$$\begin{aligned} F(x) = F_0(x) &\qquad \text{für } x \in L\{x_0\}, \\ |F(x)| \leq p(x) &\qquad \text{für } x \in X. \end{aligned}$$

Da p stetig ist, ist auch F stetig; wegen $F(x_0) = p(x_0) \neq 0$ ist F nicht–trivial. ■

3.3 Schwache Topologien

Sei X eine Menge, F eine Menge von Funktionen auf X. Man sagt, die Funktionen aus F *trennen die Punkte von* x, wenn zu je zwei Punkten $x, y \in X, x \neq y$ ein $f \in F$ existiert mit $f(x) \neq f(y)$. Ist X ein Vektorraum über $\mathbb{K}$ und F eine Menge von linearen Funktionalen, so ist dies äquivalent mit: Zu jedem $x \in X \backslash \{0\}$ existiert ein $f \in F$ mit $f(x) \neq 0$. (Sei dies der Fall und $x, y \in X, x \neq y$. Dann existiert ein $f \in F$ mit $f(x-y) \neq 0$, also $f(x) \neq f(y)$.)

Satz 3.6 *Ist $(x, \mathcal{T})$ ein separierter lokalkonvexer Raum (z. B. ein normierter Raum), so trennen die stetigen linearen Funktionale auf $(X, \mathcal{T})$ die Punkte von X.*

Beweis. Nach Satz 2.8 gibt es eine Familie $\{p_\alpha : \alpha \in A\}$ von Halbnormen, durch die die Topologie $\mathcal{T}$ erzeugt wird; die p_α sind offensichtlich $\mathcal{T}$–stetig. Sei $x \in X \setminus \{0\}$. Da die Topologie $\mathcal{T}$ separiert ist, existiert nach Satz 2.5 ein $\alpha \in A$ mit $p_\alpha(x) \neq 0$. Wir definieren

$$F_0 : L\{x\} \to \mathbb{K}, \quad F_0(ax) := ap_\alpha(x).$$

Dann ist F_0 ein auf $L\{x\}$ definiertes lineares Funktional mit

$$F_0(z) \leq p_\alpha(z) \text{ für alle } z \in L\{x\}.$$

Dieses kann unter Erhalt dieser Ungleichung zu einem Funktional F auf ganz X fortgesetzt werden. Da p_α stetig ist, ist auch F stetig und es gilt $F(x) \neq 0$. ■

Sei $(X, \mathcal{T})$ ein topologischer Vektorraum für den die stetigen linearen Funktionale die Punkte trennen (dies gilt z. B. dann, wenn $(X, \mathcal{T})$ ein separierter lokalkonvexer Raum ist). Die *schwache Topologie* $\mathcal{T}(X^*)$ auf X wird definiert durch die Nullumgebungsbasis

$$U(f_1, \ldots, f_n; \varepsilon_1 \ldots, \varepsilon_n) := \{x \in X : |f_j(x)| \leq \varepsilon_j \text{ für } j = 1, \ldots, n\} (f_j \in X^*, \varepsilon_j > 0).$$

Da die Abbildungen $x \mapsto |f_j(x)|$ Halbnormen sind, ist klar, daß hierdurch eine separierte lokalkonvexe Topologie erklärt wird. Auf Grund der Definition ist klar, daß $\mathcal{T}(X^*)$ die schwächste Topologie auf X ist, für die alle $f \in X^*$ stetig sind. Sie ist schwächer als die ursprüngliche Topologie $\mathcal{T}$ (oder gleich dieser).

Satz 3.7 *Für jede Topologie $\tilde{\mathcal{T}}$ mit $\mathcal{T}(X^*) \subset \tilde{\mathcal{T}} \subset \mathcal{T}$ ist der Dualraum von $(X, \tilde{\mathcal{T}})$ gleich dem Dualraum von $(X, \mathcal{T})$.*

Beweis. Wegen $\mathcal{T}(X^*) \subset \tilde{\mathcal{T}} \subset \mathcal{T}$ gilt offenbar

$$(X, \mathcal{T}(X^*))^* \subset (X, \tilde{\mathcal{T}})^* \subset (X, \mathcal{T})^*.$$

Andererseits ist nach Konstruktion jedes bezüglich $\mathcal{T}$ stetige Funktional auch stetig bezüglich $\mathcal{T}(X^*)$ und somit

$$(X, \mathcal{T})^* \subset (X, \mathcal{T}(X^*))^*.$$

■

Bei der Untersuchung linearer Funktionale ist häufig aus das folgende rein algebraische Lemma nützlich.

Hilfssatz 3.8 *Sei X ein Vektorraum. $f, f_1, \ldots, f_n$ seien lineare Funktionale auf X mit*

$$\bigcap_{j=1}^{n} N(f_j) \subset N(f).$$

Dann ist f eine Linearkombination von $f_1, \ldots, f_n$.

Beweis. Sei $F : X \to \mathbb{K}^n$ definiert durch $F(x) := (f_1(x), \ldots, f_n(x)))$. Dann ist

$$N(F) = \bigcap_{j=1}^{n} N(f_j) \subset N(f).$$

Also gibt es ein $g : \mathbb{K}^n \to \mathbb{K}$ mit $f = g \circ F$; z. B. kann man auf $R(F)$ definieren $g(F(x)) = f(x)$ und linear auf $\mathbb{K}^n$ fortsetzen. Mit $c_j := g(e_j)$ ist dann für alle $x \in X$

$$f(x) = g(F(x)) = g\Big(\sum_{j=1}^{n} f_j(x) e_j\Big) = \sum_{j=1}^{n} c_j f_j(x).$$

■

Ganz analog zur schwachen Topologie auf X kann eine lokalkonvexe Topologie $\mathcal{T}(X)$ auf dem Dualraum $X^* = (X, \mathcal{T})^*$ erklärt werden durch die Nullumgebungsbasis

$$U(x_1, \ldots, x_n; \varepsilon_1, \ldots, \varepsilon_n) = \{F \in X^* : |F(x_j)| \leq \varepsilon_j \text{für } j = 1, \ldots, n\}(x_j \in X, \varepsilon_j > 0).$$

Diese Topologie wird auch als die schwache $*$–*Topologie* bezeichnet.

Jedes $x \in X$ erzeugt durch

$$X^* \to \mathbb{K}, \quad F \to F(x)$$

ein lineares Funktional auf X^*. Die schwache $*$–*Topologie* ist offenbar die schwächste Topologie auf X^*, für die diese Funktionale stetig sind.

3.4 Übungsaufgaben

3.1 (a) Auf einem endlichdimensionalen topologischen Vektorraum ist jedes lineare Funktional stetig.

(b) Sei $(X, \|\cdot\|)$ ein normierter Raum, $x_1, \ldots, x_n \in X$ linear unabhängig, $\alpha_1, \ldots, \alpha_n \in K$. Dann gibt es ein $f \in X^*$ mit $f(x_j) = \alpha_j$ für $j = 1, \ldots, n$.

3.2 Auf $\mathbb{R}^2$ seien die Normen $\|\cdot\|_1$ und $\|\cdot\|_\infty$ definiert durch

$$\|(x_1, x_2)\|_1 := |x_1| + |x_2|, \quad \|(x_1, x_2)\|_\infty := \max\Big\{|x_1|, |x_2|\Big\}.$$

Auf dem 1–dimensionalen Teilraum $D = L\{e_1\}$ sei das lineare Funktional F_0 definiert durch $F_0(ce_1) = c$.

(a) Es gilt $|F_0(x)| \leq \|x\|_1 = \|x\|_\infty$ für alle $x \in D$.

(b) Man bestimme alle Fortsetzungen G bzw. H von F_0 auf ganz $\mathbb{R}^2$ mit

$$|G(x)| \leq \|x\|_1 \text{ bzw. } |H(x)| \leq \|x\|_\infty$$

für alle $x \in \mathbb{R}^2$.

3.3 (a) Sei X ein Vektorraum, F ein lineares Funktional auf X. Dann gibt es ein $y \in X$ mit $X = N(F) + L\{y\}$, wobei $N(F)$ der *Nullraum* von F ist, $N(F) := \{x \in X : F(x) = 0\}$.

(b) Ein lineares Funktional F auf einem normierten Raum $(X, \|\cdot\|)$ ist genau dann beschränkt (stetig), wenn der Nullraum $N(F)$ abgeschlossen ist.

Der Stoff der folgenden Aufgaben wird in Kapitel 6 ausführlicher behandelt.

3.4 Man zeige, daß die in Beispiel 3.2 benutzten „p–Normen"keine Normen sind. (z. B. gilt für

$$f(x) = \begin{cases} 1 & \text{in } [0, \frac{1}{2}), \\ 0 & \text{in } [\frac{1}{2}, 1], \end{cases} \qquad g(x) = 1 - f(x)$$

die Dreiecksungleichung nicht.)

3.5 Sei $C_0^\infty(\mathbb{R}^m)$ der Raum der beliebig oft stetig differenzierbaren Funktionen auf $\mathbb{R}^m$ mit kompaktem Träger

(a) Es gilt

$$\begin{aligned} C_0^\infty(\mathbb{R}^m) &= \cup\{\mathcal{D}(\Omega) : \Omega \subset \mathbb{R}^m \text{ offen und beschränkt}\} \\ &= \bigcup_{n \in \mathbb{N}} \mathcal{D}\Big(\{x \in \mathbb{R}^m : |x| < n\}\Big) \end{aligned}$$

(für $\mathcal{D}(\Omega)$ vgl. Aufgabe 2.3).

(b) In $C_0^\infty(\mathbb{R}^m)$ wird durch

$$U := \text{ Konvexe Hülle von } \bigcup_n U_n,$$

wobei die U_n Nullumgebungen in $\mathcal{D}\Big(\{x \in \mathbb{R}^m : |x| < n\}\Big)$ sind, eine Nullumgebungsbasis einer lokalkonvexen Topologie definiert. (Die konvexe Hülle einer Menge A ist der Durchschnitt aller konvexen Mengen, die A enthalten.) $C_0^\infty(\mathbb{R}^m)$, ausgestattet mit dieser Topologie, wird mit $\mathcal{D}(\mathbb{R}^m)$ bezeichnet.

(c) Ein lineares Funktional auf $\mathcal{D}(\mathbb{R}^m)$ ist genau dann stetig, wenn die Einschränkung auf $\mathcal{D}(\Omega)$ für jedes offene und beschränkte Ω stetig ist.

(d) Für (f_n) und $f \in \mathcal{D}(\mathbb{R}^m)$ gilt $f_n \to f$ genau dann, wenn eine kompakte Teilmenge $K \subset \mathbb{R}^m$ existiert mit $\operatorname{supp} f_n \subset K$ für alle n und $D^\alpha f_n \to D^\alpha f$ gleichmäßig für alle α.

3.6 Sei $\mathcal{S}(\mathbb{R}^m)$ der *Schwartzsche Raum* der *schnell–fallenden Funktionen*, d. h. $f : \mathbb{R}^m \to \mathbb{C}$ liegt in $\mathcal{S}(\mathbb{R}^m)$, wenn f beliebig oft stetig differenzierbar ist und zu jedem Multiindex α und jedem $p \in \mathbb{N}$ ein $C_{\alpha p}$ existiert mit

$$|D^\alpha f(x)|(1+|x|)^p \leq C_{\alpha p} \text{ für alle } x \in \mathbb{R}^m.$$

(a) Auf $\mathcal{S}(\mathbb{R}^m)$ wird eine lokalkonvexe Topologie definiert durch die Normen

$$\|f\|_n = \max\left\{|D^\alpha f(x) < (1+|x|)^n : |\alpha| \leq n, x \in \mathbb{R}^m\right\}$$

(b) Es gilt $\mathcal{D}(\mathbb{R}^m) \subset \mathcal{S}(\mathbb{R}^m)$ und die Einbettung ist stetig.

(c) Für (f_k) und f aus $\mathcal{S}(\mathbb{R}^m)$ gilt $f_k \to f$ genau dann, wenn für jeden Multiindex α und jedes $n \in \mathbb{N}$ gilt:

$$(1+|x|)^n D^\alpha f_k(x) \to (1+|x|)^n D^\alpha f(x) \text{ gleichmäßig in } \mathbb{R}^m.$$

3.7 Sei $\mathcal{E}(\mathbb{R}^m)$ der Raum der beliebig oft stetig differenzierbaren Funktionen auf $\mathbb{R}^m$ (ohne Bedingung an das Verhalten im Unendlichen).

(a) Durch die Halbnormen

$$p_n(f) = \max\left\{D^\alpha f(x)| : |\alpha| \leq n, |x| \leq n\right\}$$

wird auf $\mathcal{E}(\mathbb{R}^m)$ eine lokalkonvexe Topologie definiert.

(b) Es gilt $\mathcal{S}(\mathbb{R}^m) \subset \mathcal{E}(\mathbb{R}^m)$ und die Einbettung ist stetig.

(c) Ist F ein stetiges lineares Funktional auf $\mathcal{E}(\mathbb{R}^m)$, so existiert ein $n \in \mathbb{N}$ mit

$$F(f) = 0 \text{ für } f \in \mathcal{E}(\mathbb{R}^m) \text{ mit } f(x) = 0 \text{ in } \{x \in \mathbb{R}^m : |x| \leq n\}.$$

(d) Für (f_n) und f auf $\mathcal{E}(\mathbb{R}^m)$ gilt $f_n \to f$ genau dann, wenn für jeden Multiindex α gilt: $D^\alpha f_n$ konvergiert lokal gleichmäßig gegen $D^\alpha f$.

3.8 (a) Es gilt

$$\mathcal{E}(\mathbb{R}^m)^* \subset \mathcal{S}(\mathbb{R}^m)^* \subset \mathcal{D}(\mathbb{R}^m)^*.$$

Man bezeichnet

(i) die Elemente von $\mathcal{D}(\mathbb{R}^m)^*$ als *Distributionen* auf $\mathbb{R}^m$,

(ii) die Elemente von $\mathcal{S}(\mathbb{R}^m)^*$ als *temperierte Distributionen*,

(iii) die Elemente von $\mathcal{E}(\mathbb{R}^m)^*$ als *Distributionen mit kompaktem Träger* (vgl. Aufgabe 3.7 c).

(b) Sei X einer der Räume $\mathcal{D}(\mathbb{R}^m), \mathcal{S}(\mathbb{R}^m), \mathcal{E}(\mathbb{R}^m)$, und $\varphi : \mathbb{R}^m \to \mathbb{C}$ stetig. Dann ist

$$F(f) = \int \varphi(x) f(x)\,\mathrm{d}x$$

(α) in $\mathcal{D}(X)^*$,
(β) in $\mathcal{S}(X)^*$, falls $|\varphi(x)| \leq C(1+|x|)^p$ gilt für ein $C \geq 0$ und ein $p \in \mathbb{N}$,
(γ) in $\mathcal{E}(X)^*$, falls φ kompakten Träger hat.

(c) Man überlege sich Bedingungen an die Folgen (c_n) aus $\mathbb{C}$, (x_n) aus $\mathbb{R}^m$ und (α_n) von Multiindices so, daß

$$F(f) = \sum c_n D^{\alpha_n} f(x_n)$$

auf $\mathcal{D}(\mathbb{R}^m), \mathcal{S}(\mathbb{R}^m)$ oder $\mathcal{E}(\mathbb{R}^m)$ stetig ist.

3.9 Sei X einer der Räume $\mathcal{D}(\mathbb{R}^m), \mathcal{S}(\mathbb{R}^m), \mathcal{E}(\mathbb{R}^m)$. Für $F \in X^*$ definiert man die Ableitung $D^\alpha F$ durch

$$(D^\alpha F)(f) = (-1)^{|\alpha|} F(D^\alpha f) \text{ für } f \in X.$$

(a) Es gilt $D^\alpha F \in X^*$; die Abbildung $D^\alpha : X^* \to X^*$ ist stetig bezüglich der schwachen $*$–Topologie.

(b) Ist $F(f) = \displaystyle\int \varphi(x) f(x)\,\mathrm{d}\,x$ (vgl. Aufgabe 3.8 b) mit einer $|\alpha|$ mal stetig differenzierbaren Funktion φ, so gilt $(D^\alpha F)(f) = \displaystyle\int (D^\alpha \varphi(x)) f(x)\,\mathrm{d}x$.

4 Normierte Räume und Banachräume

4.1 Einige Begriffe

Wir erinnern zunächst an einige bereits bekannte Tatsachen: Ein Vektorraum X über $\mathbb{K}$ ausgestattet mit einer Norm $\|\cdot\|$ heißt ein *normierter Raum* $(X, \|\cdot\|)$. Die hierdurch erzeugte *(Norm–)Topologie* $\mathcal{T}_{\|\cdot\|}$ kann beschrieben werden durch die *Metrik*

$$d(x, y) = \|x - y\|.$$

Eine Nullumgebungsbasis für $\mathcal{T}_{\|\cdot\|}$ ist gegeben durch

$$U_\varepsilon = \{x \in X : \|x\| < \varepsilon\}, \quad \varepsilon > 0.$$

Eine Abbildung f von $(X, \|\cdot\|)$ in einen topologischen Raum ist genau dann *stetig* , wenn sie *folgenstetig* ist (Satz 1.7).

Eine *lineare Abbildung* L von $(X, \|\cdot\|)$ in einen topologischen Vektorraum $(T, \mathcal{T}_Y) = (Y, \|\cdot\|)$ ein normierter Raum, so ist L genau dann stetig, wenn L *beschränkt* ist (Korollar 2.10b), d. h. es gibt ein $C \geq 0$ mit

$$\|Lx\| \leq C\|x\| \text{ für alle } x \in X.$$

Eine Teilmenge von $(X, \|\cdot\|)$ ist genau dann *kompakt*, wenn sie *folgenkompakt* ist (Satz 1.13).

Ein normierter Raum $(X, \|\cdot\|)$ heißt *vollständig*, wenn jede Cauchy–Folge konvergent ist; ein vollständiger normierter Raum wird als *Banach–Raum* bezeichnet.

Eine Teilmenge A eines topologischen Raumes $(X, \mathcal{T})$ heißt *dicht* in X), wenn $\overline{A} = X$ gilt. Ist $(X, \mathcal{T})$ ein metrischer Raum (z. B. ein normierter Raum), so gilt dies nach Satz 1.10b genau dann, wenn zu jedem $x \in X$ eine Folge (x_n) aus A existiert mit $x_n \to x$; gleichbedeutend hiermit ist: zu jedem $\varepsilon > 0$ existiert ein $y \in A$ mit $d(x, y) < \varepsilon$. — Zum *Beispiel* sind die rationalen Zahlen dicht in $\mathbb{R}$, die Menge der Punkte in $\mathbb{R}^m$ mit rationalen Koordinaten ist dicht in $\mathbb{R}^m$.

Eine Teilmenge A eines topologischen Vektorraumes $(X, \mathcal{T})$ heißt *total*, wenn die lineare Hülle $L(A)$ (das ist die Menge der *endlichen* Linearkombinationen von Elementen aus A) dicht ist. Zum *Beispiel* ist die Menge A der m Einheitsvektoren aus $\mathbb{R}^m$ total in $\mathbb{R}^m$ (in diesem Fall ist sogar $L(A) = \mathbb{R}^m$, nicht nur $\overline{L(A)} = \mathbb{R}^m$).

Ein topologischer Raum heißt *separabel*, wenn es eine abzählbare dichte Teilmenge gibt.

Satz 4.1 *Jede Teilmenge eines separablen metrischen Raumes (X, d) ist separabel.*

Beweis. Sei $\{x_n : n \in \mathbb{N}\}$ dicht in X, $A \subset X$. Weiter sei J die Menge der Paare $(n, m) \in \mathbb{N}$, für die ein $y \in A$ existiert mit $d(x_n, y) \leq \frac{1}{m}$. Für jedes $(n, m) \in J$ wählen wir ein $y_{nm} \in A$ $d(x_n, y_{nm}) \leq \frac{1}{m}$. Die Menge $B := \{y_{nm} : (n, m) \in J\}$ ist abzählbar. Wir zeigen, daß $\overline{B} \supset A$ gilt: Sei $x \in A$. Da $\{x_n : n \in \mathbb{N}\}$ dicht in X ist, existiert eine Folge (n_k) aus $\mathbb{N}$ mit $d(x_{n_k}, x) \leq \frac{1}{k}$; insbesondere ist also $(n_k, k) \in J$ für jedes $k \in \mathbb{N}$ und es gilt

$$d(y_{n_k k}, x) \leq d(y_{n_k k}, x_{n_k}) + d(x_{n_k}, x) \leq \frac{2}{k} \to 0.$$

■

Satz 4.2 *Ein topologischer Vektorraum ist genau dann separabel, wenn es eine abzählbare totale Teilmenge gibt.*

Beweis. $\Rightarrow$: Dies ist offensichtlich, da eine dichte Menge auch total ist. $\Leftarrow$: Sei A abzählbar, $\overline{L(A)} = X$. Ist $L_r(A)$ die Menge der Linearkombinationen von A mit rationalen Koeffizienten, so ist offensichtlich $\overline{L_r(A)} = \overline{L(A)} = X$. Da $L_r(A)$ abzählbar ist, ist also X separabel. ■

4.2 Einfache Beispiele

Beispiel 4.3 $\mathbb{K}^m$ ($\mathbb{R}^m$ oder $\mathbb{C}^m$): Wir wissen aus Satz 2.11, daß es auf $\mathbb{K}^m$ nur eine Vektorraumtopologie gibt. Hieraus kann man folgern, daß für zwei beliebige Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ auf $\mathbb{K}^m$ stets Konstanten C_1 und C_2 existieren mit

$$\|\cdot\|_1 \leq C_1\|x\|_2, \ \|x\|_2 \leq C_2\|x\|_1 \text{ für alle } x \in \mathbb{K}^m. \qquad *$$

Jede $\|\cdot\|_1$–Nullumgebung muß eine $\|\cdot\|_2$–Nullumgebung enthalten, d. h. z. B.

$$\{x \in \mathbb{K}^m : \|x\|_1 \leq 1\} \supset \{x \in \mathbb{K}^m : \|x\|_2 \leq \varrho\} \quad (\varrho > 0);$$

aus $\|x\|_2 \leq \varrho$ folgt also $\|x\|_1 \leq 1$ und somit

$$\|x\|_1 \leq \varrho^{-1}\|x\|_2.$$

Das ist die erste Ungleichung mit $C_1 = \varrho^{-1}$. Entsprechend folgt die zweite Ungleichung. Zwei Normen, die die Beziehung $(*)$ erfüllen, heißen *äquivalent*. Äquivalente Normen haben offenbar die gleichen Cauchy–Folgen und die gleichen konvergenten Folgen.

Da $\mathbb{K}^m$ bezüglich der euklidischen Topologie vollständig und separabel ist, folgt dies hiermit für jede Vektorraumtopologie auf $\mathbb{K}^m$. Mit Korollar 2.12 folgt dies sogar für jeden endlichdimensionalen Vektorraum. □

Beispiel 4.4 $(C[a,b], \|\cdot\|_\infty)$. Sei

$C[a,b]$ der Raum der auf $[a,b]$ stetigen Funktionen,

$\|f\|_\infty := \max\{|f(x)| : x \in [a,b]\}$.

$(C[a,b], \|\cdot\|_\infty)$ *ist vollständig*: Der Beweis der Vollständigkeit eines normierten (oder metrischen) Raumes kann fast immer nach dem folgenden Schema in drei Schritten durchgeführt werden: Sei (f_n) eine $\|\cdot\|_\infty$–Cauchyfolge aus $C[a,b]$.

(i) *Konstruktion des Limes* (man gewinnt auf eine meist etwas heuristische Weise einen Kandidaten für den Grenzwert der vorgegebenen Cauchyfolge; erst die Schritte (ii) und (iii) zeigen tatsächlich, daß dies der Limes der Cauchyfolge ist): Da (f_n) eine $\|\cdot\|_\infty$–Cauchyfolge ist, existiert zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ mit

$$|f_n(x) - f_m(x)| \leq \varepsilon \text{ für alle } n, m \geq N, x \in [a,b].$$

Also ist $(f_n(x))$ eine Cauchyfolge in $\mathbb{C}$ für jedes $x \in [a,b]$, d. h. es existiert ein $f(x) \in \mathbb{C}$ mit $f_n(x) \to f(x)$.

(ii) *Es gilt* $f \in C[a,b]$: Wir haben

(α) für jedes $\varepsilon > 0$ existiert ein $N \in \mathbb{N}$ so, daß für *alle* $x \in [a,b]$ gilt

$$|f_N(x) - f(x)| \leq \frac{\varepsilon}{3}$$

(β) da f_N stetig ist, existiert ein $\delta > 0$ mit

$$|f_N(x) - f_N(y)| \leq \frac{\varepsilon}{3} \text{ für } x, y \in [a, b] \text{ mit } |x - y| < \delta.$$

Damit folgt

$$\begin{aligned} |f(x) - f(y)| &\leq |f(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f(y)| \\ &\leq 3\frac{\varepsilon}{3} = \varepsilon \text{ für } x, y \in [a, b] \text{mit} |x - y| < \delta, \end{aligned}$$

d. h. f ist stetig.

(iii) *Es gilt* $\|f_n - f\|_\infty \to 0$ *für* $n \to \infty$: Aus (i) folgt für $n \geq N$ und $m \to \infty$

$$|f_n(x) - f(x)| = \lim_{m\to\infty} |f_n(x) - f_m(x)| \leq \varepsilon \text{ für alle } x \in [a, b],$$

d. h. $\|f_n - f\|_\infty \to 0$ für $n \to \infty$.

$(C[a, b], \|\cdot\|_\infty)$ *ist separabel*: Nach dem Weierstraßschen Approximationssatz ist die (abzählbare) Menge der Monome $\{1, x, x^2, \ldots\}$ total in $C[a, b]$ bezüglich $\|\cdot\|_\infty$. Mit Satz 4.2 folgt die Behauptung. □

Besonders deutlich wird dieses Verfahren an dem folgenden etwas ausgefallenen Beispiel.

Beispiel 4.5 $(C^{k,\alpha}[a, b], \|\cdot\|_{k,\alpha})$. Es sei

$$\begin{aligned} C^{k,\alpha}[a, b] \;=\; & \text{Raum der } k\text{–mal stetig differenzierbaren Funktionen,} \\ & \text{deren } k\text{–te Ableitung hölderstetig mit Exponent } \alpha \text{ ist} \\ & (k \in \mathbb{N}_0, \alpha \in [0, 1]), \end{aligned}$$

$$\|f\|_{k,\alpha} = \sup\left\{|f(x)|, |f'(x)|, \ldots |f^{(k)}(x)|, \frac{|f^{(k)}(x) - f^{(k)}(y)|}{|x - y|^\alpha} : x, y \in [a, b], x \neq y\right\}$$

$(C^{k,\alpha}[a, b], \|\cdot\|_{k,\alpha})$ *ist vollständig*: Sei (f_n) eine $\|\cdot\|_{k,\alpha}$–Cauchyfolge.

(i) Für $\ell = 0, \ldots, k$ sind $(f_n^{(\ell)})$ offensichtlich $\|\cdot\|_\infty$–Cauchyfolgen, d. h. es gibt stetige Funktionen $f^{[\ell]}$ mit

$$f_n^{(\ell)} \to f^{[\ell]} \text{ gleichmäßig in } [a, b], \ell = 0, \ldots, k.$$

(ii) $f := f^{[0]}$ *liegt in* $C^{k,\alpha}[a, b]$: Für $\ell = 1, \ldots, k$ gilt

$$f_n^{(\ell-1)}(x) = f_n^{(\ell-1)}(a) + \int_a^x f_n^{(\ell)}(t)\,\mathrm{d}t.$$

Wegen (i) kann der Grenzübergang $n \to \infty$ durchgeführt werden und liefert

$$f^{[\ell-1]}(x) = f^{[\ell-1]}(a) + \int_a^x f^{[\ell]}(t)\,\mathrm{d}t \text{ für } \ell = 1, \ldots, k,$$

d. h. $f = f^{[0]}$ ist k–mal stetig differenzierbar und es gilt $f^{(\ell)} = f^{[\ell]}$ für $\ell = 1, \ldots, k$. Da (f_n) eine $\|\cdot\|_{k,\alpha}$–Cauchyfolge ist, gibt es ein C mit

$$|f_n^{(k)}(x) - f_n^{(k)}(y)||x-y|^{-\alpha} \leq C \text{ für alle } x, y \in [a,b], x \neq y.$$

Also ist $f \in C^{k,\alpha}[a,b]$.

(iii) *Es gilt* $\|f - f_n\|_{k,\alpha} \to 0$: Für jedes $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}$ mit $\|f_n - f_m\|_{k,\alpha} < \varepsilon$ für $n, m \geq N$. Zusammen mit (i) folgt

$$\begin{aligned} \|f_n - f\|_{k,\alpha} &= \sup\Big\{|f_n(x) - f(x)|, \ldots, |f_n^{(k)}(x) - f^{(k)}(x)|, \\ &\quad |x-y|^{-\alpha}|f_n^{(k)}(x) - f_n^{(k)}(y) - \Big(f^{(k)}(x) - f^{(k)}(y)\Big)| : x, y \in [a,b], x \neq y\Big\} \\ &\leq \lim_{m\to\infty} \sup\{|f_n(x) - f_m(x)|, \ldots : x, y \in [a,b], x \neq y\} \\ &\leq \varepsilon \text{ für } m \geq N, \end{aligned}$$

also $\|f_n - f\|_{k,\alpha} \to 0$ für $n \to \infty$. □

Beispiel 4.6 $(H_1(D), \|\cdot\|_1)$. Es sei

$$\begin{aligned} H_1(D) &= \text{Raum der auf der offenen Einheitskreisscheibe } D \text{ in} \\ &\quad \mathbb{C} \text{ holomorphen Funktionen mit } \int\limits_D |f(x+iy)|d(x,y) < \infty, \end{aligned}$$

$$\|f\|_1 = \int\limits_D |f(x+iy)|d(x,y).$$

$(H_1(D), \|\cdot\|_1)$ *ist vollständig*: Sei Γ_r die Kreislinie um 0 mit Radius r (positiv orientiert). Für jedes $\eta \in (0, \frac{1}{2})$, $z \in \mathbb{C}$ mit $|z| \leq 1 - 2\eta$ und $r \in (1-\eta, 1)$ gilt

$$f(z) = (2\pi)^{-1} \int\limits_{\Gamma_r} (\xi - z)^{-1} f(\xi)\,\mathrm{d}\xi,$$

$$|f(z)| \leq (2\pi)^{-1}(r - |z|)^{-1} \int\limits_0^{2\pi} |f(re^{i\varphi})|r\,\mathrm{d}\varphi \leq (2\pi\eta)^{-1} \int\limits_0^{2\pi} |f(re^{i\varphi})|r\,\mathrm{d}\varphi.$$

Integration bezüglich r über $(1-\eta, 1)$ liefert für alle z mit $|z| \le 1-2\eta$

$$|f(z)| \le (2\pi\eta)^{-1} \int\limits_{1-\eta\le|(x,y)|<1} |f(x+iy)|d(x,y) \le (2\pi\eta)^{-1}\|f\|_1.$$

Wendet man diese Ungleichung auf die Differenz $f_n - f_m$ zweier Folgenglieder einer $\|\cdot\|_1$–Cauchyfolge an, so folgt: Für jedes $\varepsilon > 0$ gibt es ein N mit

$$|f_n(z) - f_m(z)| \le \varepsilon \text{ für } n, m \ge N,\ |z| \le 1-2\eta.$$

Da dies jedes $\eta \in (0, \frac{1}{2})$ gilt, folgt: (f_n) konvergiert in jedem Kreis um 0 mit Radius kleiner als 1 gleichmäßig. Dieser Limes f ist (wie man aus der Funktionentheorie weiß) wieder holomorph.

Es bleibt zu zeigen $\|f_n - f\|_1 \to 0$ (insbesondere $f \in H_1(D)$): Zu jedem $\varepsilon > 0$ gibt es ein N mit $\|f_n - f_m\|_1 < \varepsilon$ für $n, m \ge N$. Also gilt für jedes $\eta \in (0,1)$ und $n \ge N$

$$\int\limits_{|(x,y)|\le\eta} |f_n(x+iy) - f(x+iy)|d(x,y) = \lim_{m\to\infty} \int\limits_{|(x,y)|\le\eta} |f_n(x+iy) - f_m(x+iy)|d(x,y)$$
$$\lim_{m\to\infty} \int\limits_D |f_n(x+iy) - f_m(x+iy)|d(x,y) \le \varepsilon.$$

Da dies für jedes $\eta \in (0,1)$ gilt, folgt $f_n - f \in H_1(D)$ und $\|f_n - f\|_1 \le \varepsilon$ für $n \ge N$. □

Beispiel 4.7 $(C[a,b], \|\cdot\|_1)$, der Raum $C[a,b]$ mit

$$\|f\|_1 = \int_a^b |f(x)|\,\mathrm{d}x.$$

Es ist leicht zu sehen, daß $\|f\|_1$ eine Norm auf $C[a,b]$ ist. Wegen $\|f\|_1 \le (b-a)\|f\|_\infty$ ist jede $\|\cdot\|_\infty$–Cauchyfolge auch $\|\cdot\|_1$–konvergent, d. h. auch $(C[a,b], \|\cdot\|_1)$ ist *separabel*. Aber es gilt:

$(C[a,b], \|\cdot\|_1)$ ist *nicht vollständig*: Die Folge

$$f_n(x) = \begin{cases} 0 & \text{für } a \le x \le c - \frac{1}{n}, \\ n(x - c + \frac{1}{n}) & \text{für } c - \frac{1}{n} \le x \le c, \\ 1 & \text{für } c \le x \le b. \end{cases}$$

mit einem $c \in (a, b)$ ist eine $\|\cdot\|_1$–Cauchyfolge, denn für $n \leq m$ gilt

$$\|f_n - f_m\|_1 = \int\limits_{c-1/n}^{c} |f_n(x) - f_m(x)| \, dx \leq \int\limits_{c-1/n}^{c} 1 \, dx = \frac{1}{n}.$$

Wir nehmen an, daß ein $f \in C[a, b]$ existiert mit $\|f_n - f\| \to 0$. Für jedes $\varepsilon > 0$ gilt

$$\int\limits_{a}^{c-\varepsilon} |f(x) - f_n(x)| \, dx = \int\limits_{a}^{c-\varepsilon} |f(x)| \, dx \text{ für } n > \frac{1}{\varepsilon},$$

also für $n \to \infty$

$$0 = \int\limits_{a}^{c-\varepsilon} |f(x)| \, dx.$$

Da f stetig ist, folgt hieraus $f(x) = 0$ in $[a, c - \varepsilon]$ für jedes $\varepsilon > 0$, und somit

$$f(x) = 0 \text{ in } [a, c).$$

Entsprechend folgt

$$f(x) = 1 \text{ in } [c, b].$$

Das ist ein Widerspruch zur Stetigkeit von f.

(Ganz entsprechende Überlegungen gelten für $\|f\|_p = \{\int |f(x)|^p \, dx\}^{1/p}$ mit $1 < p < \infty$; wir werden in § 6 sehen, daß dies auch Normen sind.) □

4.3 ℓ_p–Räume, Höldersche Ungleichung für Reihen

Die ℓ_p–Räume sind Räume von Folgen aus $\mathbb{K}$.

$\mathbf{p} = \infty :$ $$\ell_\infty := \Big\{(x_n) : \sup_n |x_n| < \infty\Big\}, \quad \|x\|_\infty := \sup_n |x_n|.$$

Offensichtlich ist ℓ_∞ mit den üblichen linearen Operationen ein $\mathbb{K}$–Vektorraum und $\|\cdot\|_\infty$ ist eine Norm ℓ_∞.

ℓ_∞ ist vollständig: Sei $(x^m) = ((x_n^m)_n)$ eine $\|\cdot\|_\infty$–Cauchyfolge in ℓ_∞ Dann ist für jedes n $(x_n^m)_m$ eine Cauchyfolge in $\mathbb{K}$, d. h. es gibt ein x_n mit $x_n^m \to x_n$ für $m \to \infty$. Für jedes $\varepsilon > 0$ existiert ein $N \in \mathbb{N}$ mit

$$|x_n^m - x_n^\ell| \leq \|x^m - x^\ell\|_\infty \leq \varepsilon \text{ für } n, \ell \geq N, m \in \mathbb{N}.$$

Daraus folgt

$$\sup_n |x_n| \leq \sup_n |x_n^N| + \varepsilon \leq \|x^N\|_\infty + \varepsilon,$$

d. h. es gilt $(x_n) \in \ell_\infty$. Für $m \geq N$ und $\ell \to \infty$ folgt

$$\|x^m - x\|_\infty = \sup_n |x_n^m - x_n| \leq \varepsilon, \text{ also } x^m \to x.$$

ℓ_∞ ist nicht separabel: Die Menge M der 0–1–Folgen ist überabzählbar und für $x, y \in M$, $x \neq y$ gilt $\|x - y\|_\infty = 1$. Dann kann es aber keine abzählbare dichte Teilmenge geben (vgl. Aufgabe 4.4).

$$\mathbf{p} \in [\mathbf{1}, \infty): \quad \ell_p := \Big\{(x_n) : \sum_n |x_n|^p < \infty\Big\}, \quad \|x\|_p := \{\sum_n |x_n|^p\}^{1/p}.$$

ℓ_p ist ein Vektorraum: die einzige nichttriviale Tatsache ist hierbei: Aus $x, y \in \ell_p$ folgt $x + y \in \ell_p$. Dies ergibt sich aus

$$\begin{aligned} \sum_n |x_n + y_n|^p &\leq \sum_n \Big\{|2x_n|^p + |2y_n|^p\Big\} \\ &= 2^p \Big\{\sum_n |x_n|^p + \sum_n |y_n|^p\Big\} < \infty. \end{aligned}$$

$\|\cdot\|_p$ ist eine Norm auf $\ell_\mathbf{p}$: Dabei ist nur die Dreiecksungleichung nicht–trivial. Um sie beweisen zu können, benötigen wir die Höldersche Ungleichung.

Satz 4.8 (Höldersche Ungleichung) *Seien $p, q \in [1, \infty]$ mit $\frac{1}{p} + \frac{1}{q} = 1$, dabei sind auch die Fälle $p = 1, q = \infty$ und $p = \infty, q = 1$ zugelassen. Für $x \in \ell_p$ und $y \in \ell_q$ ist die Folge $(x_n y_n) \in \ell_1$ und es gilt*

$$\sum |x_n y_n| \leq \|x\|_p \|y\|_q.$$

Zum Beweis benötigen wir das

Lemma 4.9 *Seien $p, q \in (1, \infty)$ mit $\frac{1}{p} + \frac{1}{q} = 1$, $a \geq 0$ und $b \geq 0$. Dann gilt*

$$ab \leq \frac{1}{p}a^p + \frac{1}{q}b^q$$

Beweis. Wir zeigen zunächst für $0 < \alpha < 1$, $c \geq 0$, $d \geq 0$ die Ungleichung

$$c^\alpha d^{1-\alpha} \leq \alpha c + (1 - \alpha)d.$$

Für $c = 0$ und $c = d$ ist die Ungleichung trivial. Sei also o. E. $0 < c < d$. Aus dem Mittelwertsatz folgt

$$\begin{aligned} d^{1-\alpha} - c^{1-\alpha} &= (1-\alpha)(d-c)\xi^{-\alpha} \text{ mit } c < \xi < d \\ &\leq (1-\alpha)(d-c)c^{-\alpha}. \end{aligned}$$

Multiplikation mit c^α ergibt

$$c^\alpha d^{1-\alpha} \leq c + (1 - \alpha](d - c) = \alpha c + (1 - \alpha)d.$$

Wenden wir die hiermit bewiesene Ungleichung an auf $\alpha = \frac{1}{p}$, $1 - \alpha = \frac{1}{q}$, $c = a^p$, $d = b^q$, so folgt die Behauptung des Lemmas. ∎

Beweis von Satz 4.8. Die Fälle $p = 1$, $q = \infty$ und $p = \infty$, $q = 1$ sind trivial. Sei also $1 < p$, $q < \infty$. Ist $x = 0$ *oder* $y = 0$, so ist die Ungleichung (Gleichung!) offensichtlich. Ohne Einschränkung können wir also $x \neq 0$ und $y \neq 0$ annehmen. Anwendung von Lemma 4.9 auf $a = \|x\|_p^{-1}|x_n|$ *und* $b = \|y\|_q^{-1}|y_n|$ ergibt

$$(\|x\|_p\|y\|_q)^{-1}|x_n y_n| \leq \frac{1}{p}\|x\|_p^{-p}|x_n|^p + \frac{1}{q}\|y\|_q^{-q}|y_n|^q.$$

Durch Summation über n folgt

$$\begin{aligned} \sum_n |x_n y_n| &= \|x\|_p\|y\|_q \sum (\|x\|_p\|y\|_q)^{-1}|x_n y_n| \\ &\leq \|x\|_p\|y\|_q \Big\{\frac{1}{p}\|x\|_p^{-p}\|x\|_p^p + \frac{1}{q}\|y\|_q^{-q}\|y\|_q^q\Big\} = \|x\|_p\|y\|_q. \end{aligned}$$

∎

Satz 4.10 (Minkowski–Ungleichung) *Für $x, y \in \ell_p$ gilt die* Minkowski–Ungleichung, Dreiecksungleiching,

$$\|x+y\|_p \leq \|x\|_p + \|y\|_p \qquad (p \in [1,\infty)).$$

Beweis. $p = 1$ und $p = \infty$ sind offensichtlich (für $p = \infty$ haben wir dies bereits oben benutzt). Sei also $1 < p < \infty$. Dann gilt wegen der Hölderschen Ungleichung

$$\begin{aligned}
\|x+y\|_p^p &= \sum |x_n+y_n|^p \leq \sum \Big\{|x_n||x_n+y_n|^{p-1} + |y_n||x_n+y_n|^{p-1}\Big\} \\
&\leq \Big\{\sum |x_n|^p\Big\}^{1/p}\Big\{\sum |x_n+y_n|^{(p-1)q}\Big\}^{1/q} + \Big\{\sum |y_n|^p\Big\}^{1/p}\Big\{\sum |x_n+y_n|^{(p-1)q}\Big\}^{1/q} \\
&= \Big(\|x\|_p + \|y\|_p\Big)\|x+y\|_p^{p/q} \text{ (wegen } (p-1)q = p),
\end{aligned}$$

$\|x+y\|_p = \|x+y\|_p^{p-p/q} = \|x\|_p + \|y\|_p$.

■

$\ell_p (1 \leq p < \infty)$ *ist vollständig*: Sei (x^m) eine $\|\cdot\|_p$–Cauchyfolge in ℓ_p. Dann ist für jedes n $(x_n^m)_m$ eine Cauchyfolge in $\mathbb{K}$, d. h. es gibt ein x_n mit $x_n^m \to x_n$ für $m \to \infty$. Für jedes $\varepsilon > 0$ existiert ein $N \in \mathbb{N}$ mit

$$\|x^m - x^\ell\|_p \leq \varepsilon \text{ für } m, \ell \geq N.$$

Daraus folgt für jedes $K \in \mathbb{N}$

$$\sum_{n=1}^K |x_n^m - x_n|^p = \lim_{\ell\to\infty} \sum_{n=1}^K |x_n^m - x_n^\ell|^p \leq \varepsilon^p,$$

also auch

$$\sum_{n=1}^\infty |x_n^m - x_n|^p \leq \varepsilon^p,$$

d. h. es gilt $(x_n) \in \ell_p$ und $x^m \to x$ für $m \to \infty$.

$\ell_p (1 \leq p < \infty)$ *ist separabel*: Die Menge der Einheitsvektoren ist abzählbar und total in ℓ_p.

Aus der Hölderschen Ungleichung folgt, daß jedes $y \in \ell_q$ ($\frac{1}{p} + \frac{1}{q} = 1$ für $1 < p < \infty$, $q = \infty$ für $p = 1$, und $q = 1$ für $p = \infty$) ein *stetiges lineares Funktional*

$$F_y : \ell_p \to \mathbb{K}, x \mapsto \sum_n x_n y_n$$

erzeugt mit $\|F_y\| \leq \|y\|_q$. Es ist leicht zu zeigen, daß sogar $\|F_y\| = \|y\|_q$ gilt; mehr dazu in § 6 und § 10.

4.4 Beschränkte lineare Operatoren und Funktionale

Seien $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ normierte Räume (wenn keine Verwechslung möglich ist, lassen wir den Index X bzw. Y an den Normen weg). $B(X;Y)$ ist der *Raum der beschränkten* (stetigen) *linearen Operatoren* von X nach Y (auf ganz X definiert). Durch

$$\begin{aligned}\|A\| &:= \sup\Big\{\|Ax\| : x \in X, \|x\| \leq 1\Big\} \\ &= \inf\Big\{C : \|Ax\| \leq C\|x\| \text{ für alle } x \in X\Big\}\end{aligned}$$

wird auf $B(X;Y)$ eine *Norm* definiert: $\|A\| \geq 0$ ist offensichtlich auf Grund der Definition.

(N 1)
$$\begin{aligned}\|aA\| &= \sup\{\|aAx\| : x \in X, \|x\| \leq 1\} \\ &= |a| \sup\{\|Ax\| : x \in X, \|x\| \leq 1\} = |a|\|A\|.\end{aligned}$$

(N 2)
$$\begin{aligned}\|A+B\| &= \sup\{\|Ax+Bx\| : x \in X, \|x\| \leq 1\} \\ &\leq \sup\{\|Ax\| : x \in X, \|x\| \leq 1\} + \sup\{\|Bx\| : x \in X, \|x\| \leq 1\} \\ &= \|A\| + \|B\|.\end{aligned}$$

(N 3) Aus $\|A\| = 0$ folgt $\|Ax\| = 0$ für alle $x \in X$, also $A = 0$.

Offensichtlich gilt außerdem

$$\|Ax\| \leq \|A\|\|x\| \text{ für alle } x \in X$$

und

$$\|AB\| \leq \|A\|\|B\|;$$

letzteres folgt aus

$$\|ABx\| \leq \|A\|\|Bx\| \leq \|A\|\|B\|\|x\|.$$

Satz 4.11 *Ist* $(X, \|\cdot\|_X)$ *normiert und* $(Y, \|\cdot\|_Y)$ *ein Banachraum, so ist auch* $(B(X,Y), \|\cdot\|)$ *ein Banachraum. Speziell ist* $(X, \|\cdot\|_X)^*$ *stets ein Banachraum.*

Beweis. Nur die Vollständigkeit ist noch zu beweisen. Sei (A_n) eine Cauchyfolge in $(B(X,Y), \|\cdot\|)$. Dann existiert zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ mit

$$\|A_n x - A_m x\| \leq \varepsilon \|x\| \text{ für } n, m \geq N \text{ und alle } x \in X.$$

Also ist $(A_n x)$ für jedes $x \in X$ eine Cauchyfolge in $(Y, \|\cdot\|_Y)$, d. h. es existiert ein $y \in Y$ mit $A_n x \to y$; wir definieren $Ax := \lim A_n x$. Die Abbildung $A : X \to Y$ ist offenbar linear und es gilt

$$\|Ax\| = \lim_n \|A_n x\| \leq C\|x\|, \text{ d. h. } A \in B(X;Y)$$

(da (A_n) eine Cauchyfolge ist, existiert ein C mit $\|A_n\| \leq C$ für alle N). Für $n \geq N$ und $m \to \infty$ folgt

$$\|A_n x - Ax\| = \lim_m \|A_n x - A_m x\| \leq \varepsilon \|x\|,$$
$$\|A_n - A\| \leq \varepsilon \text{ für } n \geq N,$$

also $\|A_n - A\| \to 0$. □

Bemerkung 4.12 *Zur Berechnung der Norm* eines Operators $A \in B(X,Y)$ bzw. eines $F \in X^*$ ist folgende Äquivalenzaussage nützlich:

$$\|A\| = C \iff \begin{array}{l} \text{(i) } \|Ax\| \leq C\|x\| \text{ für alle } x \in X \ (\text{d. h. } \|A\| \leq C); \\ \text{(ii) es existiert eine Folge } (x_n) \text{ aus } x \in X\backslash\{0\} \text{ mit} \\ \qquad \lim \|Ax_n\| \|x_n\|^{-1} \geq C \text{ d. h. } \|A\| \geq C). \end{array}$$

Statt (ii) gelingt es häufig sogarzu zeigen: Es gibt ein $x \in X \setminus \{0\}$ mit $\|Ax\| = C\|x\|$.

Beispiel 4.13 $A : C[0,1] \to C[0,1], \quad Af(x) = xf(x).$
Wir untersuchen A in Bezug auf die Normen $\|\cdot\|_p$ für $1 \leq p \leq \infty$. In jedem Fall gilt offensichtlich

$$\|Af\|_p \leq \|f\|_p \text{ für alle } f \in C[0,1];$$

insbesondere ist $A \in B(C[0,1]) := B(C[0,1], C[0,1])$.

$\mathbf{p} = \infty$: Ist $f \in C[0,1]$ so, daß $\max |f(x)| = |f(1)|$ ist, so gilt

$$\|Af\|_\infty = \max\Big\{|xf(x)| : 0 \le x \le 1\Big\} = |f(1)| = \|f\|_\infty,$$

also $\|A\| \ge 1$ und somit $\|A\| = 1$.

$\mathbf{1} \le \mathbf{p} < \infty$: Sei $\|f_n\|_p = 1$ mit $f_n(x) = 0$ in $[0, 1 - \frac{1}{n}]$. Dann ist

$$\|Af_n\|_p = \Big\{\int\limits_0^1 |xf(x)|^p \,\mathrm{d}x\Big\}^{1/p} \ge (1 - \frac{1}{n})\|f_n\|_p = (1 - \frac{1}{n}) \to 1,$$

also gilt auch in diesem Fall $\|A\| = 1$. □

4.5 Teilräume, Produkträume, Quotientenräume

Aus vorgegebenen normierten oder Banach–Räumen können auf verschiedenen Arten neue normierte Räume konstruiert werden. In Abschnitt 4.4 lernten wir bereits $B(X;Y)$ und $B(X;\mathbb{K}) = X^*$ kennen, den Raum der stetigen linearen Abbildungen von X nach Y bzw. von X nach $\mathbb{K}$; ist Y ein Banachraum, so ist $B(X;Y)$ ebenfalls ein Banachraum; insbesondere ist X^* stets ein Banachraum

In diesem Abschnitt lernen wir noch drei weitere Verfahren zur Konstruktion neuer Räume kennen (in Abscnitt 4.6 werden wir dann noch die Vervollständigung behandeln).

Sei $(X, \|\cdot\|)$ ein normierter Raum, T ein *Teilraum* (Untervektorraum) von X. $\|\cdot\|$ ist natürlich auch auf T eine Norm und somit ist $(T, \|\cdot\|)$ ein normierter Raum.

Satz 4.14 *Sei $(X, \|\cdot\|)$ ein Banachraum, T ein Teilraum von X. T ist genau dann abgeschlosen, wenn $(T, \|\cdot\|)$ ein Banachraum ist.*

Beweis. $\Rightarrow$: Sei (x_n) eine Cauchyfolge in $(T, \|\cdot\|)$. Dann ist (x_n) eine Cauchyfolge in $(X, \|\cdot\|)$; also existiert ein $x \in X$ mit $x_n \to x$. Da T abgeschlossen ist, gilt $x \in T$.

$\Leftarrow$: Sei (x_n) eine Folge aus $T, x_n \to x \in X$. Dann ist (x_n) eine Cauchyfolge in $(T, \|\cdot\|)$ und somit in T konvergent, d. h. $x \in T$. □

Seien $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ normierte Räume. Das *Produkt* $X \times Y$, d. h. die Menge der Paare (x, y) mit $x \in X$ und $y \in Y$ wird durch

$$(x, y) + (x', y') = (x + x', y + y'), \quad a(x, y) = (ax, ay)$$

zu einem Vektorraum. Mit

$$\|(x, y)\| = \|x\|_X + \|y\|_Y$$

wird auf $X \times Y$ eine Norm definiert , also $(X \times Y, \|\cdot\|)$ zu einem normierten Raum. Die Normen

$$\|(x, y)\|_p = \begin{cases} (\|x\|_X^p + \|y\|_Y^p)^{1/p} & \text{für } 1 < p < \infty, \\ \max\{\|x\|_X, \|y\|_Y\} & \text{für } p = \infty \end{cases}$$

sind offenbar äquivalent zu $\|\cdot\|$.

Satz 4.15 *Die Räume* $(X, \|\cdot\|_X)$ und $Y, \|\cdot\|_Y)$ *sind genau dann Banachräume, wenn* $(X \times Y, \|\cdot\|)$ *ein Banachraum ist.*

Beweis. $\Rightarrow$: Ist $\big((x_n, y_n)\big)$ eine Cauchyfolge in $(X \times Y, \|\cdot\|)$, so sind (x_n) und (y_n) Cauchyfolgen in $(X, \|\cdot\|_X)$ bzw. $(Y, \|\cdot\|_Y)$, d. h. es gibt $x \in X$ und $y \in Y$ mit $x_n \to x$ und $y_n \to y$. Dann gilt aber auch $(x_n, y_n) \to (x, y)$.

$\Leftarrow$: Sei (x_n) Cauchyfolge in $(X, \|\cdot\|_X), y_0 \in Y$. Dann ist $\big((x_n, y_0)\big)$ eine Cauchyfolge in $(X \times Y, \|\cdot\|)$, d. h. es gibt ein $x \in X$ und ein $\overline{y_0} \in Y$ mit $(x_n, y_0) \to (x, \overline{y_0})$, also $x_n \to x$ (und natürlich auch $\overline{y_0} = y_0$). Entsprechend zeigt man, daß Y vollständig ist. □

Sei X ein Vektorraum, T ein Teilraum von X. Durch $x \sim_T y \Leftrightarrow x - y \in T$ wird offenbar eine *Äquivalenzrelation* auf X erklärt (für alle x gilt $x \sim_T x$; aus $x \sim_T y$ und $y \sim_T z$ folgt $x \sim_T z$; aus $x \sim_T y$ folgt $y \sim_T x$). Mit $\hat{x}$ bezeichnen wir die zu x gehörige Äquivalenzklasse. Durch

$$\hat{x}_1 + \hat{x}_2 = \widehat{x_1 + x_2}, \quad a\hat{x}_1 = \widehat{ax_1} \qquad (x_j \in X,\ a \in \mathbb{K})$$

wird die Menge der Äquivalenzklassen zu einem Vektorraum. Dieser wird als *Quotientenraum* X/T bezeichnet.

Satz 4.16 *Sei* $(X, \|\cdot\|)$ *ein normierter Raum,* T *ein abgeschlossener Teilraum von* X*.*

(a) $\|\hat{x}\| := \inf\{\|z\| : z \in \hat{x}\}$ *ist eine Norm auf* X/T.

(b) Ist X *ein Banachraum, so ist auch* $(X/T, \|\cdot\|)$ *ein Banachraum.*

(c) Die Abbildung $X \to X/T,\ x \mapsto \hat{x}$ *hat die Norm 1.*

Beweis. a) Natürlich ist $\|\hat{x}\| \geq 0$

(N 1)
$$\begin{aligned} \|a\hat{x}\| &= \|\widehat{ax}\| = \inf\{\|z\| : z \in (a\hat{x})\} \\ &= \inf\{\|az\| : z \in \hat{x}\} = |a|\|\hat{x}\|. \end{aligned}$$

(N 2)
$$\begin{aligned} \|\hat{x} + \hat{y}\| &= \|\widehat{x+y}\| = \inf\{\|z\| : z \in \widehat{x+y}\} \\ &= \inf\{\|(u+v)\| : u \in \hat{x},\ v \in \hat{v}\} \\ &\leq \inf\{\|u\| + \|v\| : u \in \hat{x}, v \in \hat{y}\} = \|\hat{x}\| + \|\hat{y}\|. \end{aligned}$$

(N 3) Ist $\|\hat{x}\| = 0$, so existiert eine Folge (z_n) aus $\hat{x}$ mit $z_n \to 0$. Da T abgeschlossen ist, ist auch $\hat{x} = x + T$ abgeschlossen, also $0 \in x + T$. Das ist nur möglich, wenn $x \in T$ ist, also $\hat{x} = \hat{0} \in X/T$.

b) Sei $(\hat{x}_n)$ eine Cauchyfolge in $(X/T, \|\cdot\|)$. Dann existiert eine Teilfolge $(\hat{x}_{n_k})$ mit

$$\|\hat{x}_n - \hat{x}_{n_k}\| < 2^{-k} \text{ für } n \geq n_k,$$

also speziell

$$\|\hat{x}_{n_{k+1}} - \hat{x}_{n_k}\| < 2^{-k}.$$

Nach Definition der Norm auf X/T existiert für jedes $k \in \mathbb{N}$ ein $y_k \in \hat{x}_{n_{k+1}} - \hat{x}_{n_k}$ mit $\|y_k\| < 2^{-k}$. Deshalb ist $\sum y_k$ konvergent; sei $y := \sum y_k$. Wegen

$$\|(\hat{y}_1 + \ldots + \hat{y}_k) - \hat{y}\| \leq \|(y_1 + \ldots + y_k) - y\|$$

ist dann auch $\sum \hat{y}_k$ konvergent gegen $\hat{y}$,

$$\hat{x}_{n_{k+1}} - \hat{x}_{n_1} = (\hat{x}_{n_2} - \hat{x}_{n_1}) + \ldots + (\hat{x}_{n_{k+1}} - \hat{x}_{n_k}) = \hat{y}_1 + \ldots + \hat{y}_k \to \hat{y},$$

$$\hat{x}_{n_{k+1}} \to \hat{y} + \hat{x}_{n_1}.$$

Da $(\hat{x}_n)$ eine Cauchyfolge ist, gilt dann auch für die gesamte Folge

$$\hat{x}_n \to \hat{y} + \hat{x}_{n_1}.$$

■

4.6 Vervollständigung metrischer und normierter Räume

Oft ist es nützlich zu wissen, daß jeder metrische Raum (normierte Raum) X als ein dichter Teilraum eines vollständigen metrischen Raumes (Banach–Raumes $\tilde{X}$ aufgefaßt werden kann. Ein solcher Raum $\tilde{X}$ wird als *Vervollständigung* von X bezeichnet.

Satz 4.17 *Sei (X, d) ein metrischer Raum ($(X, \|\cdot\|)$ ein normierter Raum). Dann existiert ein bis auf Isomorphie eindeutig bestimmter vollständiger metrischer Raum $(\tilde{X}, \tilde{d})$ (Banachraum $(\tilde{X}, |||\cdot|||)$) so, daß X zu einem dichten Teilraum von $\tilde{X}$ isomorph ist. (Als* Isomorphie *wird hier eine bijektive Abbildung bezeichnet, die im Fall des metrischen Raumes die Abstände erhält, im Fall des normierten Raumes linear ist und die Norm erhält.)*

Beweis. *Konstruktion von $\tilde{X}$:* Sei zunächst $\hat{X}$ die Menge der Cauchyfolgen in X. Wir sagen zwei Cauchyfolgen (x_n) und (y_n) sind *äquivalent*, wenn gilt $d(x_n, y_n) \to 0$ bzw. $\|x_n - y_n\| \to 0$. Mit $[(x_n)]$ bezeichnen wir die Äquivalenzklasse, die die Cauchyfolge (x_n) enthält. $\tilde{X}$ sei die Menge der Äquivalenzklassen aus $\hat{X}$. Im Falle des normierten Raumes X sind die Vektorraumoperationen auf $\tilde{X}$ definiert durch

$$[(x_n)] + [(y_n)] = [(x_n + y_n)], \qquad a[(x_n)] = [(ax_n)].$$

Schließlich definieren wir

$$\tilde{d}([(x_n)] + [(y_n)] := \lim_n d(x_n, y_n) \text{ bzw.}$$
$$|||[(x_n)]||| := \lim_n \|x_n\|.$$

Es ist leicht zu sehen, daß diese Definition von der Wahl der Repräsentanten unabhängig ist und daß hierdurch eine Metrik bzw. eine Norm definiert wird.

Wir betten X in $\tilde{X}$ ein durch

$$E : X \to \tilde{X}, x \mapsto [x],$$

wobei $[x]$ die Äquivalenzklasse der gegen X konvergenten Folgen ist. Offenbar gilt

$$\tilde{d}(Ex, Ey) = d(x, y), \text{ bzw. } |||Ex||| = \|x\|.$$

$E(X)$ *ist dicht in* $\tilde{X}$ und im Vektorraum–Fall ist E linear: Sei $\tilde{x} = [(x_n)] \in \tilde{X}$. Dann gilt offenbar $[x_m] \to \tilde{x}$, denn

$$\tilde{d}([x_m], \tilde{x}) = \tilde{d}([x_m], [(x_n)] = \lim_n d(x_m, x_n) \to 0 \text{ für } m \to \infty, \text{ bzw.}$$

$$|||[x_m] - \tilde{x}||| = \lim_n \|x_m - x_n\| \to 0.$$

$\tilde{X}$ *ist vollständig*: Sei $(\tilde{x}_k)$ eine Cauchyfolge in $\tilde{X}$. Da $E(X)$ dicht ist, existiert für jedes k ein $x_k \in X$ mit $|||\tilde{x}_k - [x_k]||| \leq \frac{1}{k}$ (wir betrachten von jetzt ab nur noch den Fall des normierten Raumes X; für metrische Räume geht es entsprechend). Wegen

$$\begin{aligned} \|x_k - x_\ell\| &= |||[x_k] - [x_\ell]||| \\ &\leq |||[x_k] - \tilde{x}_k||| + |||\tilde{x}_k - \tilde{x}_\ell||| + |||\tilde{x}_\ell - [x_\ell]||| \\ &\leq \frac{1}{k} + |||\tilde{x}_k - \tilde{x}_\ell||| + \frac{1}{\ell} \end{aligned}$$

ist (x_k) eine Cauchyfolge in X, also $[(x_k)] \in \tilde{X}$ und

$$\begin{aligned} |||[(x_n)] - \tilde{x}_k||| &\leq |||[(x_n)] - [x_k]||| + |||[x_k] - \tilde{x}_k||| \\ &\leq \lim_n \|x_n - x_k\| + \frac{1}{k} \to 0 \text{ für } k \to \infty, \end{aligned}$$

d. h. $\tilde{x}_k \to [(x_n)]$ für $k \to \infty$.

Wir zeigen nun noch, daß zwei Vervollständigungen $\hat{X}$ und $\tilde{X}$ *isomorph* sind. Seien $\hat{E}$ und $\tilde{E}$ die entsprechenden Einbettungen, d. h. $\hat{E}(X)$ ist dicht in $\hat{X}$, $\tilde{E}(X)$ ist dicht in $\tilde{X}$. Dann ist $\hat{E}\tilde{E}^{-1}$ eine Isomorphie von $\tilde{E}(X)$ auf $\hat{E}(X)$. Wir zeigen, daß $\hat{E}\tilde{E}^{-1}$ zu einem Isomorphismus von $\tilde{X}$ auf $\hat{X}$ fortgesetzt werden kann.

Sei $\tilde{x} \in \tilde{X}$. Dann existiert eine Folge (x_n) aus X mit $\tilde{E}(x_n) \to \tilde{x}$, also ist

$$\hat{E}\tilde{E}^{-1}(\tilde{E}(x_n)) = \hat{E}(x_n) \text{ eine Cauchyfolge in } \hat{X},$$

$$\hat{E}\tilde{E}^{-1}(\tilde{E}(x_n)] \to \hat{x}, \quad |||\hat{x}||| = |||\tilde{x}|||.$$

Wir definieren $V : \tilde{X} \to \hat{X}$ durch

$$V\tilde{x} := \hat{x}.$$

V erhält offenbar die Norm.

$V(\tilde{X})$ *ist vollständig*: Ist $(\hat{x}_n)$ eine Cauchyfolge in $V(\hat{X})$, so ist $V^{-1}\hat{x}_n$ eine Cauchyfolge in $\tilde{X}$, also konvergent, $V^{-1}\hat{x}_n \to \tilde{x}, \hat{x}_n \to V\tilde{x} \in V(\tilde{X})$.

Wegen $\tilde{E}(X) \subset V(\tilde{X})$ ist $V(\tilde{X})$ dicht in $\tilde{X}$, also (da $V(\tilde{X})$ vollständig ist) notwendig gleich $\hat{X}$, d. h. V ist ein Isomorphismus von $\tilde{X}$ auf $\hat{X}, \tilde{X}$ und $\hat{X}$ sind isomorph. ■

4.7 Übungsaufgaben

4.1 Sei $(X, \|\cdot\|)$ ein normierter Raum T ein abgeschlossener Teilraum, N ein 1–dimensionaler Teilraum, $T \cap N = 0$.

(a) Ist $(x_n + z_n)(x_n \in T, \in N)$ eine in X konvergente Folge aus $T + N$, so sind (x_n) und (z_n) konvergent.

(b) Ist M ein endlich dimensionaler Teilraum von X, so ist $T+M$ abgeschlossen (Induktion nach $\dim M$).

(c) Teil b) gilt i. allg. nicht, wenn M ein abgeschlossener unendlichdimensionaler Teilraum ist.

Anleitung: In ℓ_1 wähle man
$T = \Big\{(x_n) : x_{2n} = 0 \text{ für } n \in \mathbb{N}\Big\}$, $M = \Big\{(x_n) : x_{2n} = \dfrac{1}{n}x_{2n-1} \text{ für } n \in \mathbb{N}\Big\}$.

4.2 Sei $(X, \|\cdot\|)$ ein normierter Raum, T ein abgeschlossener Teilraum vom $X, T \neq X$.

(a) Ist $z \in X \backslash T$, so gibt es ein stetiges lineares Funktional F auf LT, z mit $\|F = 1$ und $F(x) = 0$ für $x \in T$. (Man benutze Aufgabe 4.1 a).

(b) Zu jedem $delta < 0$ gibt es ein $x \in X$ mit $\|x\| = 1$ und $d(x, T) := \inf\{\|x - y\| : y \in T\} \geq 1 - \delta$.

(c) Genau dann sind alle beschränkten Mengen in X relativ kompakt, wenn X endlichdimensional ist.

(d) Ist T endlichdimensional, so gilt Teil b) auch mit $\delta = 0$.

4.3 Man zeige die Behauptung von Aufgabe 4.2b (*Rieszsches Lemma*) direkt. Anleitung: a) für $y \notin x \backslash T$ ist $d := \inf\{\|y - x\| : x \in T\} < 0$.
b) Es existiert ein $z \in T$ mit $\|y - z\| < \dfrac{d}{1-\delta}$
c) $x = \|y - z\|^{-1}(y - z)$ hat die gewünschte Eigenschaft.

4.4 Ein normierter Raum $(X, \|\cdot\|)$ ist genau dann *nicht* separabel, wenn es eine überabzählbare Teilmenge A von X und ein $a > 0$ gibt mit $\|x - y\| \geq a$ für $x, y \in A, x \neq y$.

(a) Der Raum $(C(\mathbb{R}), \|\cdot\|_\infty$ der stetigen beschränkten Funktionen auf $\mathbb{R}$ mit $\|f\|_\infty := \sup\{|f(x)| : x \in \mathbb{R}\}$ ist ein Banach–Raum.

(b) $(C(\mathbb{R}), \|\cdot\|_\infty)$ ist nicht separabel
Anleitung: Man zeige, daß ℓ_∞ als Teilraum aufgefaßt werden kann.

4.5 Sei c_0 bzw. c der Raum der Nullfolgen bzw. der konvergenten Folgen $x = (x_n)_{n\in\mathbb{N}}$ mit der Norm $\|x\|_\infty := \sup\{|x_n| : n \in \mathbb{N}\}$.

(a) c_0 und c mit dieser Norm sind Banachräume. Es gilt $c = c_0 + L\{(1, 1, 1, \ldots)\}$.

(b) Die stetigen Funktionale F auf c_0 sind gegeben durch
$F(x) = \sum\limits_{n\in\mathbb{N}} f_n x_n + f_0 \lim\limits_{n\to\infty} x_n$ mit $f = (f_n)_{n\in\mathbb{N}} \in \ell_1(\mathbb{N}_0)$, $\|f\|_1 = \|F\|$.

(c) Die stetigen Funktionale F auf c sind gegeben durch
$F(x) = \sum\limits_{n\in\mathbb{N}} f_n x_n$ mit $f = (f_n)_{n\in\mathbb{N}_0} \in \ell_1(\mathbb{N}), \|f\|_1 = \|F\|$.

Bemerkung. Die abgekürzte Aussage $c^* = c_0^* = \ell_1$ ist also mißverständlich.

4.6 (a) Es gibt ein $F \in \ell_\infty^*$ mit $F(x) = \lim x_n$ für alle $x \in c$ (vgl. Aufgabe 4.6).

(b) Nicht jedes stetige lineare Funktional auf ℓ_∞ läßt sich durch ein $f \in \ell_1$ darstellen.

4.7 Eine Teilmenge B eines Vektorraumes X heißt eine algebraische Basis (auch: /*Hamelsche Basis*), wenn B linear unabhängig ist (d. h. jede endliche Teilmenge ist linear unabhängig) und $L(B) = X$.

(a) Jedes maximale linear unabhängige System in X ist eine algebraische Basis.

(b) Mit Hilfe des Zornschen Lemmas (vgl. Abschnitt 3.2) beweise man, daß jeder Vektorraum eine algebraische Basis besitzt.

4.8 (a) Eine algebraische Basis eines Banachrumes X ist endlich oder überabzählbar.
Anleitung: Die Existenz einer abzählbaren Basis liefert zusammen mit dem Satz von Baire: $\dim X < \infty$.

(b) Es gibt normierte Räume mit abzählbarer algebraischer Basis.

(c) Zwei algebraische Basen eines Vektorraumes haben die gleiche Mächtigkeit.

4.9 Man zeige, daß $d_p (1 < p < \infty)$ aus Beispiel 1.1 Metriken sind.

4.10 Sei $1 < p < \infty, \dfrac{1}{p} + \dfrac{1}{q} = 1$.

(a) Für die unendliche Matrix $(a_{nm})_{n,m\in\mathbb{N}}$ gelte $a_{nm} = b_{nm}c_{nm}$ mit

$$\sum_m |b_{nm}|^q \le b^q \text{ für alle } n,$$

$$\sum_n |c_{nm}|^p \le c^p \text{ für alle } m.$$

Dann wird durch $Ax := (\sum\limits_m a_{nm}x_m)_{n\in\mathbb{N}}$ ein Operator $A \in B(\ell_p)$ erzeugt mit $\|A\| \le bc$.

(b) Die Matrix (a_{nm}) mit

$$a_{nm} = \begin{cases} n^{-1} & \text{für } m \le n, \\ 0 & \text{sonst} \end{cases}$$

erzeugt einen beschränkten Operator A in ℓ_2 mit $\|A\| \le \sqrt{6}$ (tatsächlich gilt $\|A\| = 2$).

5 Lebesguesche Integrationstheorie

5.1 Prämaße und Nullmengen

Sei X eine beliebige Menge. Eine Familie $\mathcal{R}$ von Teilmengen von X heißt ein *Mengenring* in X, wenn gilt

$$\emptyset \in \mathcal{R}, \text{aus } A, B \in \mathcal{R} \text{ folgt } A \cup B \in \mathcal{R} \text{ und } A \setminus B (:= A \cap \complement B) \in \mathcal{R}$$

$$(\text{dann gilt auch: } A \cap B = A \setminus (A \setminus B) \in \mathcal{R}).$$

Zum Beispiel ist in $\mathbb{R}^m$ die Familie der Mengen, die sich aus endlich vielen Quadern zusammensetzten, gelegentlich *Figuren* genannt, ein Mengenring (dabei dürfen die Quader auch entartet, d. h. niedrigerdimensional sein).

Sei $\mathcal{R}$ ein Mengenring in X. Eine Funktion $\mu : \mathcal{R} \to [0, \infty)$ heißt ein *Prämaß* auf $(X, \mathcal{R})$, wenn gilt:

$$\text{aus } A_n \in \mathcal{R} \quad (n \in \mathbb{N}),\ A_n \cap A_m = \emptyset \text{ für } n \neq m,\ \cup A_n \in \mathcal{R}$$

$$\text{folgt } \mu(\bigcup_n A_n) = \sum_n \mu(A_n) \qquad\qquad (\sigma\text{–}\textit{additiv}).$$

Speziell gilt also: $\mu(\emptyset) = 0$; aus $A, B \in \mathcal{R}$, $A \cap B = \emptyset$ folgt $\mu(A \cup B) = \mu(A) + \mu(B)$.

(Man beachte, daß hier (im Gegensatz zu später beim Maß) stets vorausgesetzt wird, daß $A := \cup_n A_n \in \mathcal{R}$ ist.)

Beispiel 5.1 In $\mathbb{R}^m$ sei $\mu(A) :=$ Volumen von A, für jede Figur A in $\mathbb{R}^m$. A läßt sich darstellen als disjunkte Vereinigung von Quadern; $\mu(A)$ ist die Summe der Volumina dieser Quader (niedrigdimensionale Quader haben dabei das Volumen 0). Dieses μ heißt das *Lebesguesche Prämaß*, häufig mit λ bezeichnet. □

Beispiel 5.2 Für $\varphi : \mathbb{R} \to \mathbb{R}$ rechtsstetig[1] und nicht fallend sei

$$\mu((a, b)) = \varphi(b-) - \varphi(a), \quad \mu(\{a\}) = \varphi(a) - \varphi(a-).$$

[1]Man kann ohne weiteres auf die Rechtsstetigkeit verzichten, wenn man definiert $\mu((a, b)) := \varphi(b-) - \varphi(a+)$, $\mu(\{a\}) := \varphi(a+) - \varphi(a-)$, allerdings ist dann φ durch μ nicht mehr eindeutig bestimmt.

Jede Vereinigung von endlich vielen (abgeschlossen, offen, halboffen) Intervallen läßt sich als disjunkte Vereinigung von Intervallen der Form (a, b) und $\{a\}$ schreiben; damit läßt sich $\mu(A)$ für jede Teilmenge A von $\mathbb{R}$ definieren, die Vereinigung von endlich vielen Intervallen ist. Dieses μ heißt ein *Lebesgue–Stieltjes* Prämaß. □

Der Nachweis der σ–Additivität erfordert hier, wie in fast allen konkreten Fällen, einige Mühe. Für Beispiel 5.2 wird in Aufgabe 5.1 eine Anleitung gegeben.

Sei μ ein Prämaß auf $(X; \mathcal{R})$. Eine Teilemnge N von X heißt eine *μ–Nullmenge*, wenn zu jedem $\varepsilon > 0$ eine Folge (A_n) aus $\mathcal{R}$ existiert mit

$$N \subset \bigcup_n A_n \quad \text{und} \quad \sum_n \mu(A_n) < \varepsilon.$$

Jede Teilmenge einer μ–Nullmenge ist offenbar eine μ–Nullmenge.

Satz 5.3 *Die Vereinigung von abzählbar vielen μ–Nullmengen N_k ist eine μ–Nullmenge.*

Beweis. Für jedes $k \in \mathbb{N}$ gibt es Folgen $(A_n^k)_n$ aus $\mathcal{R}$

mit $$N_k \subset \bigcup_n A_n^k \qquad \text{und} \qquad \sum_n \mu(A_n^k) < \varepsilon \cdot 2^{-k}.$$

also $$\bigcup_k N_k \subset \bigcup_{k,n} A_n^k \qquad \text{und} \qquad \sum_{k,n} \mu(A_n^k) < \varepsilon.$$

■

Zum Beispiel ist in $\mathbb{R}^m$ jeder Punkt, und damit jede abzählbare Menge eine μ–Nullmenge bezüglich dem Lebesgueschen Prämaß.

Man sagt, eine Eigenschaft in X gilt *μ–fast überall* (μ–f. ü.), wenn es eine μ–Nullmenge N in X gibt so, daß die Eigenschaft für alle $x \in X \setminus N$ gilt, z. B.

$$f(x) = g(x), f(x) > g(x), f_n(x) \to f(x) \quad \text{für} \quad n \to \infty, \ldots \quad \mu\text{–f. ü.}$$

5.2 Integrale für Elementarfunktionen

Eine Funktion $f : X \to \mathbb{C}$ heißt eine $\mathcal{R}$*–Elementarfunktion*, oder Elementarfunktion (bezügl. $\mathcal{R}$), wenn endlich viele $A_1, \ldots, A_n \in \mathcal{R}$ und $c_1, \ldots, c_n \in \mathbb{C}$ existieren mit $A_i \cap A_j = \emptyset$ für $i \neq j$ und

$$f(x) = \begin{cases} c_j & \text{für } x \in A_j, \\ 0 & \text{für } x \in X \setminus \bigcup\limits_j A_j. \end{cases}$$

Diese Funktionen bilden einen *Vektorraum* $E(X, \mathcal{R})$ über $\mathbb{C}$. Im obigen Beispiel ($X = \mathbb{R}^m$, $\mathcal{R} =$ Familie der Figuren in $\mathbb{R}^m$) sind dies die *Treppenfunktionen*.

Für $\mathcal{R}$–Elementarfunktionen definieren wir das μ*–Integral* durch

$$\int f(x)\,\mathrm{d}\mu(x) = \int f\,\mathrm{d}\mu := \sum_j c_j \mu(A_j).$$

Obwohl die Mengen A_j durch f nicht eindeutig bestimmt sind, ist dieses Integral eindeutig definiert. Offensichtlich ist das Integral

linear, d. h. $\displaystyle\int (af + bg) d\mu = a \int f\,\mathrm{d}\mu + b \int g\,\mathrm{d}\mu$,

positiv, d. h. aus $f \geq 0$ folgt $\displaystyle\int f\,\mathrm{d}\mu \geq 0$.

Es gilt

$$\Big| \int f\,\mathrm{d}\mu \Big| = \Big| \sum c_j \mu(A_j) \Big| \leq \sum |c_j| \mu(A_j) = \int |f|\,\mathrm{d}\mu,$$

und somit

$$\Big| \int (f+g)\,\mathrm{d}\mu \Big| \leq \int |f+g|\,\mathrm{d}\mu \leq \int (|f| + |g|)\,\mathrm{d}\mu = \int |f|\,\mathrm{d}\mu + \int |g|\,\mathrm{d}\mu.$$

Also ist

$$\|f\|_1 := \int |f(x)|\,\mathrm{d}\mu(x)$$

eine *Halbnorm* auf $E(X, \mathcal{R})$.

Im folgenden sagen wir: Eine Folge (f_n) von Funktionen $X \to \mathbb{C}$ konvergiert μ*–fast gleichmäßig* gegen eine Funktion $f : X \to \mathbb{C}$, wenn es zu jedem $\varepsilon > 0$ eine Folge (A_k^ε) aus $\mathcal{R}$ gibt mit

$$\sum_k \mu(A_k^\varepsilon) < \varepsilon \quad \text{und} \quad f_n(x) \to f(x) \quad \text{gleichmäßig in} \quad X \setminus \bigcup_k A_k^\varepsilon.$$

Z. B. ist $f_n(x) = x^n$ auf $[0,1]$ λ–fast gleichmäßig gegen 0 konvergent für das Lebesguesche Prämaß λ.

Satz 5.4 *Sei (f_n) eine $\|\cdot\|_1$–Cauchyfolge in $E(X,\mathcal{R})$. dann gilt*

(a) Es gibt eine Funktion $f : X \to \mathbb{C}$ und eine Teilfolge (f_{n_k}) von (f_n) so, daß gilt

$$f_{n_k}(x) \to f(x) \qquad \mu\text{–fast gleichmäßig (insbesondere}\mu\text{–f. ü.).}$$

(b) Sind (g_k) und (h_k) zwei μ–f. ü. konvergente Teilfolgen von (f_n), so gilt

$$g_k(x) - h_k(x) \to 0 \qquad \mu\text{–f. ü.}$$

Beweis. Sei (f_{n_k}) so gewählt, daß gilt

$$\|f_{n_k} - f_{n_{k+1}}\| \leq 3^{-k-1} \qquad \text{für alle} \quad k \in \mathbb{N}.$$

Mit $g_k(x) := |f_{n_k}(x) - f_{n_{k+1}}(x)|$ sei:

$$M_k := \left\{x \in X : g_k(x) > 2^{-k-1}\right\} \in \mathcal{R}, \quad N_k := \bigcup_{\ell \geq k} M_\ell.$$

Dann gilt

$$\mu(M_k) \leq 2^{k+1}3^{-k-1} = \left(\frac{2}{3}\right)^{k+1}, \quad \sum_{\ell \geq k} \mu(M_\ell) \leq \sum_{\ell \geq k} \left(\frac{2}{3}\right)^{\ell+1} \to 0 \quad \text{für} \quad k \to \infty,$$

d. h. $\sum g_\ell(x)$ ist für jedes $k \in \mathbb{N}$ gleichmäßig konvergent in $\complement N_k$; dort ist also $(f_{n_\ell}(x))$ gleichmäßig konvergent, d. h. (f_{n_ℓ}) ist μ–fast gleichmäßig konvergent.

b) Es gelte $g_k(x) \to g(x)$, $k_k(x) \to h(x)$ μ–f. ü. Es ist zu zeigen: $g(x) = h(x)\mu$–f. ü. Offenbar ist $(p_n) = (g_1, h_1, g_2, h_2, \ldots)$ eine $\|\cdot\|_1$–Cauchyfolge in $E(X,\mathcal{R})$. Eine Teilfolge (p_{n_k}) gemäß Teil a kann hiervon so gewählt werden, daß die Folgenglieder mit ungeradem Index aus der Folge (g_k), die mit geradem Index aus der Folge (h_k) entnommen sind.

Es gibt eine μ–Nullmenge N und eine Funktion $f : X \to \mathbb{C}$ mit $p_{n_k}(x) \to f(x)$ für $x \in X \setminus N$. Dies gilt natürlich dann auch für die Teilfolgen mit ungeraden bzw. geraden k; andererseits gibt es Nullmengen N_1 und N_2 so, daß diese in $X \setminus N_1$ bzw. $X \setminus N_2$ gegen g bzw. h konvergieren. In $X \setminus (N_1 \cup N_2 \cup N)$ gilt also $g(x) = f(x) = h(x)$; da $N_1 \cup N_2 \cup N$ eine Nullmenge ist, folgt hiermit die Behauptung. ∎

Satz 5.5 *(a) Sind* A, A_j $(j \in \mathbb{N})$ *aus* $\mathcal{R}$ *und* $A \subset \cup A_j$*, so gilt* $\mu(A) \leq \sum \mu(A_j)$.
(b) Ist (f_n) *aus* $E(X, \mathcal{R})$ *mit* $f_n(x) \searrow 0$ μ*–f. ü., so gilt* $\|f_n\|_1 \to 0$.

Beweis. a) Mit $A_1' := A_1$, $A_{n+1}' := A_{n+1} \setminus \Big(\bigcup\limits_{j=1}^{n} A_j\Big)$ gilt $A = \cup(A \cap A_j')$ (disjunkt), also wegen der σ–Additivität von μ

$$\mu(A) = \sum_j \mu(A \cap A_j') \leq \sum_j \mu(A_j') \leq \sum_j \mu(A_j).$$

b) Wegen $0 \leq \int f_{n+1}\, \mathrm{d}\mu \leq \int f_n\, \mathrm{d}\mu \leq \int f_1\, \mathrm{d}\mu$ ist (f_n) eine $\|\cdot\|_1$–Cauchyfolge. Nach Satz 5.4 a konvergiert also $(f_n)\mu$–fast gleichmäßig gegen 0. Für jedes $\varepsilon > 0$ existiert eine Folge (A_j) aus $\mathcal{R}$ mit $\sum \mu(A_j) < \varepsilon$ und $f_n \to 0$ gleichmäßig in $X \setminus (\cup A_j)$. Wählt man n_0 so, daß

$$f_n(x) < \varepsilon \quad \text{für} \quad n \geq n_0 \quad \text{und} \quad x \in X \setminus (\cup A_j),$$

so folgt für $n \geq n_0$

$$\begin{aligned} \|f_n\|_1 &= \int f_n\, \mathrm{d}\mu \leq \mu\Big(\Big\{x \in X : f_n(x) > \varepsilon\Big\}\Big) \max f_1 + \varepsilon\mu\Big(\Big\{x \in X : f_1(x) \neq 0\Big\}\Big) \\ &\leq \varepsilon\Big(\max f_1 + \mu(\{x \in X : f_1(x) \neq 0\})\Big), \end{aligned}$$

also $\|f_n\|_1 \to 0$. ∎

Satz 5.6 *Sei* (f_n) *eine* $\|\cdot\|_1$*–Cauchyfolge in* $E(X, \mathcal{R})$*, die* μ*–fast überall konvergiert. Dann gilt:* $f_n(x) \to 0$ μ*–f. ü.* $\Longleftrightarrow$ $\|f_n\|_1 \to 0$.

Beweis. $\Leftarrow$: $(f_1, 0, f_2, 0, \ldots)$ ist ebenfalls eine $\|\cdot\|_1$–Cauchyfolge. (f_n) und (0) sind f. ü. konvergente Teilfolgen, die gegen $f(x) := \lim f_n(x)$ bzw. 0 konvergieren. Nach Satz 5.4 b gilt also $f(x) = 0$ μ–f. ü.

$\Rightarrow$: Annahme: $\|f_n\|_1 \not\to 0$ (z. z. es gilt **nicht**: $f_n(x) \to 0$ μ–f. ü.).
O, E. $f_n \geq 0$ (sonst betrachte man $|f_n|$ statt f_n) und $\|f_n\|_1 \to c > 0$. Sei (n_j) eine Teilfolge von $\mathbb{N}$ mit

$$\|f_{n_1}\|_1 \geq \frac{2}{3}c, \|f_{n_j} - f_{n_{j+1}}\|_1 < \frac{2}{3}c\Big(\frac{1}{2}\Big)^j = \frac{1}{3}c\Big(\frac{1}{2}\Big)^{j+1}.$$

Die Folge $(\tilde{f}_j)$ mit

$$\tilde{f}_j(x) := \min(f_{n_1}(x), \dots, f_{n_j}(x))$$

ist nicht wachsend mit

$$\begin{aligned}
\tilde{f}_j &\geq f_{n_1} - |f_{n_1} - f_{n_2}| - \dots - |f_{n_{j-1}} - f_{n_j}|, \\
\|\tilde{f}_j\|_1 &= \int \tilde{f}_j \,\mathrm{d}\mu \geq \|f_{n_1}\|_1 - \|f_{n_1} - f_{n_2}\|_1 - \dots - \|f_{n_{j-1}} - f_{n_j}\|_1 \\
&\geq \frac{2}{3}c - \frac{1}{3}c = \frac{1}{3}c.
\end{aligned}$$

Also gilt $\|\tilde{f}_j\|_1 \not\to 0$, und nach Satz 5.5 b gilt somit *nicht* $\tilde{f}_j(x) \to 0 \mu$–f. ü. Wegen $f_{n_j} \geq \tilde{f}_j$ folgt hieraus: es gilt *nicht* $f_n(x) \to 0 \quad \mu$–f. ü. ■

Korollar 5.7 *(a) Sind (f_n), (g_n) $\|\cdot\|_1$–Cauchyfolgen in $E(X, \mathcal{R})$ mit $f_n(x) - g(x) \to 0 \quad \mu$–f. ü., so folgt*

$$\left| \int (f_n - g_n) \,\mathrm{d}\mu \right| \leq \|f_n - g_n\|_1 \to 0, \quad \|f_n\|_1 - \|g_n\|_1 \to 0.$$

(b) Ist (f_n) eine $\|\cdot\|_1$–Cauchyfolge in $E(X, \mathcal{R})$ mit $\lim_n f_n(x) \geq 0 \quad \mu$–f. ü., so folgt

$$\lim \int f_n \,\mathrm{d}\mu \geq 0 \qquad \textit{entsprechend für „} \leq \textit{“).}$$

Beweis. Da mit (f_n) und (g_n) auch $(f_n - g_n)$ eine $\|\cdot\|_1$–Cauchyfolge ist, folgt aus Satz 5.6 $\|f_n - g_n\|_1 \to 0$. Mit $\left| \|f_n\|_1 - \|g_n\|_1 \right| \leq \|f_n - g_n\|_1$ folgt auch die zweite Behauptung.

b) $g_n(x) := \max\{0, f_n(x)\}$ ist eine $\|\cdot\|_1$–Cauchyfolge mit $f_n(x) - g_n(x) \to 0 \quad \mu$–f. ü. Mit Teil a folgt hieraus

$$\lim_n \int f_n \,\mathrm{d}\mu = \lim_n \int g_n \,\mathrm{d}\mu \geq 0 \quad (\text{da } g_n \geq 0).$$

■

5.3 Integrierbare Funktionen

Eine Funktion $f : X \to \mathbb{C}$ heißt *μ–integrierbar*, wenn eine $\|\cdot\|_1$–Cauchyfolge in $E(X, \mathcal{R})$ existiert mit $f_n(x) \to f(x)\mu$–f. ü. Man definiert dann das *μ–Integral*:

$$\int f(x)\,\mathrm{d}\mu(x) = \int f\,\mathrm{d}\mu := \lim_n \int f_n\,\mathrm{d}\mu.$$

Aus Korollar 5.7 folgt die *Eindeutigkeit* dieser Definition und die *Positivität* des Integrals. Die *Linearität* überträgt sich beim Grenzübergang. Mit f ist auch $|f|\mu$–integrierbar und mit der Positivität folgt

$$\int |f+g|\,\mathrm{d}\mu \le \int (|f|+|g|)\,\mathrm{d}\mu = \int |f|\,\mathrm{d}\mu + \int |g|\,\mathrm{d}\mu.$$

Bemerkung 5.8 Ist $f : I \to R$ Riemann–integrierbar, so ist es auch Lebegue–integrierbar mit gleichem Integral. Sind f *und* $|f|$ über $\mathbb{R}^m$ uneigentlich Riemann–integrierbar, so sind sie Lebesgue–integrierbar mit gleichen Integral (Beweis ?!).

Die Menge der μ–integrierbaren Funktionen auf X bezeichnen wir mit $\mathcal{L}_1(X, \mu)$. $\mathcal{L}_1(X, \mu)$ ist ein komplexer (ggf. reeller) *Vektorraum*. $\|f\|_1 := \int |f|\,\mathrm{d}\mu$ ist eine *Halbnorm* auf $\mathcal{L}_1(X, \mu)$, aber i. allg. keine Norm. Der folgende Satz macht deutlich, daß abgesehen von der Tatsache, daß $\|\cdot\|_1$ keine Norm ist, $\mathcal{L}(X, \mu)$ die „Vervollständigung" von $E(X, \mathcal{R})$ bezüglich $\|\cdot\|_1$ ist.

Satz 5.9 *Ist (f_n) eine $\|\cdot\|_1$–Cauchyfolge in $\mathcal{L}_1(X, \mu)$, so gilt: Es gibt ein $f \in \mathcal{L}_1(X, \mu)$ mit $\int f\,\mathrm{d}\mu = \lim_n \int f_n\,\mathrm{d}\mu$ und eine Teilfolge (f_{n_k}) mit $f_{n_k}(x) \to f(x)\mu$–fast gleichmäßig.*

Beweis. Aus der Definition der μ–Integrierbarkeit und mit Satz 5.4 folgt: Für jedes $n \in \mathbb{N}$ existiert ein $h_n \in E(X, \mathcal{R})$ und eine Folge $(A_\ell^n)_\ell$ aus $\mathcal{R}$ mit

$$\|f_n - h_n\|_1 \le 2^{-n}, \ \sum_\ell \mu(A_\ell^n) < 2^{-n} \text{ und } |f_n(x) - h_n(x)| \le 2^{-n} \text{ in } X \setminus \bigcup_\ell A_\ell^n.$$

Offenbar ist (h_n) eine $\|\cdot\|_1$–Cauchyfolge in $E(X, \mathcal{R})$. Hieraus ergibt sich ebenfalls mit Satz 5.4: Es existiert ein $f : X \to \mathbb{C}$, und eine Teilfolge (h_{n_k}) so, daß für jedes $\varepsilon > 0$ eine Folge (A_q) aus $\mathcal{R}$ existiert mit

$$\sum_q \mu(A_q) < \varepsilon \text{ und } h_{n_k}(x) \to f(x) \text{ gleichmäßig in } X \setminus \bigcup_q A_q.$$

Also gilt $f_{n_k}(x) \to f(x)$ gleichmäßig in $X \setminus \Big\{(\bigcup\limits_q A_q) \bigcup (\bigcup\limits_{n \geq N} \bigcup\limits_{\ell} A_\ell^n)\Big\}$. Da dies für alle $\varepsilon > 0$ und $n \in \mathbb{N}$ gilt, folgt

$$f_{n_k}(x) \to f(x) \mu\text{–fast gleichmäßig (insbes.} \mu\text{–f. ü.).}$$

und somit

$$\lim_n \int f_n \, \mathrm{d}\mu = \lim_k \int f_{n_k} \, \mathrm{d}\mu = \lim_k \int h_{n_k} \, \mathrm{d}\mu = \int f \, \mathrm{d}\mu.$$

■

5.4 Grenzwertsätze

Satz 5.10 (Beppo Levi) *Ist (f_n) aus $\mathcal{L}_1(X, \mu)$ mit $f_n(x) \leq f_{n+1}(x)$ und $\int f_n d\mu \leq C$ für alle n (*monotone Folge mit beschränkter Integralfolge*), so existiert ein $f \in \mathcal{L}_1(X, \mu)$ mit*

$$f_n(x) \to f(x) \mu\text{–f. ü. und } \int f_n \, \mathrm{d}\mu \to \int f \, \mathrm{d}\mu.$$

Beweis. Aus $\int f_n \, \mathrm{d}\mu \leq \int f_{n+1} \, \mathrm{d}\mu \leq C$ und $\int |f_n - f_m| \, \mathrm{d}\mu = \int f_n \, \mathrm{d}\mu - \int f_m \, \mathrm{d}\mu$ für $n \geq m$ folgt, daß (f_n) eine $\| \cdot \|_1$–Cauchyfolge ist. Mit Satz 5.8 folgt hieraus die Behauptung. ■

Satz 5.11 *(a) Sind $f, g \in \mathcal{L}_1(X, \mu)$ reellwertig, so folgt* $\min(f, g)$, $\max(f, g)$ f_+, $f_- \in \mathcal{L}(X, \mu)$.

(b) Mit $f \in \mathcal{L}(X, \mu)$ gilt auch $\mathrm{Re}\ f$, $\mathrm{Im}\ f \in \mathcal{L}_1(X, \mu)$.

Beweis. a) Zu f, g seien (f_n), (g_n) die $\| \cdot \|_1$–Cauchyfolgen gemäß Definition der μ–Integrierbarkeit. Dann ist auch $(\min(f_n, g_n))$ eine $\| \cdot \|_1$–Cauchyfolge mit $\min(f_n, g_n)(x) \to \min(f, g)(x) \mu$–f. ü.; also ist $\min(f, g) \in \mathcal{L}_1(X, \mu)$. Entsprechend für $f_+ = \max(f, 0)$ und $f_- = (-f)_+$.
b) Ist (f_n) eine $\| \cdot \|_1$–Cauchyfolge für f, so sind $(\mathrm{Re}\ f_n))$, $(\mathrm{Im}\ f_n)$ $\| \cdot \|_1$–Cauchyfolgen für $\mathrm{Re}\ f$ bzw. $\mathrm{Im}\ f$. ■

Satz 5.12 (Lebesgue, Satz von der dominierten Konvergenz) *Gilt* $f_n \in \mathcal{L}_1(X,\mu)$ *für* $n \in \mathbb{N}$, $|f_n(x)| \le g(x)$*–f. ü. mit einem* $g \in \mathcal{L}_1(X,\mu)$, $f_n(x) \to f(x)$*–f. ü. (f. ü.* konvergente Folge mit integrierbarer Majorante*), so folgt*

$$f \in \mathcal{L}_1(X,\mu), \qquad \int f \,\mathrm{d}\mu = \lim_n \int f_n \,\mathrm{d}\mu.$$

Beweis. Es gilt

$$g_n := \sup\{f_n, f_{n+1}, \ldots\} = \lim_k \max\{f_n, \ldots, f_{n+k}\}.$$

Mit Satz 5.10 a) folgt $\max\{f_n, \ldots, f_{n+k}\} \in \mathcal{L}_1(X,\mu)$ und Satz 5.9 (B. Levi) liefert $g_n \in \mathcal{L}_1(x,\mu)$. Auf die nichtwachsende Folge $(g_n)(\int g_n \,\mathrm{d}\,\mu \ge -\int g \,\mathrm{d}\,\mu)$ kann ebenfalls Satz 5.9 (B. Levi) angewandt werden:

$$g_n(x) \searrow f(x) \quad \mu\text{–f. ü.}, \quad f \in \mathcal{L}_1(X,\mu), \quad \int g_n \,\mathrm{d}\mu \to f \,\mathrm{d}\mu.$$

Mit $h_n := \inf\{f_n, \ldots\}$ folgt entsprechend $\int h_n \,\mathrm{d}\mu \to \int f \,\mathrm{d}\mu$. Wegen $h_n \le f_n \le g_n$ ergibt sich $\int f_n \,\mathrm{d}\mu \to \int f \,\mathrm{d}\mu$. ■

Satz 5.13 (Lemma von Fatou) *Gilt* $f_n \in \mathcal{L}_1(X,\mu)$ *für* $n \in \mathbb{N}$, $f_n \ge 0$, $\int f_n \,\mathrm{d}\,\mu \le C$ *und* $f_n \to f\mu$*–f. ü. (f. ü.* konvergente positive Folge mit beschränkter Integralfolge*), so folgt*

$$f \in \mathcal{L}_1(X,\mu), \qquad \int f \,\mathrm{d}\mu \le \liminf_n \int f_n \,\mathrm{d}\mu.$$

Beweis. Mit $h_n := \inf\{f_n, f_{n+1}, \ldots\}$ gilt $h_n(x) \nearrow f(x)\mu$–f. ü. Aus $h_n \le f_{n+k}$ $(n, k \in \mathbb{N})$ folgt

$$\int h_n \,\mathrm{d}\mu \le \liminf_k \int f_{n+k} \,\mathrm{d}\mu = \liminf_k \int f_k \,\mathrm{d}\mu \le C.$$

Mit Satz 5.9 (B. Levi) ergibt sich also

$$f \in \mathcal{L}_1(X,\mu) \quad \text{und} \quad \int f \,\mathrm{d}\mu = \lim_n \int h_n \,\mathrm{d}\mu \le \liminf_k \int f_k \,\mathrm{d}\mu.$$

■

5.5 Meßbare Mengen und Funktionen, Maße

Zur Vereinfachung nehmen wir im folgenden an, daß (X, μ) σ*–endlich* ist, d. h. es gibt eine Folge (X_n) aus $\mathcal{R}$ mit $X = \cup X_n$.

Eine Funktion $f : X \to \mathbb{C}$ heißt μ*–meßbar*, wenn eine Folge (f_n) aus $E(X, \mathcal{R})$ existiert mit $f_n(x) \to f(x)$ μ–f. ü. (Im nicht σ–endlichen Fall müßte man definieren: f heißt μ–meßbar, wenn in jedem $A \in \mathcal{R}$ die Funktion f μ–f. ü. Limes von Elementarfunktionen ist. Wir überlassen es dem Leser, sich die im folgenden nötigen Modifikationen zu überlegen. In den Anwendungen spielen nicht σ–endliche Maße kaum eine Rolle.)

Satz 5.14 *Ist f μ–meßbar, $g \in \mathcal{L}_1(X, \mu)$ und $|f(x)| \le g(x)$ μ–f. ü., so ist $f \in \mathcal{L}_1(X, \mu)$.*

Beweis. Sei $f_n \in E(X, \mathcal{R})$ mit $f_n(x) \to f(x) \mu$–f. ü., dann gilt

$$\tilde{f}_n := \max\Big\{ -g, \min\{f_n, g\}\Big\} \in \mathcal{L}_1(X, \mu) \qquad \text{(Satz 5.10)}$$

und

$$\tilde{f}_n(x) \to f(x) \quad \mu\text{–f. ü.}, \qquad |\tilde{f}_n| \le g.$$

Daraus ergibt sich die Behauptung mit Satz 5.11 (Lebesgue). ■

Satz 5.15 *(a) Sind $f, g\mu$–meßbar, so sind auch $f+g, f\cdot g, \dfrac{f}{g}, |f|, \ldots$ μ–meßbar.*

(b) Ist $f_n\mu$–meßbar $(n \in \mathbb{N})$ und $f_n(x) \to f(x)\mu$–f. ü., so ist auch $f\mu$–meßbar.

Beweis. a) Offensichtlich.

b) Es genügt offenbar, reellwertige Funktionen zu betrachten.Seien X_n disjunkte Mengen aus $\mathcal{R}$ mit $X = \cup X_n$. Für $h(x) = (n^2\mu(X_n))^{-1}$ für $x \in X_n$ gilt $h \in \mathcal{L}_1(X, \mu)$ und $h(x) > 0$ für alle $x \in X$. Dann ist

$$g_n(x) := \frac{h(x) f_n(x)}{h(x) + |f_n(x)|}, \quad \mu\text{–meßbar}, \quad |g_n(x)| \le h(x),$$

und somit $g_n \in \mathcal{L}_1(X, \mu)$ (Satz 5.14). Es gilt

$$g_n(x) \to g(x) := \frac{h(x)f(x)}{h(x) + |f(x)|}, \qquad |g(x)| \leq h(x)$$

also $g \in \mathcal{L}_1(X, \mu)$ und somit $f = \dfrac{hg}{h - |g|}$ μ–meßbar. ■

Eine Menge $A \subset X$ heißt μ*–meßbar*, wenn die charakteristische Funktion χ_A μ–meßbar ist. Mit A ist offenbar auch $\complement A$ μ–meßbar. Das Prämaß μ wird (ohne, daß wir dafür eine neue Bezeichnung einführen) auf die μ–meßbaren Mengen ausgedehnt durch:

$$\mu(A) := \begin{cases} \int \chi_A d\mu & \text{falls } \chi_A \in \mathcal{L}_1(X, \mu), \\ \infty & \text{sonst.} \end{cases}$$

Satz 5.16 *(a) Sind A_n μ–meßbar mit $A_n \cap A_m = \emptyset$, so ist $\bigcup\limits_n A_n$ μ–meßbar und es gilt $\mu(\bigcup\limits_n A_n) = \sum\limits_n \mu(A_n)$.*

(b) Sind A_n μ–meßbar mit $A_n \supset A_{n+1}$ und ist $\mu(A_n) < \infty$ für ein n, so ist $\cap A_n$ μ–meßbar und es gilt $\mu(\bigcap\limits_n A_n) = \lim\limits_{n\to\infty} \mu(A_n)$.

Beweis. Zunächst gilt

(a) $\chi_{\bigcup_{n=1}^m A_n} \leq \chi_{\bigcup A_n}$, $\chi_{\bigcup_{n=1}^m A_n} \nearrow \chi_{\bigcup A_n} = \sum\limits_n \chi_{A_n}$.

(b) $\chi_{\bigcap_{n=1}^m A_n} \leq \chi_{\bigcap A_n}, \chi_{\bigcap_{n=1}^m A_n} \searrow \chi_{\bigcap A_n} = \lim\limits_n \chi_{A_n}$.

Hieraus folgen die Meßbarkeitsaussagen sofort. Die Aussagen über die Maße folgen aus den Grenzwertsätzen. ■

Eine Familie $\mathcal{A}$ von Teilmengen von X heißt eine σ*–Algebra* in X, wenn gilt:

$\emptyset \in \mathcal{A}$, aus $A \in \mathcal{A}$ folgt $\complement A \in \mathcal{A}$, und
aus $A_n \in \mathcal{A}$ $(n \in \mathbb{N})$ folgt $\bigcup A_n \in \mathcal{A}$.

Eine Abbildung $\mu : \mathcal{A} \to [0, \infty]$ mit der Eigenschaft:

aus $A_n \in \mathcal{A}(n \in \mathbb{N})$, $A_n \bigcap A_m = \emptyset$ für $n \neq m$
folgt $\mu(\bigcup A_n) = \sum \mu(A_n)$ $(\sigma$*–additiv*$)$

heißt ein *Maß* auf $(X, \mathcal{A})$. $(X, \mathcal{A}, \mu)$ oder kürzer (X, μ) wird als *Maßraum* bezeichnet.

Satz 5.15 besagt, daß $(X, \mathcal{A}_\mu, \mu) = (X, \mu)$ ein Maßraum ist, wenn $\mathcal{A}_\mu$ die Familie der μ–meßbaren Mengen ist.

Satz 5.17 *Eine Teilmenge N von X ist genau dann eine μ–Nullmenge, wenn sie μ–meßbar ist ist $\mu(N) = 0$.*

Beweis. $\Rightarrow$: Ist N eine μ–Nullmenge, so ist χ_N μ–f. ü. gleich der Nullfunktion, also $\chi_N \in \mathcal{L}_1(X, \mu)$ und somit N μ–meßbar mit $\mu(N) = \int \chi_N \, d\mu = 0$.
$\Leftarrow$: Sei N μ–meßbar mit $\mu(N) = 0$. Dann ist $(n\chi_N)$ eine nicht–fallende Folge aus $\mathcal{L}_1(X, \mu)$ mit $\int n\chi_N \, d\mu = n\mu(N) = 0$ (also beschränkter Integralfolge). Also ist $n\chi_N$ μ–f. ü. konvergent. Da die Folge aber nur außerhalb von N konvergiert, ist N eine μ–Nullmenge. ∎

Ein Maß heißt *vollständig*, wenn jede Teilmenge einer Nullmenge meßbar ist. Die hier konstruierten Maße sind also stets vollständig.

Ist $A \subset X$ μ–meßbar, so definiert man

$$\int_A f \, d\mu := \int \chi_A f \, d\mu \quad \text{falls} \quad \chi_A f \in \mathcal{L}_1(X, \mu) \quad \text{ist.}$$

Für $A \cap B = \emptyset$ gilt dann offenbar.

$$\int_A f \, d\mu + \int_B f \, d\mu = \int_{A \cup B} f \, d\mu.$$

Satz 5.18 *Sei $f : X \to \mathbb{R}$. f ist genau dann μ–meßbar, wenn $M_c := \{x \in X : f(x) \geq c\}$ für jedes $c \in \mathbb{R}$ μ–meßbar ist (entsprechend für $> c, \leq c, < c$).*

Beweis. $\Leftarrow$: Sei M_c μ–meßbar für jedes $c \in \mathbb{R}$. Dann sind auch die Mengen

$$\{x \in X : f(x) < t\} = X \setminus M_t, \quad M(s, t) := \{x \in X : s \leq f(x) < t\}$$

μ–meßbar. Damit sind die Funktionen

$$f_n(x) := \begin{cases} \frac{k}{n} & \text{für } x \in M((k-1)n^{-1}, kn^{-1}) \text{ für } k = -n^2 + 1, \ldots, n^2, \\ 0 & \text{sonst} \end{cases}$$

μ–meßbar und es gilt $f_n(x) \to f(x)$ für alle $x \in X$.

$\Rightarrow$: Sei o. E. $c = 1$ (sonst ist $f - c + 1$ statt f zu betrachten). $g := \min\{1, \max\{0, f\}\}$ ist μ–meßbar und es gilt $g^n(x) \to \chi_{M_c}(x)$ für $n \to \infty$ und alle $x \in X$; also ist χ_{M_c} μ–meßbar. ■

5.6 Produktmaße und der Satz von Fubini–Tonelli

Seien $\mathcal{R}_X, \mathcal{R}_Y$ Mengenringe in X bzw. Y, μ_X, μ_Y Prämaße auf X bzw. Y. Die Mengen $A \times B$ mit $A \in \mathcal{R}_X$, $B \in \mathcal{R}_Y$ erzeugen einen Mengenring $\mathcal{R}_X \times \mathcal{R}_Y$ in $X \times Y$. Durch

$$\mu(\cup(A_j \times B_j)) := \sum \mu_X(A_j)\mu_Y(B_j) \text{ für } A_j \in \mathcal{R}_X, B_j \in \mathcal{R}_Y,\ A_j \times B_j \cap A_k \times B_k = \emptyset$$

wird ein Maß μ in $X \times Y$ erzeugt (*Produktmaß*).

Hilfssatz 5.19 *Ist $N \subset X \times Y$ eine μ–Nullmenge, so ist*

$$N_X = \Big\{x \in X : \{y \in Y : (x.y) \in N\} \quad \textit{ist nicht } \mu_Y\textit{–Nullmenge} \Big\}$$

eine μ_X–Nullmenge, d. h. für μ_X–fast alle x ist $\{y \in Y : (x, y) \in N\}$ eine μ_Y–Nullmenge.

Beweis. Sei N eine μ–Nullmenge. Dann existieren (Beweis!)

$$C_k = A_k \times B_k (A_k \in \mathcal{R}_X,\ B_k \in \mathcal{R}_Y) \text{ mit}$$

$$N \subset \bigcup_k C_k, \sum_k \mu(C_k) < \infty, \text{ jedes } z \in N \text{ ist } \textit{in } \infty \textit{ vielen } C_k \text{ enthalten.}$$

Aus

$$\sum_k \int \chi_{A_k}\, \mathrm{d}\mu_X \int \chi_{B_k}\, \mathrm{d}\mu_Y = \sum \mu_X(A_k)\mu_Y(B_k) = \sum \mu(C_k) < \infty$$

folgt: Die Folge

$$\left(\sum_{\ell=1}^{n} \chi_{A_\ell}(x) \int \chi_{B_\ell}\, \mathrm{d}\mu_Y \right)_n \quad \text{in} \quad E(X, \mathcal{R}_X)$$

ist nicht fallend mit beschränkter Integralfolge. Also existiert eine μ–Nullmenge F_X mit

$$\sum_\ell \chi_{A_\ell}(x) \int \chi_{B_\ell}\, \mathrm{d}\mu_y < \infty \quad \text{in} \quad X \setminus F_X.$$

Es genügt also $N_X \subset F_X$ zu zeigen: Sei $x_0 \in X \setminus F_X$. Dann gilt

$$\sum_\ell \int\limits_Y \chi_{A_\ell}(x_0)\chi_{A_\ell}(y)\, \mathrm{d}\mu_Y(y) < \infty.$$

Ist $y \in Y$ so, daß $(x_0, y) \in N$ gilt, so ist (x_0, y) in ∞ vielen $C_k = A_k \times B_k$ enthalten, d. h. die Folge

$$\left(\sum_{\ell=1}^n \chi_{A_\ell}(x_0)\chi_{B_\ell}(y)\right)_n \quad \text{ist divergent.}$$

Da die zugehörige Integralfolge (über y) beschränkt ist, ist die Menge der $y \in Y$ mit $(x_0, y) \in N$ eine μ_Y–Nullmenge; also ist $x_0 \notin N_X$, d. h. $N_X \subset F_X$. ■

Satz 5.20 (Fubini-Tonelli) *Es gilt: $f \in \mathcal{L}_1(X \times Y, \mu) \iff f$ ist μ–meßbar, $|f(x,\cdot)|$ ist für μ_X–f. a. $x \in X$ μ_Y–integrierbar, und*

$$F(x) := \begin{cases} \int\limits_Y |f(x,y)|\, \mathrm{d}\mu_Y(y) & \textit{falls } |f(x,\cdot)|\mu_Y\textit{–integrierbar,} \\ 0 & \textit{sonst} \end{cases}$$

ist μ_X–integrierbar. (Die Aussage „$\Longrightarrow$“ wird als Satz von Fubini bezeichnet, die Aussage „$\Longleftarrow$“ als Satz von Tonelli.)

Es ist dann auch $\int_Y f(x,y)\, \mathrm{d}\mu_Y(y)$ μ_X–integrierbar und es gilt

$$\int f\, \mathrm{d}\mu = \int\limits_X \Big\{ \int\limits_Y f(x,y)\, \mathrm{d}\mu_Y(y) \Big\}\, \mathrm{d}\mu_X(x).$$

Entsprechendes gilt, wenn die Rollen von x und y vertauscht werden. Für $f \in \mathcal{L}_1(X \times Y, \mu)$ ist die Integrationsreihenfolge vertauschbar.

Beweis. $\Rightarrow$: Ist $f \in \mathcal{L}_1(X \times Y, \mu)$, so existiert eine $\|\cdot\|_1$–Cauchyfolge (f_n) aus $E(X \times Y, \mathcal{R}_X \times \mathcal{R}_Y)$, mit $f_n(x,y) \to f(x,y)$ μ–f. ü. Dabei können wir o. E. annehmen

$$f_n(x,y) = \sum_{j=1}^{n} h_j(x,y) \quad \text{mit} \quad \sum_{j=1}^{\infty} \|h_j\|_1 < \infty.$$

Nach Satz 5.17 existiert eine μ_X–Nullmenge N_1 mit

$$(1) \qquad f_n(x,y) \to f(x,y) \quad \text{für} \quad \mu_Y\text{–f. a.,} \quad x \in X \setminus N_1.$$

Außerdem gilt offenbar, da die Behauptung für Treppenfunktionen gilt

$$(2') \qquad \left(\sum_{j=1}^{n} \int_Y |h_j(x,y)| \, \mathrm{d}\mu_Y(y) \right)_n \quad \text{ist} \quad \|\cdot\|_1\text{–Cauchyfolge in} \quad E(X, \mathcal{R}_X)$$

und

$$(2) \qquad \left(\int_Y f_n(x,y) \, \mathrm{d}\mu_Y(y) \right) \quad \text{ist} \quad \|\cdot\|_1\text{–Cauchyfolge aus} \quad E(X, \mathcal{R}_X).$$

Es existiert eine μ_X–Nullmenge N_2 so, daß die Folge aus $(2')$ in $(X \setminus N_2)$ konvergiert, also ist

$$(3) \qquad (f_n(x,\cdot)) \text{ eine} \|\cdot\|_1\text{–Cauchyfolge in } E(Y, \mathcal{R}_Y) \text{ für } x \in X \setminus N_2.$$

Aus (1) und (3) folgt

$$(4) \qquad f(x,\cdot) \in \mathcal{L}_1(Y, \mu_Y), \ \int f_n(x,y) \, \mathrm{d}\mu_Y(y) \to \int f(x,y) \, \mathrm{d}\mu_Y(y) \text{ für } x \in X \setminus (N_1 \cup N_2).$$

Aus (2) und (4) folgt

$$\begin{aligned} \int \left\{ \int f(x,y) \, \mathrm{d}\mu_Y(y) \right\} \mathrm{d}\mu_X(x) &= \lim \int \left\{ \int f_n(x,y) \, \mathrm{d}\mu_Y(y) \right\} \mathrm{d}\mu_X(x) = \\ &= \lim \int f_n \, \mathrm{d}\mu_X \times \mu_Y = \int f \, \mathrm{d}\mu_X \times \mu_Y = \int f \, \mathrm{d}\mu. \end{aligned}$$

Anwendung dieses Resultats auf $|f|$ liefert die Aussage über F.

$\Rightarrow$: Seien $A_n \in \mathcal{R}_X, B_n \in \mathcal{R}_Y$ mit $A_n \subset A_{n+1}, B_n \subset B_{n+1}, X = \cup A_n, Y = \cup B_n$,

$$f_n(x,y) := \begin{cases} f(x,y) & \text{falls } |f(x,y)| \leq n \text{ und } (x,y) \in A_n \times B_n, \\ 0 & \text{sonst.} \end{cases}$$

Dann folgt:

$$
\begin{aligned}
&f_n \in \mathcal{L}_1(X \times Y, \mu_X \times \mu_Y), |f_n|) \nearrow, f_n(x,y) \to f(x,y),\\
&\int |f_n| \,\mathrm{d}\mu_X \times \mu_Y = \int \Big\{ \int |f_n| \,\mathrm{d}\mu_Y \Big\} \,\mathrm{d}\mu_X \le \int \Big\{ \int |f| \,\mathrm{d}\mu_Y \Big\} \,\mathrm{d}\mu_X < \infty \text{ (Fubini)},\\
&|f| \in \mathcal{L}_1(X \times Y, \mu_X \times \mu_Y) \quad \text{B. Levi},\\
&f \in \mathcal{L}_1(X \times Y, \mu_X \times \mu_Y) \quad \text{(da } f\text{–meßbar ist).}
\end{aligned}
$$

■

Für das durch das Lebesguesche Prämaß erzeugte Integral (das Lebesgue–Integral) in $\mathbb{R}$ bzw. $\mathbb{R}^m$ schreiben wir kurz

$$\int f(x)\,\mathrm{d}x \text{ bzw. } \int f(x_1,\dots,x_m)\,\mathrm{d}(x_1,\dots,x_m) = \int \dots \int f(x_1,\dots,x_m)\,\mathrm{d}x_1 \dots \mathrm{d}x_m.$$

5.7 Absolut stetige Funktionen, partielle Integration

Eine Funktion $F : \mathbb{R} \to \mathbb{C}$ heißt *absolutstetig* (bezüglich dem Lebesgue–Maß), wenn es eine über jedes beschränkte Intervall Lebesgue–integrierbare (*lokal integrierbare*) Funktion $f : \mathbb{R} \to \mathbb{C}$ gibt mit

$$F(x) - F(c) = \int_c^x f(t)\,\mathrm{d}t := \begin{cases} \displaystyle\int_{(c,x)} f(t)\,\mathrm{d}t & \text{für } c \le x,\\ \displaystyle -\int_{(x,c)} f(t)\,\mathrm{d}t & \text{für } x < c; \end{cases}$$

f heißt die (*Lebesgue–*) *Ableitung* von F. — Jede stetige differenzierbare Funktion ist absolut stetig; die klassische Ableitung und die Lebesgue–Ableitung stimmen fast überall überein.

Satz 5.21 (Partielle Integration) *Sind $F, G : \mathbb{R} \to \mathbb{C}$ absolut stetig und f, g ihre Lebesgue–Ableitungen, so gilt für $c, x \in \mathbb{R}$*

$$\int_c^x F(t)g(t)\,\mathrm{d}t = F(x)G(x) - F(c)G(c) - \int_c^x f(t)G(t)\,\mathrm{d}t.$$

Beweis. Es gibt Folgen $(f_n), (g_n)$ von Treppenfunktionen so, daß für jedes beschränkte Intervall I gilt

$$\int_I |f_n(t) - f(t)|\,\mathrm{d}t \to 0, \int_I |g_n(t) - g(t)|\,\mathrm{d}t \to 0.$$

Dann konvergieren

$$F_n(t) = F(c) + \int_c^t f_n(s)\,\mathrm{d}s, \quad G_n(t) = G(c) + \int_c^t g_n(s)\,\mathrm{d}s$$

lokal gleichmäßig gegen F bzw. G. Da die gewünschte Gleichung für f_n, g_n, F_n, G_n statt f, g, F, G gilt, folgt die Behauptung durch Grenzübergang. ∎

5.8 Beispiel einer nicht Lebesgue–meßbaren Teilmenge von $\mathbb{R}$

In den Anwendungen spielt natürlich das Lebesguesche Maß λ auf $\mathbb{R}$ und $\mathbb{R}^m$ die größte Rolle. Offenbar sind *sehr viele* Funktionen Lebesgue–integrierbar, also sind auch *sehr viele* Teilmengen von $\mathbb{R}$ bzw. $\mathbb{R}^m$ ebesgue–meßbar. Da es durchaus nicht–trivial ist, nicht Lebesgue–meßbare Funktionen bzw. Mengen zu konstruieren, wollen wir hier noch ein Beispiel vorstellen.

Sei $X = [0,1)$. Zwei Punkte $x, y \in [0,1)$ nennen wir *äquivalent*, $x \sim y$, falls $x - y$ rational ist (dies ist eine Äquivalenzrelation!). Hierdurch wird X in *Äquivalenzklassen* zerlegt (x, y sind genau dann aus zwei verschiedenen Äquivalenzklassen, wenn $x - y$ irrational ist).
Sei $M \subset X$ eine Teilmenge von X, die aus jeder Äquivalenzklasse genau ein Element enthält (Auswahlaxiom). Für jedes $r \in \mathbb{Q}_0 := \mathbb{Q} \cap [0,1)$ sei

$$M_r := \Big\{x + r (\operatorname{mod} 1) : x \in M\Big\} \quad \text{(speziell ist} \quad M_0 = M).$$

Es ist nun leicht zu zeigen, daß M *nicht* Lebesgue–meßbar ist.

Wir nehmen dazu an, daß M Lebesgue–meßbar ist. Dann gilt

$$(\alpha) \quad \lambda(M) = 0 \qquad \text{oder} \qquad (\beta) \quad \lambda(M) > 0.$$

Wir benutzen im folgenden die offensichtliche Tatsache, daß das Lebesgue–Maß λ *translationsinvariant* ist, $\lambda(A) = \lambda(x+A)$ für jede Lebesgue–meßbare Menge und jedes $x \in \mathbb{R}$ (vgl. Aufgabe 5.8). Es gilt also $\lambda(M_r) = \lambda(M)$ für alle $r \in \mathbb{Q}_0$. Wegen $M_r \cap M_s = \emptyset$ für $r \neq s$ und $\bigcup_{r\in\mathbb{Q}_0} M_r = [0,1)$ folgt mit Hilfe der σ–Additivität

$$\begin{aligned} &\text{im Fall } (\alpha) \quad \lambda([0,1)) = \sum_{r\in\mathbb{Q}} \lambda(M_r) = 0, \\ &\text{im Fall } (\beta) \quad \lambda([0,1)) = \sum_{r\in\mathbb{Q}} \lambda(M_r) = \infty, \end{aligned}$$

das ist in beiden Fällen ein Widerspruch zu $\lambda([0,1)) = 1$. Also ist M nicht Lebesgue–meßbar.

Damit ist χ_M ein Beispiel einer nicht Lebesgue–meßbaren Funktion.

5.9 Übungsaufgaben

5.1 Man zeige, daß das Lebesgue–Stieltjes Prämaß (vgl. Beispiel 5.2) σ–additiv ist.
Anleitung: Man benutze z. B. die folgenden Beweisschritte:

(a) Zu jedem Intervall I und jedem $\varepsilon > 0$ gibt es ein offenes Intervall J mit $I \subset J$ und $\mu(J) \leq \mu(I) + \varepsilon$. Ist I ein abgeschlossenes Intervall und (I_n) eine disjunkte Folge von Intervallen mit $I = \cup I_n$, so gilt $\mu(I) = \sum \mu(I_n)$.

(b) Beweis der Behauptung mit Hilfe von Teil b.

5.2 Man zeige: Sind $f : \mathbb{R}^m \to \mathbb{C}$ und $|f|$ im Unendlichen uneigentlich Riemann–integrierbar, so ist f Lebesgue–integrierbar.

5.3 Sei $f : [a,b] \to \mathbb{R}$ Lebesgue–integrierbar. Ist $\displaystyle\int_1 f \,\mathrm{d}x = 0$ für jedes Intervall $I \subset [a,b]$, so gilt $f(x) = 0$fast überall. (Es genügt auch, offene oder abgeschlossene Intervalle zu benutzen).
Anleitung: Man benutze z. B. die folgenden Beweisschritte:

(a) $\displaystyle\int gf \,\mathrm{d}x = 0$ für jede Treppenfunktion g,

(b) $\displaystyle\int_A f \,\mathrm{d}x = 0$ für jede Lebesgue–meßbare Teilmenge A von $[a,b]$,

(c) ist $h : [a,b] \to \mathbb{R}$ Lebesgue–integrierbar, $h \geq 0$ f. ü. und $\displaystyle\int_{[a,b]} h \,\mathrm{d}\mu = 0$, so gilt $h(x) = 0$ f. ü.

5.4 (a) Zu jeder Familie von Teilmengen einer Menge X gibt es eine kleinster σ–Algebra, die diese Familie enthält. Die auf diese Weise aus den offenen Mengen des $\mathbb{R}^m$ erzeugte Familie ist die *Borelsche σ–Algebra*; ihre Mengen heißen *Borel*–Mengen.

(b) Eine Funktion auf $\mathbb{R}^m$ heißt *Borel–meßbar*, wenn das Urbild jeder offenen Menge eine Borel–Menge ist. Ist ϱ ein Maß auf $\mathbb{R}^m$, das im Sinne dieses Paragraphen aus einem Prämaß auf den Figuren in $\mathbb{R}^m$ erzeugt wird (z. B. das Lebesgue–Maß), so ist jede Borel–Menge eine ϱ–meßbare Menge; jede Borel–meßbare Funktion ist ϱ–meßbare Menge; jede Borel–meßbare Funktion ist ϱ–meßbar.

5.5 Sei $f : \mathbb{R}^m \to \mathbb{R}$ beschränkt, $f(x) = 0$ für $|x| \geq R$,

$$\overline{f}(x) := \inf_Q \{ \sup\{f(y) : y \in Q\} : x \in \mathring{Q}\},$$
$$\underline{f}(x) := \sup_Q \{ \inf\{f(y) : y \in Q\} : x \in \mathring{Q}\},$$

wobei Q Quader in $\mathbb{R}^m$ sind und $\mathring{A}$ das Innere der Menge A bezeichnet.

(a) Es gilt

$$\begin{aligned}\overline{f}(x) &= \lim_{n\to\infty} \sup\{f(y) : y \in Q_n\},\\ \underline{f}(x) &= \lim_{n\to\infty} \inf\{f(y) : y \in Q_n\}\end{aligned}$$

für jede Folge (Q_n) von Quadern mit $x \in \mathring{Q}_n$, deren Kantenlängen gegen 0 konvergieren.

(b) Es gilt $\overline{f}(x) = \underline{f}(x) = f(x)$ genau dann, wenn f im Punkt x stetig ist.

(c) $\overline{f}$ und $\underline{f}$ sind Lebesgue–integrierbar und es gilt

$$\overline{\int} f(x)\,\mathrm{d}x = \int \overline{f}(x)\,\mathrm{d}x, \quad \underline{\int} f(x)\,\mathrm{d}x = \int \underline{f}(x)\,\mathrm{d}x = \int \underline{f}(x)\,\mathrm{d}x,$$

wobei $\overline{\int}$ bzw. $\underline{\int}$ das Ober– bzw. Unterintegral, $\displaystyle\int$ das Lebesgue–Integral bedeuten.

(d) f ist genau dann Riemann–integrierbar, wenn f fast überall (bezüglich dem Lebesgue–Maß) stetig ist.

5.6 Die *Cantor–Menge C* ensteht aus dem Intervall $[0,1]$, indem zunächst das offene mittlere Drittel herausgenommen wird, aus den zwei verbleibenden Intervallen wieder jeweils das offene mittlere Drittel, usw., also

$$C = [0,1] \setminus \left\{ \left(\frac{1}{3}, \frac{2}{3}\right) \cup \left(\frac{1}{9}, \frac{2}{9}\right) \cup \left(\frac{7}{8}, \frac{8}{9}\right) \cup \dots \right\}.$$

(a) C ist eine Jordan–Nullmenge, also auch eine Lebesgue–Nullmenge.

(b)) C besteht aus den Punkten $a \in \mathbb{R}$ mit $a = \sum\limits_{j=1}^{\infty} a_j 3^{-j}$ mit $a_j \in \{0,2\}$.

(c) C hat die Mächtigkeit von $[0,1]$ (ist also insbesondere nicht abzählbar).

5.7 Sei $0 < a_n < 1$ mit $\prod_{n=1}^{\infty} a_n > 0$, $b_n = 1 - a_n$. Eine *verallgemeinerte Cantor–Menge* $C_{(a_n)}$ wird dadurch definiert, daß im n–ten Schritt nicht ein Drittel, sondern der b_n–te Anteil des jeweiligen Intervalls herausgenommen wird. Die Menge $C_{(a_n)}$ ist Lebesgue–meßbar mit Maß $\prod\limits_{n=1}^{\infty} a_n$, aber nicht Jordan–meßbar. Was bedeutet das für die entsprechenden Integrale?

5.8 Man beweise die Substitutionsregel für das Lebesgue–Integral in $\mathbb{R}^m$ mit Hilfe der Substitutionsregel für das Riemann–Integral.

5.9 In Aufgabe 2b b kann $\varphi : \mathbb{R}^m \to \mathbb{C}$ wie folgt verallgemeinert werden:

(a) φ Lebesgue–integrierbar über jede kompakte Teilmenge von $\mathbb{R}^m$ (lokal Lebesgue–integrierbar),

(b) $(1 + |\cdot|)^p \varphi$ Lebesgue–integrierbar über $\mathbb{R}^m$ für ein $p \in \mathbb{N}$,

(c) φ Lebesgue–integrierbar und $\varphi(x) = 0$ für λ–fast alle $x \in \mathbb{R}^m$ mit $|x| \geq R_\varphi$ (λ = Lebesgue–Maß).

5.10 Sei $\mathcal{R}$ der Mengenring in $\mathbb{R}$, der aus endlichen Vereinigungen von Intervallen besteht, $\varrho : \mathcal{R} \to [0,\infty)$ sei definiert durch

$$\varrho(A) := \begin{cases} 1 & \text{falls ein } \varepsilon > 0 \text{ existiert mit } (0,\varepsilon) \subset A, \\ 0 & \text{sonst.} \end{cases}$$

Dann gilt: Aus $A_1, \dots, A_m \in \mathcal{R}$ und $A_n \cap A_\ell = \emptyset$ für $n \neq \ell$ folgt

$$\varrho\Big(\bigcup_{n=1}^{m} A_n\Big) = \sum_{n=1}^{m} \varrho(A_n) \qquad \text{(endlich additiv);}$$

ϱ ist aber nicht σ–additiv.

5.11 Jedes $f \in \mathcal{L}_1(X,\mu)$ läßt sich schreiben in der Form $f = f_1 - f_2 + i(f_3 - f_4)$, wobei für jedes f_j eine nicht–fallende $\| \cdot \|_1$–Cauchyfolge $(g_n^{(j)})$ aus $E(X, \mathcal{R}_X)$ existiert mit $g_n^{(j)}(x) \to f_j(x)$ μ–f. ü.

6 Distributionen, verallgemeinerte Funktionen

Distributionen sind stetige lineare Funktionale auf Räumen aus glatten Funktionen mit geeigneten Topologien (sogenannten *Testräumen* oder *Grundräumen*) . Wir führen zunächst die wichtigsten Testräume ein, und untersuchen dann einige grundlegende Operationen für Distributionen. Selbstverständlich können wir hier die Theorie der Distributionen nur ganz knapp anreißen. Für weitere Informationen sei z. B. auf F. Constantinescu [2] und L. Jantscher [5] verwiesen.

6.1 Der Grundraum $\mathcal{D}(K)$

Für kompakte Teilmengen $K \subset \mathbb{R}^m$ (kurz $K \subset\subset \mathbb{R}^m$) sei

$$\mathcal{D}(K) := \Big\{ f \in C^\infty(\mathbb{R}^m) : f(x) = 0 \ \text{ für } x \notin K \Big\}.$$

Es ist zunächst gar nicht leicht, sich viele solcher Funktionen vorzustellen. Ein wichtiges Beispiel ist

$$\varphi_\varepsilon(x) := \begin{cases} c_\varepsilon \exp\left(\dfrac{1}{|x|^2 - \varepsilon^2}\right) & \text{für } |x| < \varepsilon, \\ 0 & \text{für } |x| \geq \varepsilon, \end{cases}$$

wobei c_ε so gewählt wird, daß $\int \varphi_\varepsilon(x)\,\mathrm{d}x = 1$ wird. Diese Funktionen sind in $|x| < \varepsilon$ und $|x| > \varepsilon$ offensichtlich beliebig oft differenzierbar. Daß dies, auch an der Sphäre $|x| = \varepsilon$ gilt, folgt leicht aus der Tatsache, daß jede Ableitung von φ_ε in $|x| < \varepsilon$ geschrieben werden kann als Produkt

- einer gebrochen rationalen Funktion, in deren Nenner eine Potenz von $|x|^2 - \varepsilon^2$ steht und
- dem Term $\exp\left(\frac{1}{|x|^2-\varepsilon^2}\right)$.

Weitere glatte Funktionen mit kompaktem Träger erhält man z. B. so: Ist $f \in L_1(\mathbb{R}^m)$ mit kompaktem Träger, so ist

$$f_\varepsilon(x) := \int \varphi_\varepsilon(x-y) f(y)\,\mathrm{d}y \ \text{ für alle } x \in \mathbb{R}^m$$

eine beliebig oft differenzierbare Funktion mit kompaktem Träger (der Träger von f_ε enthält nur Punkte, die höchstens ε vom Träger von f entfernt sind). In

Kapitel 8 werden wir u.a. sehen, daß für $f \in L_p(\mathbb{R}^m)$ gilt $f_\varepsilon \to f$ im Sinne von $L_p(\mathbb{R}^m)$ für $\varepsilon \to 0$.

Die *Topologie* auf $\mathcal{D}(K)$ wird erzeugt durch die (Halb–)Normen

$$\|f\|_k := \max\Big\{|D^\alpha f(x)| : x \in K, |\alpha| \leq k\Big\}.$$

Dies macht $\mathcal{D}(K)$ zu einem lokalkonvexen topologischen Vektorraum $(\mathcal{D}(K), \|\cdot\|_k (k \in \mathbb{N}))$. Dieser ist metrisierbar mit

$$d(f,g) := \sum_{k\in\mathbb{N}} 2^{-k} \frac{\|f-g\|_k}{1+\|f-g\|_k}.$$

Satz 6.1 $(\mathcal{D}(K), \|\cdot\|_k (k \in \mathbb{N}))$ *ist ein vollständiger Montelraum (jede abgeschlossene beschränkte Menge ist kompakt, vgl. Satz von Montel in der Funktionentheorie. – Es sei daran erinnert, was beschränkt heißt: Eine Teilmenge A eines topologischen Verktorraumes X heißt beschränkt, wenn zu jeder Nullumgebung U in X ein $\lambda \in \mathbb{R}$ existiert mit $A \subset \lambda U$. Dies gilt genau dann, wenn jede Norm $\|\cdot\|_k$ auf U beschränkt ist.).*

Beweis. a) Für eine Cauchyfolge (f_n) (bezüglich der Metrik d) sind alle Ableitungen $(D^\alpha f_n)$ gleichmäßig konvergent gegen Funktionen $f^{(\alpha)}$. Durch „zurückintegrieren“rkennt man, daß für alle α gilt $f^{(\alpha)} = D^\alpha f^{(0)} = D^\alpha f$.

b) Ist A eine beschränkte Teilmenge von $\mathcal{D}(K)$, so ist $\{\|f\|_k : f \in A\}$ beschränkt für jedes k. Für jedes α sind also $D^\alpha f_n$ $(n \in \mathbb{N})$ gleichgradig stetig. Mit dem Satz von Arzela–Ascoli (vgl. Analysis III) und einem Diagonalfolgen–Verfahren erhält man eine bezüglich der Metrik d konvergente Teilfolge. ■

6.2 Der Grundraum $\mathcal{D}(\Omega)$

Sei nun $\Omega \subset \mathbb{R}^m$ offen (insbes. auch $\Omega = \mathbb{R}^m$),

$$\mathcal{D}(\Omega) := \Big\{f \in C_0^\infty(\mathbb{R}^m) : \operatorname{supp} f \subset\subset \Omega\Big\} = \bigcup_{K\subset\subset\Omega} \mathcal{D}(K).$$

Sei $\mathcal{B}$ die Familie der Teilmengen V von $\mathcal{D}(K)$ mit

(a) jedes $V \in \mathcal{B}$ ist kreisförmig und absorbierend,

(b) für jedes $K \subset\subset \Omega$ und jedes $V \in \mathcal{B}$ ist $V \cap \mathcal{D}(K)$ eine Nullumgebung in $\mathcal{D}(K)$ im Sinne des obigen Abschnitts.

$\mathcal{B}$ hat die Eigenschaften einer Nullumgebungsbasis in $\mathcal{D}(\Omega)$, definiert also eine lokalkonvexe Topologie auf $\mathcal{D}(\Omega)$. Für dieses gilt (was nicht schwer zu beweisen ist; vgl. z. Jantscher [5]):

Satz 6.2 *(a) Eine lineare Abbildung $F : \mathcal{D}(\Omega) \to \mathbb{C}$ ist genau dann stetig, wenn sie auf jedem $\mathcal{D}(K)$ mit $K \subset\subset \Omega$ stetig ist.*

(b) Es gilt $\varphi_n \to 0$ in $\mathcal{D}(\Omega)$ genau dann, wenn gilt

i. es existiert ein $K \subset\subset \Omega$ mit $\varphi_n \in \mathcal{D}(K)$ für alle n,

ii. $D^\alpha \varphi_n \to 0$ gleichmäßig in K für alle α.

(Entsprechend gilt also $\varphi_n \to \varphi$ in $\mathcal{D}(\Omega)$, wenn ein $K \subset\subset \Omega$ existiert mit $\varphi_n, \varphi \in \mathcal{D}(K)$ und $D^\alpha \varphi_n \to D^\alpha \varphi$ gleichmäßig für jedes α.)

6.3 Der Grundraum $\mathcal{E}(\Omega)$

Sei wieder $\Omega \subset \mathbb{R}^m$ offen (einschließlich $\Omega = \mathbb{R}^m$),

$$\mathcal{E}(\Omega) := \Big\{ f \in C^\infty(\Omega) \Big\}$$

(ohne eine Bedingung über das Verhalten am Rand von Ω bzw. bei Unendlich). Mit einer Folge (K_n) kompakter Teilmengen von Ω mit $K_n \subset K_{n+1}$ und $\cup_n K_n = \Omega$ definieren wir (Halb–)Normen

$$\|\varphi\|_n := \max \Big\{ |D^\alpha \varphi(x)| : |\alpha| \le n, x \in K_n \Big\}.$$

Damit wird auch $\mathcal{E}(\Omega)$ zu einem lokalkonvexen Raum; dieser ist metrisierbar mit der Metrik

$$d(f,g) := \sum_k 2^{-k} \frac{\|f-g\|_k}{1 + \|f-g\|_k}.$$

Satz 6.3 *In $\mathcal{E}(\Omega)$ gilt: $f_n \to f \Leftrightarrow$ auf jedem Kompaktum $K \subset\subset \Omega$ konvergieren alle Ableitungen gleichmäßig. $\mathcal{E}(\Omega)$ ist vollständig.*

6.4 Der Schwartzsche Raum

Ein weiterer wichtiger Raum auf $\mathbb{R}^m$ ist der *Schwartzsche Raum der schnell fallenden Funktionen*

$$\mathcal{S}(\mathbb{R}^m) := \Big\{ f \in C^\infty(\mathbb{R}^m) : \forall \alpha \in \mathbb{N}_0^m, k \in \mathbb{N}_0 \ \exists C_{\alpha k} \geq 0 \ \text{ mit } |D^\alpha f(x)| \leq C_{\alpha k}(1+|x|)^{-k} \Big\}.$$

Mit den (Halb–)Normen

$$\|f\|_k := \sup \Big\{ |D^\alpha f(x)|(1+|x|)^k : |\alpha| \leq k, x \in \mathbb{R}^m \Big\}$$

wird $\mathcal{S}(\mathbb{R}^m)$ ein lokal konvexer Raum, dessen Topologie wieder durch eine Metrik erzeugt werden kann.

Satz 6.4 *In $\mathcal{S}(\mathbb{R}^m)$ gilt: $f_n \to f \Leftrightarrow$ für alle α und k gilt $|D^\alpha f_n(x) - D^\alpha f(x)|(1+|x|)^k \to 0$ für $n \to \infty$. $\mathcal{S}(\mathbb{R}^m)$ ist vollständig.*

Die Bedeutung des Schwartzschen Raumes beruht darauf, daß er durch die *Fouriertransformation* auf sich abgebildet wird. Es ist leicht zu sehen, daß

$$(Ff)(x) := \frac{1}{(2\pi)^{m/2}} \int e^{-ixy} f(y)\,\mathrm{d}y$$

für $f \in \mathcal{S}(\mathbb{R}^m)$ wohldefiniert und beliebig oft differenzierbar ist.

Satz 6.5 *Für $\varphi \in \mathcal{S}(\mathbb{R}^m)$ gilt (mit $M_\beta =$ Multiplikation mit x^β)*

$$F(D^\alpha \varphi) = (i)^{|\alpha|} M_\alpha F(\varphi), \quad F(M_\beta \varphi) = (i)^{|\beta|} D^\beta F(\varphi).$$

F bildet $\mathcal{S}(\mathbb{R}^m)$ bijektiv auf sich ab, und es gilt

$$(F^{-1}\varphi)(x) = \frac{1}{(2\pi)^{m/2}} \int e^{ixy} \varphi(y)\,\mathrm{d}y.$$

Insbesondere gilt $F^2\varphi(|x|) = \varphi(-x)$.

Beweis. Die beiden Identitäten rechnet man leicht nach:

$$\begin{aligned} F(D^\alpha\varphi)(x) &= \frac{1}{(2\pi)^{m/2}} \int e^{-ixy} D^\alpha \varphi(y)\,\mathrm{d}y = \frac{(-1)^{|\alpha|}}{(2\pi)^{m/2}} \int \Big(D_y^\alpha e^{-ixy} \Big) \varphi(y)\,\mathrm{d}y \\ &= \frac{(i)^{|\alpha|}}{(2\pi)^{m/2}} \int x^\alpha e^{-ixy} \varphi(y)\,\mathrm{d}y = (i)^{|\alpha|} x_\alpha (F\varphi)(x), \\ F(M_\beta\varphi)(x) &= \frac{1}{(2\pi)^{m/2}} \int e^{-ixy} y^\beta \varphi(y)\,\mathrm{d}y = \frac{(-i)^{-|\beta|}}{(2\pi)^{m/2}} \int (D_x^\beta e^{-ixy}) \varphi(y)\,\mathrm{d}y \\ &= (i)^{|\beta|} D^\beta (F\varphi)(x). \end{aligned}$$

Daraus folgt leicht, daß $F\varphi \in \mathcal{S}(\mathbb{R}^m)$ ist für jedes $\varphi \in \mathcal{S}(\mathbb{R}^m)$. Für den Rest des Beweises vgl. man J. Weidmann [11], Satz 11.5 (man beachte, daß obige Formeln dort etwas anders aussehen, weil dort $D^\alpha = (-i)^{|\alpha|}\, \partial^\alpha/\partial x^\alpha$ gesetzt ist).. ■

6.5 Reguläre Distributionen

Die stetigen linearen Funktionale auf den oben beschriebenen Räumen heißen *Distributionen*. Spezielle Distributionen werden durch Funktionen erzeugt, genauer:

$\mathcal{D}(K)$: Ist $f \in L_1(K)$, so ist offenbar T_f mit $T_f(\varphi) = \int_K f(x)\varphi(x)\,\mathrm{d}x$ stetig auf $\mathcal{D}(K)$, $|T_f(\varphi)| \le \|f\|_1 \max|\varphi| = \|f\|_1\|\varphi\|_0$ mit $\|\cdot\|_0$ gemäß Abschnitt 1.

$\mathcal{D}(\Omega)$: Ist $f \in L_{1,\mathrm{lok}}(\Omega)$, so ist das entsprechend definierte T_f stetig, da es auf jedem $\mathcal{D}(K)$ mit $K \subset\subset \Omega$ stetig ist.

$\mathcal{E}(\Omega)$: Ist $f \in L_1(\Omega)$ mit kompaktem Träger in Ω, so ist T_f stetig auf $\mathcal{E}(\Omega)$. Allgemein gilt, daß für ein $T \in \mathcal{E}(\Omega)'$ der Wert von $T(\varphi)$ nicht vom Verhalten von φ außerhalb einer (von T abhängigen) kompakten Teilmenge von Ω abhängt; man nennt diese Distributionen *Distributionen mit kompaktem Träger*.

$\mathcal{S}(\mathbb{R}^m)$: Ist $f \in L_{1,\mathrm{lok}}$ und bei ∞ höchstens polynominal wachsend, so ist T_f stetig auf $\mathcal{S}(\mathbb{R}^m)$. Auf Grund dieser Tatsache nennt man die Distributionen aus $\mathcal{S}(\mathbb{R}^m)'$ *temperierte Distributionen*.

Diese durch Funktionen erzeugten Distributionen heißen *reguläre Distributionen*. Wir kommen weiter unten darauf zurück.

6.6 Differentiation und Multiplikation von Distributionen

Ist T_f eine reguläre Distribution mit einer glatten erzeugenden Funktion, so gilt (auf Grund des jeweiligen Verhaltens von f bzw. φ treten bei der partiellen Integration keine Randterme auf)

$$T_{D^\alpha f}(\varphi) = \int (D^\alpha f)(x)\varphi(x)\,\mathrm{d}x = (-1)^{|\alpha|} \int f(x) D^\alpha \varphi(x)\,\mathrm{d}x.$$

Man definiert deshalb für eine beliebige Distribution T die Ableitung $D^\alpha T$ durch

$$(D^\alpha T)(\varphi) := (-1)^{|\alpha|} T(D^\alpha \varphi).$$

Da in jedem Fall $\varphi \mapsto D^\alpha \varphi$ eine stetige Abbildung des Grundraumes in sich ist (Beweis!), ist $D^\alpha T$ wieder eine Distribution. Bemerkenswert ist, daß jede Distribution in diesem Sinn beliebig oft differenzierbar ist und daß die Ableitungen beliebig vertauschbar sind. (Wendet man diese Art der Ableitung auf reguläre Distributionen an, so könnte dies zu Verwirrungen führen, da die partiellen Ableitungen von Funktionen i. allg. nicht vertauschbar sind. Man redet deshalb von „Ableitungen im Sinne der Distributionentheorie“. Bei stetig differenzierbaren Funktionen stimmen natürlich beide Ableitungen überein.)

Die *Ordnung* einer Distribution F ist das kleinste $k \in \mathbb{N}_0$, für das F als Ableitung der Ordnung k einer regulären Distribution dargestellt werden kann; eine reguläre Distribution hat also die Ordnung 0.

Auf allen genannten Grundräumen ist die DIRACSCHE *Deltadistribution* (häufig als δ–*Funktion* bezeichnet) eine Distribution: $\delta(\varphi) := \varphi(0)$ (oder allgemeiner $\delta_a(\varphi) = \varphi(a)$.

δ ist nicht regulär: Gäbe es ein lokal integrierbares f, das δ erzeugt, so würde mit

$$\psi_\varepsilon(x) = \begin{cases} \exp\left(\frac{1}{\varepsilon^2} + \frac{1}{|x|^2-\varepsilon^2}\right\} & \text{für } |x| < \varepsilon \\ 0 & \text{für } |x| \geq 0 \end{cases}$$

gelten

$$\int f(x)\psi_\varepsilon(x)\,\mathrm{d}x = \psi_\varepsilon(0) = 1.$$

Für $\varepsilon \to 0$ liefert der Satz von Lebesgue

$$0 = \lim_{\varepsilon \to 0} \int f(x)\psi_\varepsilon(x)\,\mathrm{d}x = 1,$$

ein Widerspruch.

Andererseits ist in $\mathbb{R}$ die δ–Distribution die erste Ableitung der durch die *Heaviside–Funktion*

$$H(x) = \begin{cases} 0 & \text{für } x < 0, \\ 1 & \text{für } x \geq 0, \end{cases}$$

$$(T_H)'(\varphi) = -\int H(x)\varphi'(x)\,\mathrm{d}x = -\int_0^\infty \varphi'(x)\,\mathrm{d}x = \varphi(0) = \delta(\varphi).$$

erzeugten Distribution. δ ist also eine Distribution der Ordnung 1.

Die *Multiplikation von Distributionen* ist i. allg. schwierig. Einfach geht es, mit glatten Funktionen (die ggf. noch geeignetes Verhalten am Rand/bei Unendlich haben) zu multiplizieren. Weil man bei regulären Distributionen die Multiplikation auf die Testfunktionen überwälzen kann, definiert man die Distribution ψT durch

$$(\psi T)(\varphi) := T(\psi\varphi).$$

6.7 Fouriertransformation in $\mathcal{S}'(\mathbb{R}^m)$

Für $T \in \mathcal{S}(\mathbb{R}^m)$ definiert man die Fouriertransformierte FT durch

$$FT(\varphi) = T(F\varphi).$$

Auch diese Definition wird durch die entsprechende Formel für T_{Ff} nahegelegt, wenn man die Funktion f Fourier–transformieren kann.

Zusammen mit den früheren Formeln für die Fouriertransformierte von Ableitungen von Testfunktionen erhält man hieraus

$$F(D^\alpha T) = (i)^{|\alpha|} x^\alpha F(T).$$

Als Beispiel soll die Fouriertransformierte von δ berechnet werden:

$$\begin{aligned}(F\delta)(\varphi) = \delta(F\varphi) &= \frac{1}{(2\pi)^{m/2}} \int e^{-ixy}\varphi(y)\,\mathrm{d}y\Big|_{x=0} \\ &= \frac{1}{(2\pi)^{m/2}} \int \varphi(y)\,\mathrm{d}y,\end{aligned}$$

d. h. $F\delta$ ist die durch die konstante Funktion $C := (2\pi)^{-m/2}$ erzeugte reguläre Distribution. Zusammen mit Satz 6.5 erhält man daraus (mit $\varphi_-(x) := \varphi(-x)$)

$$(FC)(\varphi) = (F^2\delta)(\varphi) = \delta(F^2\varphi) = \delta(\varphi_-) = \varphi(0),$$

also $FC = \delta$.

Hier wird die Bedeutung der Distributionentheorie für die Theorie der partiellen Differentialgleichungen erkennbar: Eine Differentialgleichung mit konstanten Koeffizienten

$$\sum a_\alpha D^\alpha f = g$$

geht durch Fouriertransformation über in

$$\sum (i)^{|\alpha|} a_\alpha x^\alpha F(f) = F(g).$$

Wenn man es beherrscht, $F(g)$ durch ein Polynom zu dividieren, erhält man daraus eine Formel für $F(f)$.

7 L_p–Räume

Im folgenden sei, wenn nichts anderes vorausgesetzt wird, μ ein beliebiges Maß auf einer Menge X; von besonderem Interesse ist der Fall, in dem X eine Lebesgue–meßbare Teilmenge von R^m ist und μ das Lebesgue–Maß. Die Räume $\ell_p = \ell_p(\mathbb{N})$ erhält man als Spezialfall, wenn man $X = \mathbb{N}$ wählt und für μ das Zählmaß auf $\mathbb{N}$, d. h. $\mu(A) :=$ Anzahl der Punkte in A; völlig analog kann man $\ell_p(A)$ mit einer beliebigen Menge A betrachten.

7.1 Der Raum $L_\infty(X, \mu)$

Analog zum Raum $C[a, b]$ in Beispiel 4.4 könnte man den Raum der beschränkten μ–meßbaren Funktionen auf X mit der Supremums–Norm versehen. Der Leser überzeugt sich leicht, daß dies einen vollständigen normierten Raum ergeben würde. Im Hinblick auf Anwendungen ist dies völlig uninteressant, da man Funktionen als verschieden betrachtet, die μ–fast überall übereinstimmen; es ist aber meistens nicht sinnvoll, solche Funktionen zu unterscheiden. Man geht deshalb etwas anders vor:

Sei zunächst

$$\mathcal{L}_\infty(X, \mu) := \{f : X \to \mathbb{K}\ \mu\text{–meßbar},\ \exists\ C_f \text{ mit } |f(x)| \leq C_f\ \mu\text{–f. ü.}\}$$

Für jedes $f \in \mathcal{L}_\infty(X, \mu)$ gibt es also eine μ–Nullmenge N_f mit $|f(x)| \leq C_f$ für $x \in X \backslash N_f$; die Elemente von $\mathcal{L}_\infty(X, \mu)$ heißen *wesentlich beschränkt*. Es ist offensichtlich, daß $\mathcal{L}_\infty(X; \mu)$ ein Vektorraum ist. Auf $\mathcal{L}_\infty(X, \mu)$ definiert man

$$\begin{aligned} \|f\|_\infty &:= \text{wes–}\sup |f(x)| \ (\textit{wesentliches Supremum}) \\ &:= \inf \Big\{C : |f(x)| \leq C\ \mu\text{–f. ü.}\Big\}. \end{aligned}$$

Offenbar gilt

$$\|f\|_\infty \geq 0, \quad \|af\|_\infty = |a|\|f\|_\infty, \quad \|f + g\|_\infty \leq \|f\|_\infty + \|g\|_\infty;$$

die letzte Ungleichung ergibt sich aus

$$|f(x) + g(x)| \leq |f(x)| + |g(x)| \leq \|f\|_\infty + \|g\|_\infty\text{–f. ü.}$$

Also ist $\| \cdot \|_\infty$ eine *Halbnorm*. Da jedoch aus $\|f\|_\infty = 0$ nur folgt $f(x) = 0$ μ–f. ü., aber *nicht* $f(x) = 0$ für *alle* x, ist $\| \cdot \|_\infty$ i. allg. *keine* Norm auf $\mathcal{L}_\infty(X, \mu)$.

Sei

$$\begin{aligned} \mathcal{N}(X,\mu) &:= \{f : X \to \mathbb{K} \text{ mit } f(x) = 0\ \mu\text{–f. ü.}\} \\ &:= \{f \in \mathcal{L}_\infty(X,\mu) : \|f\|_\infty = 0\}. \end{aligned}$$

Offensichtlich ist $\mathcal{N}(X,\mu)$ ein Teilraum von $\mathcal{L}_\infty(X,\mu)$ und $\|\cdot\|_\infty$ definiert eine *Norm* auf dem Quotientenraum

$$L_\infty(X,\mu) := \mathcal{L}_\infty(X,\mu)/\mathcal{N}(X,\mu).$$

Es ist zu beachten, daß $L_\infty(X,\mu)$ streng genommen kein Raum von Funktionen ist, sondern ein Raum von Äquivalenzklassen von Funktionen. Zwei wesentlich beschränkte μ–meßbare Funktionen repräsentieren das gleiche Element von $L_\infty(X,\mu)$, wenn sie μ–f. ü. übereinstimmen. Alle Operationen in $L_\infty(X,\mu)$ werden mit Hilfe von Repräsentanten definiert; es ist dabei darauf zu achten (was wir meist nicht weiter erwähnen werden), daß die Definition von der Wahl des Repräsentanten unabhängig ist.

Je nachdem, ob $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ ist, ist $L_\infty(X,\mu)$ ein *reeller* oder ein *komplexer* Raum. Meist wird der komplexe Fall betrachtet.

Satz 7.1 *$L_\infty(X,\mu)$ ist vollständig.*

Beweis. Sei (f_n) eine $\|\cdot\|_\infty$–Cauchyfolge in $L_\infty(X,\mu)$ (wir bezeichnen mit f_n nicht nur das L_∞–Element [die Äquivalenzklasse], sondern auch einen Repräsentanten von f_n). Dann gilt: Für jedes $k \in \mathbb{N}$ existiert ein $N(k) \in \mathbb{N}$ so, daß für alle $n, m \geq N(k)$ eine μ–Nullmenge $N_{n,m,k}$ existiert mit

$$|f_n(x) - f_m(x)| < \frac{1}{k} \text{ für } x \in X \backslash N_{n,m,k}.$$

Die Folge (f_n) ist also außerhalb der μ–Nullmenge

$$N := \bigcup_{\substack{k\in\mathbb{N}\\ n,m\geq N(k)}} N_{n,m,k}$$

gleichmäßig konvergent gegen eine μ–meßbare Funktion

$$\tilde{f} : X\backslash N \to \mathbb{K}.$$

Setzen wir (z. B.)

$$f(x) = \begin{cases} \tilde{f}(x) & \text{für } x \in X \backslash N, \\ 0 & \text{für } x \in N, \end{cases}$$

so ist f μ–meßbar (Satz 5.16), wesentlich beschränkt, und es gilt $\|f_n - f\|_\infty \to 0$ für $n \to \infty$ (wobei hier mit f auch die durch f erzeugte Äquivalenzklasse bezeichnet wird.). ■

7.2 Die Räume $L_p(X, \mu)$ für $1 \leq p < \infty$

Sei zunächst

$$\mathcal{L}_p(X,\mu) := \Big\{ f : X \to \mathbb{K} \ \mu\text{–meßbar und } |f|^p \mu \text{ –integrierbar} \Big\},$$
$$\|f\|_p := \Big\{ \int\limits_X |f(x)|^p \, \mathrm{d}\mu(x) \Big\}^{1/p} \quad \text{für } f \in \mathcal{L}_p(X,\mu).$$

Wegen $|f(x) + g(x)|^p \leq 2^p\{|f(x)|^p + |g(x)|^p\}$ ist $\mathcal{L}_p(X, \mu)$ ein Vektorraum. Offensichtlich gilt

$$\|f\|_p \geq 0 \ \text{ und } (N1) : \|af\|_p = |a| \|f\|_p.$$

Um die Dreiecksgleichung (N 2) für $\|\cdot\|_p$ zu beweisen, benötigen wir (wie bei den ℓ_p–Räumen, Abschnitt 4.3) die Höldersche Ungleichung.

Satz 7.2 (Höldersche Ungleichung) *Seien $p, q \in [1, \infty]$ mit $\frac{1}{p} + \frac{1}{q} = 1$ (einschließlich $p = 1, q = \infty$ und $p = \infty, q = 1$). Für $f \in \mathcal{L}_p(X, \mu)$ und $g \in \mathcal{L}_q(X, \mu)$ gilt*

$$fg \in \mathcal{L}^1(X, \mu) \ \ \textit{und} \ \int |f(x)g(x)| \, \mathrm{d}\mu(x) \leq \|f\|_p \|g\|_q.$$

Beweis. Aus Lemma 4.9 folgt mit

$$a = \|f\|_p^{-1} |f(x)|, \quad b = \|g\|_q^{-1} |g(x)|$$

offenbar

$$(\|f\|_p\|g\|_q)^{-1}|f(x)g(x)| \le \frac{1}{p}\|f\|_p^{-p}|f(x)|^p + \frac{1}{q}\|g\|_q^{-q}|g(x)|^q.$$

Integration liefert auf der rechten Seite 1 und somit offenbar die Höldersche Ungleichung für $1 < p, q < \infty$. Die Grenzfälle $p = 1, q = \infty$ und $p = \infty, p = 1$ sind offensichtlich. ■

Völlig analog wie in ℓ_p können wir nun die *Dreiecksungleichung* für $\|\cdot\|_p$ beweisen:

Satz 7.3 (Minkowskische Ungleichung in $\mathcal{L}^p$) *Für $f, g \in L^p(X,\mu)$ gilt $\|f+g\|_p \le \|f\|_p + \|g\|_p$.*

Beweis. Für $f, g \in \mathcal{L}_p(X,\mu)$ folgt aus der Hölderschen Ungleichung

$$\begin{aligned}
&\|f+g\|_p^p = \int |f+g|^p \,\mathrm{d}\mu \le \int \Big\{|f||f+g|^{p-1} + |g||f+g|^{p-1}\Big\}\,\mathrm{d}\mu \\
&\le \Big\{\int |f|^p \,\mathrm{d}\mu\Big\}^{1/p}\Big\{\int |f+g|^{(p-1)q}\,\mathrm{d}\mu\Big\}^{1/q} + \Big\{\int |g|^p\,\mathrm{d}\mu\Big\}^{1/p}\Big\{\int |f+g|^{(p-1)q}\,\mathrm{d}\mu\Big\}^{1/q} \\
&= (\|f\|_p + \|g\|_p)\|f+g\|_p^{p/q}, \\
&\|f+g\|_p^{p-p/q} \le \|f\|_q + \|g\|_p.
\end{aligned}$$

Wegen $p - p/q = 1$ ist dies die gewünschte Ungleichung. ■

Damit ist also $\|\cdot\|_p$ eine *Halbnorm*, aber wieder i. allg. keine Norm, da (N 3) nicht erfüllt ist. Wie im Fall $p = \infty$ wird das Problem durch Einführung des Quotientenraumes

$$L_p(x,\mu) := \mathcal{L}_p(X,\mu)/\mathcal{N}(X,\mu)$$

behoben. $\|\cdot\|_p$ ist auf $L_p(X,\mu)$ eine *Norm*. Auch dieser Raum ist ein Raum von Äquivalenzklassen von Funktionen; wir verweisen aus die entsprechende Anmerkung in Abschnitt 6.1.

Satz 7.4 *Sei $1 \le p < \infty$. $L_p(X,\mu)$ ist vollständig. Jede Cauchyfolge in $L_p(X,\mu)$ enthält eine μ–f. ü. konvergente Teilfolge (die Folge selbst ist i. allg. nicht f. ü. konvergent, vgl. Aufgabe 6.7a).*

Beweis. Sei (f_n) eine Cauchyfolge in $L_p(X,\mu)$; wieder bezeichnen wir mit f_n gleichzeitig Repräsentanten von f_n aus $\mathcal{L}_p(X,\mu)$. Für jedes $j \in \mathbb{N}$ existiert ein $n_j \in \mathbb{N}$ mit

$$\|f_n - f_m\|_p \leq 2^{-j} \text{ für } n, m \geq n_j$$

also insbesondere

$$\|f_{n_{j+1}} - f_{n_j}\|_p \leq 2^{-j} \text{ für alle } j.$$

Sei

$$g_k(x) := \sum_{j=1}^{k} \Big| f_{n_{j+1}}(x) - f_{n_j}(x) \Big|.$$

Dann ist die Folge $(g_k(\cdot)^p)_k$ nichtfallend und wegen der Dreiecksungleichung für $\|\cdot\|_p$ gilt

$$\int g_k(x)^p \, \mathrm{d}\mu(x) = \|g_k\|_p^p \leq \Big\{ \sum_{j=1}^{k} 2^{-j} \Big\}^p < 1\,.$$

Nach dem Satz von B.Levi (Satz 5.9) ist also die Folge $(g_k(x)^p)_k$, und somit auch die Folge $(g_k(x))_k$, μ–f.ü. konvergent. Dann ist auch

$$f_{n_k} = f_{n_1} + \sum_{j=1}^{k-1} \Big(f_{n_{j+1}} - f_{n_j} \Big)$$

μ–f.ü. konvergent gegen eine μ–meßbare Funktion f. Wir zeigen

$\mathbf{f \in \mathcal{L}_p(X,\mu)}$ **und** $\mathbf{\|f_n - f\|_p \to 0}$: Für jedes $\varepsilon > 0$ seien $n(\varepsilon)$ und $j(\varepsilon)$ so gewählt, daß

$$\Big| \int |f_{n_j}(x) - f_n(x)\Big|^p \mathrm{d}\mu(x) = \|f_{n_j} - f_n\|_p^p < \varepsilon \text{ für } j > j(\varepsilon) \text{ und } n > n(\varepsilon)\,.$$

Die Folge $(|f_{n_j}(x) - f_n(x)|^p)_j$ ist also nicht–negativ mit beschränkter Integralfolge und es gilt für $j \to \infty$

$$|f_{n_j}(x) - f_n(x)|^p \to |f(x) - f_n(x)|^p \ \mu\text{–f.ü.}$$

Nach dem Lemma von Fatou (Satz 5.12) ist also

$$|f - f_n|^p \ \mu\text{–integrierbar}, \qquad \int |f - f_n|^p \, \mathrm{d}\mu \leq \varepsilon \text{ für } n > n(\varepsilon).$$

Insbesondere ist $f \in \mathcal{L}_p(X, \mu)$ und $\|f_n - f\|_p^p \leq \varepsilon$ für $n \geq n(\varepsilon)$. Mit dem durch f rzeugten Element aus $L_p(X; \mu)$ ergibt dies die Vollständigkeit von $L_p(X, \mu)$. ∎

7.3 Dichte Teilräume von $L_p(X, \mu)$

Sei $\mathcal{R}$ ein Mengenring in X so, daß jedes $A \in \mathcal{R}$ μ–meßbar ist mit $\mu(A) < \infty$ und die Einschränkung von μ auf $\mathcal{R}$ das Maß μ erzeugt. (vgl. §5).

Satz 7.5 *Der Raum der $\mathcal{R}$–Elementarfunktionen $E(X, \mathcal{R})$ ist dicht in $L_p(X; \mu)$ für $1 \leq p < \infty$.*

Beweis. (Für $p = 1$ ist dies die Definition der μ–integrierbaren Funktionen).

a) Sei $\mathcal{F}_p(X, \mu)$ der Teilraum der Funktionen aus $\mathcal{L}_p(X, \mu)$ für die gilt:

(a) es gibt eine meßbare Teilmenge $X_f \subset X$ mit $\mu(X_f) < \infty$ und $f(x) = 0$ μ–f. ü. in $X \backslash X_f$ („finite“ Funktionen),

(b) f ist wesentlich beschränkt.

Wir zeigen, daß $\mathcal{F}_p(X, \mu)$ in $\mathcal{L}_p(X, \mu)$ dicht ist: Für $f \in \mathcal{L}_p(X, \mu)$ und $n \in \mathbb{N}$ sei

$$f_n(x) := \begin{cases} f(x) & \text{falls } \dfrac{1}{n} \leq |f(x)| \leq n, \\ 0 & \text{sonst.} \end{cases}$$

Wegen $|f_n(x)| \leq n$ und $\mu(\{x \in X : f_n(x) \neq 0\}) = \mu(\{x \in X : |f(x)| \geq \dfrac{1}{n}\}) < \infty$ ist $f_n \in \mathcal{F}_p(X, \mu)$. Aus

$$|f_n - f|^p \leq |f|^p \text{ und } |f_n - f|^p \to 0 \ \mu\text{–f. ü.}$$

folgt mit Hilfe des Satzes von Lebesgue (Satz 5.11)

$$\|f_n - f\|_p \to 0.$$

b) $E(X,\mathcal{R})$ ist dicht in $\mathcal{F}_p(X;\mu)$: Sei $f \in \mathcal{F}_p(X,\mu)$. Dann ist $f \in \mathcal{L}_1(X,\mu)$ und es gibt eine Folge (f_n) aus $E(X,\mathcal{R})$ mit $\|f_n - f\|_p \to 0$; offenbar kann man ohne Einschränkung annehmen, daß gilt

$$|f_n(x)| \leq C := \text{wes–sup}|f|.$$

Damit folgt

$$\|f_n - f\|_p^p = \int |f_n - f|^p \, \mathrm{d}\mu \leq (2C)^{p-1} \int |f_n - f| \, \mathrm{d}\mu \to 0.$$

∎

Im folgenden Satz spezialisieren wir unsere Überlegungen auf den Fall, in dem $X = \Omega$ eine Lebesgue–meßbare Teilmenge von $\mathbb{R}^m$ ist und μ das Lebesguesche Maß / λ. Wir schreiben dann kurz $L_2(\Omega)$ statt $L_2(\Omega,\lambda)$.

Satz 7.6 *Sei* $1 \leq p < \infty, \Omega \subset \mathbb{R}^m$ *Lebesgue–meßbar.*

(a) $L_p(\Omega)$ *ist separabel (vgl. Aufgabe 6.1 für* $p = \infty$*).*

(b) Ist Ω *offen, so ist* $C_0^\infty(\Omega)$ *dicht in* $L_p(\Omega)$*, wobei*

$$\begin{aligned} C_0^\infty(\Omega) \;\; := \;\; \Big\{ & f : \Omega \to \mathbb{C}, f \text{ beliebig oft stetig differenzierbar,} \\ & \operatorname{supp} f \text{ kompakte Teilmenge von } \Omega \Big\}. \end{aligned}$$

Beweis. a) $L_p(\Omega)$ kann als Teilraum von $L_p(\mathbb{R}^m)$ aufgefaßt werden, indem man Funktionen aus $L_p(\Omega)$ außerhalb von Ω durch 0 fortsetzt. Es genügt also (nach Satz 4.1) zu zeigen, daß $L_p(\mathbb{R}^m)$ separabel ist. Dies ergibt sich aber leicht daraus, daß die (abzählbare) Menge der charakteristischen Funktionen von Quadern mit rationalen Eckpunkten in dem dichten Teilraum der Treppenfunktionen $E(\mathbb{R}^m,\mathcal{R})$ total ist ($\mathcal{R}$ = Familie der Figuren).

b) Es genügt zu zeigen, daß jede charakteristische Funktion eines abgeschlossenen Quaders $Q \subset \Omega$ durch C_0^∞–Funktionen approximiert werden kann. Sei

$$\begin{aligned} \tilde{\delta}_\varepsilon(x) &:= \begin{cases} \exp\left(\dfrac{1}{x^2 - \varepsilon^2}\right) & \text{für } |x| < \varepsilon, \\ 0 & \text{sonst,} \end{cases} \\ \delta_\varepsilon(x) &:= \left\{ \int \tilde{\delta}_\varepsilon(y) \, \mathrm{d}y \right\}^{-1} \tilde{\delta}_\varepsilon(x). \end{aligned}$$

Dann ist $\delta_\varepsilon \in C_0^\infty(\mathbb{R}^m)$, $\int \delta_\varepsilon(x)\,\mathrm{d}x = 1$, und für

$$f_\varepsilon(x) = \int\limits_Q \delta_\varepsilon(x-y)\,\mathrm{d}y = \begin{cases} 1 & \text{für } \mathrm{d}(x, \complement Q) \geq \varepsilon, \\ 0 & \text{für } \mathrm{d}(x, Q) \geq \varepsilon, \\ \text{zwischen 0 und 1} & \text{für alle } x \end{cases}$$

gilt $f_\varepsilon \in C_0^\infty(\Omega)$ für hinreichend kleine $\varepsilon > 0$ und

$$\|f_\varepsilon - \chi_Q\|_p \to 0 \text{ für } \varepsilon \to 0.$$

■

7.4 Stetige lineare Funktionale auf $L_p(X,\mu)$ und ℓ_p

Die topologischen Dualräume von $L_p(X,\mu)$ können explizit angegeben werden. Insbesondere gilt für $1 \leq p < \infty$

$$L_p(X,\mu)^* \cong L_p(X,\mu), \qquad \ell_p^* \cong \ell_p \text{ mit } \frac{1}{p} + \frac{1}{q} = 1.$$

An dieser Stelle können wir dieses Resultat nur teilweise beweisen. Der verbleibende Teil wird in §10 nachgetragen.

Satz 7.7 *Sei* $\frac{1}{p} + \frac{1}{q} = 1$ *(einschließlich der Fälle* $p = 1, 1 = \infty$ *und* $p = \infty,\ q = 1$*). Jedes* $f \in L_q(X,\mu)$ *erzeugt durch*

$$F(g) = F_f(g) := \int f(x)g(x)\,\mathrm{d}\mu(x) \text{ für } g \in L_p(X,\mu)$$

ein stetiges lineares Funktional auf $L_p(X,\mu)$ *mit*

$$\|F\| = \|f\|_q.$$

Beweis. Aus der Hölderschen Ungleichung (Satz 6.2) folgt, daß F tatsächlich auf ganz $L_p(X,\mu)$ definiert ist mit

$$|F(g) \leq \|f\|_q \|g\|_p, \text{ also } \|F\| \leq \|f\|_q.$$

Es genügt also $\|F\| \geq \|f\|_q$ zu beweisen. Ohne Einschränkung können wir dabei $\|f\|_1 \neq 0$ voraussetzen.

$\mathbf{p = 1}$ $(q = \infty)$: Es gibt zu jedem $\varepsilon > 0$ eine μ–meßbare Teilmenge X_ε von X mit

$$\mu(X_\varepsilon) > 0 \text{ und } |f(x)| \geq \|f\|_\infty - \varepsilon \ \text{ in } X_\varepsilon.$$

Für jedes $g \in L_1(X, \mu)$ mit $g \geq 0$, $\|g\|_1 = 1$ und $g(x) = 0$ in $X \backslash X_\varepsilon$ gilt $\|g \cdot \overline{\operatorname{sgn} f}\|_1 = 1$ und

$$\begin{aligned} F(g \cdot \overline{\operatorname{sgn} f}) &= \int_X g(x)|f(x)| \,\mathrm{d}\mu(x) \geq \int_X (\|f\|_\infty - \varepsilon) g(x) \,\mathrm{d}\mu(x) \\ &\geq (\|f\|_\infty - \varepsilon)\|g\|_1 = (\|f\|_\infty - \varepsilon). \end{aligned}$$

Also ist $\|F\| \geq \|f\|_\infty - \varepsilon$. Da dies für jedes $\varepsilon > 0$ gilt, folgt die Behauptung.

$\mathbf{p = \infty}$ $(q = 1)$: mit $g := \overline{\operatorname{sgn} f}$ gilt

$$F(g) = \|f\|_1 = \|f\|_1 \|g\|_\infty,$$

also $\|F\| \geq \|f\|_1$.

$\mathbf{1 < p < \infty}$ $(1 < q < \infty)$: Mit

$$g(x) := \|f\|_q^{-q/p} \overline{\operatorname{sgn} f(x)} |f(x)|^{q/p}$$

gilt offenbar

$$\begin{aligned} \|g\|_p &= \|f\|_q^{-q/p} \Big\{ \int |f(x)|^q \,\mathrm{d}\mu(x) \Big\}^{1/p} = 1, \\ F(g) &= \|f\|_q^{-q/p} \int |f(x)||f(x)|^{q/p} \,\mathrm{d}\mu(x) \\ &= \|f\|^{-q/p} \int |f(x)|^{1+q/p} \,\mathrm{d}\mu(x) = \|f\|_q^{-q/p} \int |f(x)|^q \,\mathrm{d}\mu(x) \\ &= \|f\|_q^{q-q/p} = \|f\|_q = \|g\|_p \|f\|_q, \end{aligned}$$

also $\|F\| \geq \|f\|_q$. ■

Speziell gilt also ($X = \mathbb{N}, \mu =$ Zählmaß):

Satz 7.8 *Sei* $\frac{1}{p} + \frac{1}{q} = 1$ *(einschließlich der Fälle* $p = 1, 1 = \infty$ *und* $p = \infty,\ q = 1$*). Jedes* $f = (f_n) \in \ell_q$ *erzeugt durch*

$$F(g) = F_f(g) := \sum f_n g_n \text{ für } g = (g_n) \in \ell_p$$

ein stetiges lineares Funktional auf ℓ_p *mit*

$$\|F\| = \|f\|_q.$$

Natürlich kann man den *Beweis* auch analog zum Beweis vom Satz 6.6 führen, wobei sich einige Vereinfachungen ergeben.

Erst in §10 werden wir zeigen, daß für $p \in [1, \infty)$ tasächlich jedes stetige lineare Funktional auf $L_p(X, \mu)$ bzw. ℓ_p durch ein Element aus $L_q(X, \mu)$ bzw. ℓ_q erzeugt wird. Der Fall ℓ_p könnte bereits jetzt behandelt werden; für den allgemeineren Fall $L_p(X, \mu)$ benötigen wir ein Resultat aus § 9.

7.5 Übungsaufgaben

7.1 $L_\infty(X, \mu)$ ist entweder endlichdimensional oder nicht separabel. $L_q(X, \mu)$ ist genau dann n–dimensional, wenn es μ–meßbare Teilmengen $X_1, \ldots, X_n$ von X gibt mit $X = \bigcup_{j=1}^{n} X_j$ und $0 < \mu(X_j) < \infty$ so, daß für jede meßbare Teilmenge A_j von X_j gilt $\mu(A_j) = 0$ oder $\mu(A_j) = \mu(X_j)$.

7.2 Sei μ ein Maß auf X.

(a) Ist $1 \leq p, q, r, \leq \infty$, $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$, $f \in L_p(X, \mu)$, $g \in L_q(X, \mu)$, so gilt $fg \in L_r(X, \mu)$ und

$$\|fg\|_r \leq \|f\|_p \|g\|_q \text{ (\textit{verallgemeinerte Höldersche Ungleichung})}$$

Anleitung: Man benutze die Höldersche Ungleichung.

(b) Ist $\mu(X) < \infty$ und $1 \leq r \leq p \leq \infty$, so gilt $\|f\|_r \leq C_{p,r} \|f\|$ für $f \in L_p(X, \mu)$, d. h. insbesondere $L_p(X, \mu) \subset L_r(X, \mu)$.

(c) Dagegen gilt für $1 \leq r \leq p \leq \infty$: $\ell_r \subset \ell_p$ und $\|x\|_p \leq \|x\|_r$ für jedes $x \in \ell_r$. *Anleitung:* Es genügt, die Ungleichung für $x \in \ell_r$ mit $\|x\|_r = 1$ zu beweisen.

(d) Für $\mu(X) = \infty$ gilt i. allg. weder $L_r \subset L_p$ noch $L_p \subset L_r$.

7.3 (a) Man zeige, daß sich nicht jedes stetige lineare Funktional F auf $L_\infty(\mathbb{R})$ durch ein $f \in L_1(\mathbb{R})$ darstellen läßt.

(b) Dies gilt für jeden unendlichdimensionalen $L_\infty(X, \mu)$.

7.4 Sei μ ein Maß auf X, $t : X \to \mathbb{C}$ μ–meßbar.

(a) In $L_p(X, \mu)$ $(1 \le p \le \infty)$ wird durch

$$T : L_p(X, \mu) \to L_p(X, \mu),\ Tf(x) = t(x)f(x)$$

genau dann ein beschränkter linearer Operator definiert, wenn t aus $L_\infty(\mathbb{R}, \mu)$ ist. Es gilt $\|T\| = \|t\|_\infty$.

(b) Wie muß t beschaffen sein, damit ein $f \in L_p(X, \mu), f \neq 0$ existiert mit $\|Tf\|_p = \|t\|_\infty \|f\|_p$?

7.5 (a) In $L_p(0, \infty)$ $(1 \le p \le \infty)$ wird durch

$$Tf(x) = f(x + 1)$$

ein beschränkter Operator T mit $\|T\| = 1$ definiert.

(b) Jedes $z \in \mathbb{C}$ mit $|z| < 1$ $(|z| \le 1$ für $p = \infty)$ ist Eigenwert unendlicher Vielfachheit von T, d. h. es gibt einen unendlichdimensionalen Teilraum N_z von $L_p(0, \infty)$ mit $Tf = zf$ für alle $f \in N_z$.

7.6 Sei μ ein Maß auf X, $f : X \to \mathbb{C}$ μ–meßbar und $fg \in L_1(X, \mu)$ für jedes $g \in L_q(X, \mu)$. Dann ist $f \in L_p(X, \mu)$ mit $\dfrac{1}{p} + \dfrac{1}{q} = 1$.
Anleitung: Man zeige, daß das lineare Funktional $F : L_q(X; \mu) \to \mathbb{C}, F(g) = \int fg \, \mathrm{d}\mu$ stetig ist und benutze Satz 6.6.

7.7 (a) Für $f_n \in L_p(X, \mu)$ und $\|f_n\|_p \to 0$ $(2 \le p \le \infty)$ gilt i. allg. *nicht* $f_n(x) \to 0$ μ–f. ü.
Anleitung: In $L_p(0, 1)$ sei $f_{2^n+k}(x) = \chi_{[(k-1)2^{-n}, k2^{-n})}(x)$ für $n \in \mathbb{N}_0$ und $k \in \{1, \ldots, 2^n\}$.

(b) Für $f \in L_p(X, \mu)$ und $f_n(x) \to 0$ μ–f. ü. gilt i. allg. *nicht* $\|f_n\|_p \to 0$.
Anleitung: In $L_p(0, 1)$ sei $f_n(x) = n\chi_{[0,1/n]}(x)$.

(c) Falls $\mu(x) = \infty$ ist, genügt es auch nicht, in Teil b zusätzlich $|f_n(x)| \le C$ für alle n und x vorauszusetzen.
Anleitung: In $L_p(0, \infty)$ sei $f_n(x) = \chi_{[n,n+1]}$.

8 Lineare Operatoren in normierten und Banach–Räumen

8.1 Grundbegriffe, Fortsetzungen linearer Operatoren

Seien X und Y Vektorräume über $\mathbb{K}$ (zunächst ohne topologische Struktur). Eine lineare Abbildung T von einem Teilraum $D(T)$ von X in den Raum Y heißt ein *linearer Operator von X in* (oder *nach* Y; der Teilraum $D(T)$ heißt der *Definitionsbereich* von T. Die Menge

$$R(T) := \{Tx : x \in D(T)\}$$

ist offensichtlich ein Teilraum von Y; dieser wird als der *Wertebereich* von T bezeichnet.

Seien jetzt X und Y normierte Räume. Der lineare Operator T von X nach Y heißt *beschränkt*, wenn ein $C \geq 0$ existiert mit

$$\|Tx\| \leq C\|x\| \text{ für alle } x \in D(T)$$

(nach Korollar 2.10 und Satz 2.8 ist dies äquivalent mit *stetig* und *stetig im Nullpunkt*. Man definiert

$$\|T\| := \inf\{C : \|Tx\| \leq C\|x\| \text{ für alle } x \in D(T)\};$$

es gilt dann

$$\|Tx\| \leq \|T\|\|x\| \text{ für alle } x \in D(T).$$

Mit $B(X,Y)$ bezeichnet man die Menge der beschränkten Operatoren T von X nach Y mit $D(T) = X$. Nach Abschnitt 4.4 ist $B(X,Y)$ mit der oben angegebenen Norm stets ein normierter Raum; er ist vollständig (Banach–Raum), falls Y vollständig ist. Der Raum $B(X;X)$ wird kurz als $B(X)$ bezeichnet. Der topologische Dualraum $X^* = B(X, \mathbb{K})$ ist stets ein Banach–Raum.

Aus dem allgemeinen Fortsetzungssatz von Hahn–Banach (Satz 3.5) erhalten wir sofort den folgenden besonders wichtigen Spezialfall:

Satz 8.1 (Satz von Hahn–Banach für normierte Räume) *Sei X ein normierter Raum, F_0 ein stetiges lineares Funktional auf einem Teilraum $D(F_0)$ von X. Dann existiert ein stetiges lineares Funktional $F \in X^*$ mit $F(x) = F_0(x)$ für alle $x \in D(F_0)$ und $\|F\| = \|F_0\|$.*

Aus Aufgabe 3.6 folgt, daß F i.allg. nicht eindeutig bestimmt ist.

Ein entsprechendes Resultat für beliebige Operatoren (d. h. für Operatoren von X nach Y mit einem beliebigen normierten Raum Y) gilt nicht, vgl. N. Dunford – J. T. Schwartz, Linear Operators I, section VI. 12, Seite 554.

Eine Fortsetzung auf den Abschluß des Definitionsbereiches ist (sogar in eindeutiger Weise) immer möglich, wenn ein Y ein Banachraum ist.

Satz 8.2 *Sei X ein normierter Raum, Y ein Banachraum, T_0 ein beschränkter linearer Operator von X nach Y. Dann existiert* genau ein stetiger *linearer Operator T von X in Y mit $D(T) = \overline{D(T_0)}$. Es gilt $\|T\| = \|T_0\|$.*

Beweis. (Durch stetige Fortsetzung). *Eindeutigkeit*: Ist T ein solcher Operator, $x \in D(T) = \overline{D(T_0)}$ und (x_n) eine Folge aus $D(T_0)$ mit $x_n \to x$, so gilt wegen der Stetigkeit von T

$$Tx = \lim_n Tx_n = \lim_n T_0 x_n$$

d. h. T ist durch T_0 eindeutig bestimmt.

Existenz: Sei $x \in \overline{D(T_0)}$ und (x_n) eine Folge aus $D(T_0)$ mit $x_n \to x$. Da (x_n) eine Cauchyfolge ist, ist auch $(T_0 x_n)$ eine Cauchyfolge ($\|T_0 x_n - T_0 x_m\| \leq \|T_0\| \|x_n - x_m\|$). Da Y vollständig ist, ist $(T_0 x_n)$ konvergent. Der Limes hängt nicht von der Wahl der Folge (x_n) ab: Sind $(x_n), (y_n)$ zwei Folgen aus $D(T_0)$ mit $x_n \to x, y_n \to x$, so konvergiert auch die Folge $(z_n) = (x_1, y_1, x_2, y_2, \ldots)$ gegen x, d. h. auch $(T_0 z_n)$ ist konvergent; somit haben $(T_0 x_n)$ und $(T_0 y_n)$ den gleichen Limes.

Wir definieren

$$Tx := \lim_n T_0 x_n \text{ für } x \in \overline{D(T_0)},\ (x_n) \text{ aus } D(T_0),\ x_n \to x.$$

T ist linear: Aus $x_n \to x, y_n \to y$ folgt $ax_n + by_n \to ax + by$ und somit

$$T(ax + by) = \lim T_0(ax_n + by_n) = a \lim T_0 x_n + b \lim T_0 y_n = aTx + bTy.$$

Es gilt $\|T\| = \|T_0\|$; insbesondere ist T *stetig*: $\|T\| \geq \|T_0\|$ ist klar. Für $x \in D(T) = \overline{D(T_0)}$ sei (x_n) eine Folge aus $D(T_0)$ mit $x_n \to x$. Dann gilt

$$\|Tx\| = \lim \|T_0 x_n\| \leq \lim \|T_0\| \|x_n\| = \|T_0\| \|x\|,$$

also $\|T\| \leq \|T_0\|$. ∎

Eine Fortsetzung unter Erhaltung der Norm über den Abschluß des Definitionsbereichs hinaus ist i. allg. nicht möglich (vgl. obige Bemerkung). Ist X ein Hilbertraum (vgl. §9), so ist es leicht, die Existenz einer Fortsetzung auf den ganzen Raum mit gleicher Norm zu beweisen (vgl. Aufgabe 9.7); diese ist allerdings i. allg. nicht mehr eindeutig bestimmt.

8.2 Prinzip der gleichmäßigen Beschränktheit, starke Operatorenkonvergenz

Satz 8.3 (Prinzip der gleichmäßigen Beschränkheit) *Sei X ein Banachraum, Y ein normierter Raum, $\{T_\alpha : \alpha \in A\} \subset B(X;Y)$. Ist $\{T_\alpha : \alpha \in A\}$* punktweise beschränkt *(d. h. zu jedem $x \in X$ gibt es ein C_x mit $\|T_\alpha x\| \leq C_x$ für alle $\alpha \in A$), so ist $\{T_\alpha : \alpha \in A\}$ beschränkt (d. h. es gibt ein $C \geq 0$ mit $\|T_\alpha\| \leq C$ für alle $\alpha \in A$).*

Beweis. a) Es existieren $x_0 \in X, \varrho_0 > 0$ und $K_0 \geq 0$ mit

$$\|T_\alpha x\| \leq K_0 \text{ für alle } x \in K(x_0, \varrho_0),\ \alpha \in A.$$

Beweis von a: Sei für $k \in \mathbb{N}$

$$\begin{aligned} M_k &:= \{x \in X : \|T_\alpha x\| \leq k \text{ für alle } \alpha \in A\} \\ &= \bigcap_{\alpha \in A} \{x \in X : \|T_\alpha x\| \leq k\}. \end{aligned}$$

M_k ist abgeschlossen, da die T_α stetig sind (woraus die Abgeschlossenheit von $\{x \in X : \|T_\alpha x\| \leq k\}$ folgt). Nach Voraussetzung liegt jedes $x \in X$ in einem M_k, also gilt

$$X = \bigcup_{k \in \mathbb{N}} M_k.$$

Nach dem Satz von Baire (Satz 1.15) gibt es also ein $k_0 \in \mathbb{N}$ so, daß M_{k_0} eine Kugel $K(x_0, \varrho_0)$ mit $\varrho_0 > 0$ enthält. Mit $K_0 := k_0$ ist Teil a bewiesen.

b) Wir können nun die Behauptung des Satzes beweisen: Wegen Teil a gilt für alle $x \in K(0, \varrho_0)$ und $\alpha \in A$

$$\begin{aligned} \|T_\alpha x\| &= \|T_\alpha(x + x_0) - T_\alpha x_0\| \leq \|T_\alpha(x + x_0)\| + \|T_\alpha x_0\| \\ &\quad (\text{da } x + x_0 \in K(x_0, \varrho)) \\ &\leq K_0 + C_{x_0}. \end{aligned}$$

Hieraus folgt für alle $x \in K(0,1)$ und $\alpha \in A$

$$\|T_\alpha x\| \leq \frac{1}{\varrho_0}(K_0 + C_{x_0}) =: C,$$

d. h. es gilt

$$\|T_\alpha\| \leq C \text{ für alle } \alpha \in A.$$

■

Seien X, Y normierte Räume, $T_n (n \in \mathbb{N})$ und T aus $B(X, Y)$. Man sagt, die Folge (T_n) *konvergiert stark* (auch: *punktweise*) *gegen* T, wenn gilt

$$T_n x \to Tx \text{ (d. h. } \|T_n x - Tx\| \to 0) \text{ für alle } x \in X.$$

Zur Abkürzung schreiben wir $T_n \xrightarrow{s} T$. (Nur zur Verdeutlichung sei hier auf einen weiteren Konvergenzbegriff hingewiesen: Die Folge $(T_n x)$ *konvergiert schwach gegen* T, wenn für jedes $x \in X$ gilt $T_n x \xrightarrow{w} Tx$, d. h. $F(T_n x) \to F(Tx)$ für jedes $F \in Y^*$ und jedes $x \in X$.)

Satz 8.4 *(a) Seien X, Y normierte Räume, $T_n, T \in B(X, Y)$. Aus $T_n \xrightarrow{s} T$ folgt*

$$\|T\| \leq \liminf \|T_n\|.$$

(b) Seien X, Z normierte Räume, Y ein Banachraum, $S_n, S \in B(Y, Z)$, $T_n, T \in B(X, Y)$. Aus $S_n \xrightarrow{s} S, T_n \xrightarrow{s} T$ folgt $S_n T_n \xrightarrow{s} ST$.

(c) Seien X, Y Banachräume, $T_n \in B(X, Y)$. Dann sind folgende Aussagen äquivalent:

(i) *Es gibt ein* $T \in B(X,Y)$ *mit* $T_n \xrightarrow{s} T$.

(ii) *Es gibt ein* $C \geq 0$ *mit* $\|T_n\| \leq C$ *für alle* n, *und eine dichte Teilmenge* M *von* X *so, daß für jedes* $x \in M$ $(T_n x)$ *eine Cauchyfolge ist.*

Beweis. a) Sei (T_{n_k}) eine Teilfolge von (T_n) mit

$$\liminf_n \|T_n\| = \lim_k \|T_{n_k}\|.$$

Dann gilt für alle $x \in X$

$$\|Tx\| = \lim \|T_{n_k}x\| \leq \lim \|T_{n_k}\|\|x\| = \liminf \|T_n\|\|x\|.$$

b) Für jedes $y \in Y$ (Banachraum) ist $(S_n y)$ konvergent, also beschränkt; nach Satz 7.3 existiert also ein C mit $\|S_n\| \leq C$ für alle n. Hiermit folgt für alle $x \in X$

$$\begin{aligned} \|(S_nT_n - ST)x\| &\leq \|S_n(T_n - T)x\| + \|(S_n - S)Tx\| \\ &\leq C\|T_nx - Tx\| + \|(S_n - S)(Tx)\| \to 0 \quad \text{für} \quad n \to \infty. \end{aligned}$$

c) **(i)** $\Rightarrow$ **(ii)**: Nur die Beschränktheit der Folge (T_n) ist zu beweisen; vgl. hierzu Teil b; dabei wird benutzt, daß X ein Banachraum ist.

(ii) $\Rightarrow$ **(i)**: Sei $x \in X$. Zu jedem $\varepsilon > 0$ existiert ein $y \in M$ mit $\|x - y\| < \frac{\varepsilon}{3C}$. Zu diesem y gibt es nach Voraussetzung ein N mit

$$\|T_n y - T_m y\| \leq \frac{\varepsilon}{3} \text{ für } n, m \geq N.$$

Damit folgt

$$\begin{aligned} \|T_nx - T_mx\| &\leq \|T_n(x-y)\| + \|(T_n - T_m)y\| + \|T_m(y-x)\| \\ &\leq C\|x-y\| + \|T_ny - T_my\| + C\|x-y\| \\ &\leq 3\frac{\varepsilon}{3} = \varepsilon \text{ für } n, m \geq N. \end{aligned}$$

Also ist $(T_n x)$ eine Cauchyfolge in Y für jedes $x \in X$ und somit konvergent (da Y ein Banachraum ist). Wir definieren

$$Tx := \lim T_n x \text{ für alle } x \in X.$$

T ist offenbar linear und es gilt

$$\|Tx\| = \lim \|T_nx\| \leq C\|x\|, \text{ d. h. } T \in B(X,Y).$$

Nach Konstruktion von T gilt $T_n \xrightarrow{s} T$. ∎

8.3 Glättung in $L_p(\mathbb{R}^m)$, Integraloperatoren in $L_p(X,\mu)$

Es ist unser erstes Ziel zu zeigen, daß es eine Schar $\{J_\varepsilon : \varepsilon > 0\}$ von Operatoren in $B(L_p(\mathbb{R}^m))$ gibt, die jedem Element $f \in L_p(\mathbb{R}^m)$ „glatte“ $J_\varepsilon f$ aus $L_p(\mathbb{R}^m)$ zuordnen mit $\|J_\varepsilon f\| \leq \|f\|$ und $J_\varepsilon f \to f$ für $\varepsilon \to 0$.

Wir definieren (wie in Abschnitt 6.3) $\tilde{\delta}_\varepsilon : \mathbb{R}^m \to \mathbb{R}$ durch

$$\begin{aligned}
\tilde{\delta}_\varepsilon(x) &= \begin{cases} \exp\left\{\dfrac{1}{x^2-\varepsilon^2}\right\} & \text{für } |x| < \varepsilon, \\ 0 & \text{für } |x| \geq \varepsilon, \end{cases} \\
\delta_\varepsilon(x) &= \left\{\int \tilde{\delta}_\varepsilon(y)\,\mathrm{d}y\right\}^{-1} \tilde{\delta}_\varepsilon(x).
\end{aligned}$$

Dann gilt $\delta_\varepsilon \in C_0^\infty(\mathbb{R}^m)$ und $\displaystyle\int \delta_\varepsilon(x)\,\mathrm{d}x = 1$.

Für $f \in L_p(\mathbb{R}^m)$ definiert man

$$J_\varepsilon f(x) := \int \delta_\varepsilon(x-y) f(y)\,\mathrm{d}y \quad \text{für } x \in \mathbb{R}^m.$$

Da für jedes $x \in \mathbb{R}^m$ gilt $\delta_\varepsilon(x-\cdot)f(\cdot) \in L_1(\mathbb{R}^m)$, ist dieser Ausdruck wohldefiniert; es ist leicht zu sehen, daß $J_\varepsilon f$ beliebig oft stetig differenzierbar („glatt“) ist.

Satz 8.5 *(a) Für $1 \leq p \leq \infty$ ist J_ε für jedes $\varepsilon > 0$ ein Operator aus $B(L_p(\mathbb{R}^m))$ mit $\|J_\varepsilon\| \leq 1$.*

(b) Für $1 \leq p < \infty$ gilt $J_\varepsilon \xrightarrow{s} I$ (Identität) für $\varepsilon \to 0$. (Dies ist der Grund, weshalb man $\{\delta_\varepsilon : \varepsilon > 0\}$ auch als Deltaschar *bezeichnet.)*

Beweis. a) $\mathbf{1 < p < \infty}$: Da $J_\varepsilon f$ stetig ist, ist es auf jeden Fall Lebesgue–meßbar. Es genügt zu zeigen, daß $|J_\varepsilon f|^p$ integrierbar ist: Sei $\dfrac{1}{p} + \dfrac{1}{q} = 1$; dann folgt

$$\begin{aligned}
\int |J_\varepsilon f(x)|^p\,\mathrm{d}x &= \int \left| \int \delta_\varepsilon(x-y) f(y)\,\mathrm{d}y \right|^p \mathrm{d}x \\
&= \int \left| \int \delta_\varepsilon(x-y)^{1/q} \left\{\delta_\varepsilon(x-y)^{1/p} f(y)\right\} \mathrm{d}y \right|^p \mathrm{d}x \\
&\qquad \text{(Hölder'sche Ungleichung)} \\
&\leq \int \left\{ \int \delta_\varepsilon(x-y)\,\mathrm{d}y \right\}^{p/q} \left\{ \int \delta_\varepsilon(x-y)|f(y)|^p\,\mathrm{d}y \right\}^{p/p} \mathrm{d}x \\
&= \int\int \delta_\varepsilon(x-y)|f(y)|^p\,\mathrm{d}y\,\mathrm{d}x \overset{\text{(Fubini)}}{=} \int\int \delta_\varepsilon(x-y)|f(y)|^p\,\mathrm{d}x\,\mathrm{d}y \\
&= \int |f(y)|^p\,\mathrm{d}y = \|f\|_q^p.
\end{aligned}$$

Hieraus ergibt sich gleichzeitig $\|J_\varepsilon\| \le 1$ für alle $\varepsilon > 0$.

$\mathbf{p} = \infty$: Wegen $\displaystyle\int \delta_\varepsilon(x)\,\mathrm{d}x = 1$ folgt für alle $x \in \mathbb{R}^m$

$$|J_\varepsilon f(x)| \le \int \delta_\varepsilon(x-y)|f(y)|\,\mathrm{d}y \le \|f\|_\infty \int \delta_\varepsilon(x-y)\,\mathrm{d}y = \|f\|_\infty,$$

also $J_\varepsilon f \in L_\infty(\mathbb{R}^m)$ und $\|J_\varepsilon\| \le 1$.

$\mathbf{p} = \mathbf{1}$: Für alle $f \in L_1(\mathbb{R}^m)$ gilt wegen $\displaystyle\int \delta_\varepsilon(x)\,\mathrm{d}x = 1$

$$\begin{aligned}\int |J_\varepsilon f(x)|\,\mathrm{d}x &= \int \Big| \int \delta_\varepsilon(x-y) f(y)\,\mathrm{d}y \Big|\,\mathrm{d}x \\ &\le \int\int \delta_\varepsilon(x-y)|f(y)|\,\mathrm{d}y\,\mathrm{d}x \overset{\text{(Fubini)}}{=} \int\int \delta(x-y)|f(y)|\,\mathrm{d}x\,\mathrm{d}y \\ &= \int |f(y)|\,\mathrm{d}y = \|f\|_1,\end{aligned}$$

also $J_\varepsilon f \in L_1$ und $\|J_\varepsilon\| \le 1$.

b) Für Treppenfunktionen $f \in E(\mathbb{R}^m, \mathcal{R})$ ist leicht zu zeigen, daß gilt $J_\varepsilon f \to f$ für $\varepsilon \to 0$ im Sinne von $L_p(\mathbb{R}^m)$, $1 \le p < \infty$ (vgl. Beweis von Satz 7.6). (Entsprechend könnte z. B. $C_0^\infty(\mathbb{R}^m)$ statt $E(\mathbb{R}^m), \mathcal{R})$ benutzt werden.) Mit Satz 8.4 folgt die Behauptung. ■

Diesen Prozess, eine L_p–Funktion durch glatte Funktionen zu approximieren, nennt man kurz *Glättung*. Er wird insbesondere in der Theorie der partiellen Differentialgleichungen häufig benutzt. Es sei noch bemerkt, daß die spezielle Form der Funktion δ_ε keine wesentliche Rolle spielt.

Die im Beweis von Teil a) des Satzes 7.5 benutzte Methode zum Beweis der Beschränktheit von J_ε wollen wir nun noch verfeinern, um ein sehr allgemeines Kriterium für die Beschränktheit von Integraloperatoren von $L_p(Y,\nu)$ nach $L_p(X,\mu)$ zu beweisen (dies ist eine Verallgemeinerung einer Abschätzung der Norm einer Matrix von I. Schur.)

Satz 8.6 *Seien (X,μ), (Y,ν) σ–endliche Maßräume, $1 < p < \infty$, $\dfrac{1}{p} + \dfrac{1}{q} = 1$, die* Kerne *$k, k_1, k_2 : X \times Y \to \mathbb{C}$ seien $(\mu \times \nu)$–meßbar, $k(x,y) = k_1(x,y)k_2(x,y)$ und*

$$\int |k_1(x,y)|^q\,\mathrm{d}\nu(y) \le C_1^q \qquad \text{für } \mu\text{–f. a.} x \in X,$$
$$\int |k_2(x,y)|^p\,\mathrm{d}\mu(x) \le C_2^p \qquad \text{für } \mu\text{–f. a.} y \in Y.$$

Dann wird durch

$$Kf(x) = \int k(x,y)f(y)\,\mathrm{d}\nu(y) \text{ für } \mu\text{–f,ä.} x \in X$$

ein Operator $K \in B(L_p(Y,\nu), L_p(X,\mu))$ definiert mit $\|K\| \le C_1C_2$.

Beweis. a) Wir zeigen zunächst, daß $Kf(x)$ für μ–fast alle $x \in X$ definiert und μ–meßbar ist: Sei $X = \cup X_n$ mit μ–meßbaren Teilmengen X_n, $\mu(X_n) < \infty$. Nach dem Satz von Fubini–Tonelli (Satz 5.18) gilt

$$\begin{aligned}
\int\limits_{X_n \times Y} |k(x,y)f(y)|\,\mathrm{d}(\mu\times\mu)(x,y) &= \int\limits_{X_n}\Big\{\int\limits_{Y} |k(x,y)f(y)|\,\mathrm{d}\nu(y)\Big\}\,\mathrm{d}\mu(x) \\
&\quad (a \le 1 + a^p \text{ für } a \ge 0,\ p \ge 1) \\
&\le \int\limits_{X_n}\Big\{1 + \Big[\int\limits_{Y} |k(x,y)f(y)|\,\mathrm{d}\nu(y)\Big]^p\Big\}\,\mathrm{d}\mu(x) \\
&= \mu(X_n) + \int\limits_{X_n}\Big\{\int\limits_{Y} |k_1(x,y)||k_2(x,y)||f(y)|\,\mathrm{d}\nu(y)\Big\}^p\,\mathrm{d}\mu(x) \\
&\le \mu(X_n) + \int\limits_{X_n}\Big\{\int\limits_{Y} |k_1(x,y)|^q\,\mathrm{d}\nu(y)\Big\}^{p/q}\int\limits_{Y} |k_2(x,y)|^p|f(y)|^p\,\mathrm{d}\nu(y)\,\mathrm{d}\mu(x) \\
&\le \mu(X_n) + C_1^p\int\limits_{Y} |f(y)|^p \int\limits_{X} |k_2(x,y)|^p\,\mathrm{d}\mu(x)\,\mathrm{d}\nu(y) \\
&\le \mu(X_n) + C_1^pC_2^p\|f\|_p^p < \infty \text{ für alle } n.
\end{aligned}$$

Nach dem Satz von Fubini–*Tonelli* (Satz 5.18) existiert also $\int_Y k(x,y)f(y)\,\mathrm{d}\nu(y)$ für μ–fast alle $x \in X_n$ und stellt dort eine μ–meßbare Funktion dar. Da dies für alle n gilt, folgt die Behauptung a).

b) Es bleibt zu zeigen, daß $Kf \in L_p(X,\mu)$ und $\|Kf\|_p \le C_1C_2\|f\|_p$ für jedes

$f \in L_p(Y, \nu)$ gilt: Dies folgt aus

$$
\begin{aligned}
\int_X |Kf(x)|^p \,\mathrm{d}\mu(x) &= \int_X \Big| \int_Y k(x,y) f(y) \,\mathrm{d}\nu(y) \Big|^p \,\mathrm{d}\mu(x) \\
&\le \int_X \Big\{ \int_Y |k_1(x,y)||k_2(x,y)||f(y)| \,\mathrm{d}\nu(y) \Big\}^p \,\mathrm{d}\mu(x) \\
&\le \int_X \Big\{ \int_Y |k_1(x,y)|^q \,\mathrm{d}\nu(y) \Big\}^{p/q} \int_Y |k_2(x,y)|^p |f(y)|^p \,\mathrm{d}\nu(y) \,\mathrm{d}\mu(x) \\
&\le C_1^p \int_X \int_Y |k_2(x,y)|^p |f(y)|^p \,\mathrm{d}\nu(y) \,\mathrm{d}\mu(x) \\
&\le C_1^p \int_Y \int_X |k_2(x,y)|^p |f(y)|^p \,\mathrm{d}\mu(x) \,\mathrm{d}\nu(y) \\
&\le C_1^p C_2^p \|f\|_p^p.
\end{aligned}
$$

■

Für die Fälle $p = 1$ und $p = \infty$ sei auf Aufgabe 7.4 verwiesen. Kriterien für die Beschränktheit von Integraloperatoren von L_p nach L_q mit $p \neq q$ werden z. B. in K. Jörgens, Lineare Integraloperatoren, Abschnitt 11.2 - 4, behandelt.

8.4 Der Satz von der offenen Abbildung

Die Resultate dieses und des folgenden Abschnitts sind von fundamentaler Bedeutung für fast alle tiefergehenden Untersuchungen linearer (insbesondere auch unbeschränkter) Operatoren in Banachräumen.

Satz 8.7 (Satz von der offenen Abbildung) *Seien X und Y Banachräume, $T \in B(X, Y)$ sei surjektiv. Dann ist T* offen, *d. h. für jede offene Menge $U \subset X$ ist auch TU offen in Y.*

Beweis. In diesem Beweis bezeichnen wir mit X_ϱ und Y_ϱ die offenen Kugeln um 0 mit Radius ϱ in X bzw. Y. Der Beweis besteht aus drei Schritten:

(i) Für jedes $\varrho > 0$ existiert ein $\tau > 0$ mit $\overline{TX_\varrho} \supset Y_\tau$:
Sei $\delta := \varrho/2$; dann gilt natürlich $X = \bigcup_n nX_\delta = \bigcup_n X_{n\delta}$. Daraus folgt wegen der

Surjektivität von T

$$Y = TX = \bigcup_n nTX_\delta = \bigcup_n \overline{nTX_\delta}.$$

Nach dem Satz von Baire (Satz 1.15) enthält mindestens eine der Mengen $\overline{nTX_\delta} = n\overline{TX_\delta}$ eine offene Kugel; also enthält $\overline{TX_\delta}$ eine offene Kugel K.

Aus $X_\varrho \supset X_\delta - X_\delta$ folgt $TX_\varrho \supset TX_\delta - TX_\delta$ und somit

$$\overline{TX_\varrho} \supset \overline{TX_\delta - TX_\delta} \supset \overline{TX_\delta} - \overline{TX_\delta} \supset K - K = \bigcup_{x \in K}(x - K).$$

Die zuletzt stehende Menge ist offen (da K offen ist) und enthält 0. Also existiert eine Kugel um 0, die in dieser Menge, und somit in $\overline{TX_\varrho}$ enthalten ist.

(ii) Für jedes $\varrho > 0$ enthält TX_ϱ (jetzt ohne Abschluß) eine offene Kugel um Null:
Sei $r_0 := \varrho/2$; $r_n > 0$ seien so gewählt, daß $\sum_{n=1}^\infty r_n < r_0$ gilt. Nach (i) existieren für alle $n \in \mathbb{N}_0$ positive τ_n mit $Y_{\tau_n} \subset \overline{TX_{r_n}}$; dabei können wir ohne Einschränkung annehmen, daß $\tau_n \to 0$ gilt.

Wir zeigen $Y_{\tau_0} \subset TX_\varrho$ (womit (ii) bewiesen ist):
Sei $y \in Y_{\tau_0}$. Dann ist $y \in \overline{TX_{r_0}}$ und somit existiert ein $x_0 \in X_{r_0}$ mit

$$\|y - Tx_0\| < \tau_1, \text{ d. h. } y - Tx_0 \in Y_{\tau_1}.$$

Dann ist aber $y - Tx_0 \in \overline{TX_{r_1}}$ und somit existiert ein $x_1 \in X_{r_1}$ mit

$$\|y - Tx_0 - Tx_1\| < \tau_2, \text{ d. h. } y - Tx_0 - Tx_1 \in Y_{\tau_2},$$

usw. ...

Dieser Prozess liefert eine Folge (x_n) aus X mit

$$x_n \in X_{r_n} \text{ und } \|y - T(x_0 + \ldots + x_n)\| < \tau_{n+1}.$$

Wegen $\|x_n\| < r_n$ ist $\sum x_n$ konvergent; sei $x := \sum x_n$. Dann gilt

$$\|x\| \leq \sum_{N=0}^\infty \|x_n\| \leq r_0 + \sum_{n=1}^\infty r_n < 2r_0 = \varrho, \text{ d. h. } x \in X_\varrho$$

und

$$\|y - Tx\| = \lim_n \|y - T(x_0 + \ldots + x_n)\| \leq \lim \tau_n = 0,$$

d. h. es gilt

$$y = Tx \in TX_\varrho.$$

(iii) Wir zeigen jetzt, daß T offen ist: Sei $M \subset X$ offen, $x \in M$, $y := Tx$. Dann existiert eine offene Kugel K um 0 in X mit $x + K \subset M$, und (2. Schritt) eine offene Kugel V um 0 in Y mit $V \subset TK$. Daraus folgt

$$y + V = Tx + V \subset Tx + TK = T(x + K) \subset TM,$$

d. h. um jedes $y \in TM$ existiert eine offene Kugel, die ganz in TM liegt, TM ist somit offen. □

Satz 8.8 (Satz von der stetigen Inversen) *Seien X und Y Banachräume, $T \in B(X,Y)$ sei bijektiv. Dann ist T^{-1} stetig.*

Beweis. Nach Satz 7.7 ist $(T^{-1})^{-1} = T$ offen, d. h. für jede offene Menge $M \subset X$ ist $(T^{-1})^{-1}M$ (das Urbild von M unter T^{-1}) offen. Das ist gerade die Definition der Stetigkeit von T^{-1} (vgl. Abschnitt 1.3). ■

8.5 Abgeschlossene Operatoren, der Satz vom abgeschlossenen Graphen

Seien X und Y normierte Räume. Dann ist auch $X \times Y$ mit

$$\|(x, y)\| = \|x\| + \|y\|$$

ein normierter Raum; er ist vollständig genau dann, wenn X *und* Y vollständig sind (vgl. Satz 4.15; ebenso könnten die anderen in Abschnitt 4.5 angegebenen Normen auf $X \times Y$ betrachtet werden, wodurch sich nichts wesentliches ändern würde, da sie alle die Produkttopologie auf $X \times Y$ erzeugen).

Sei T ein linearer Operator von X nach Y. An Anlehnung an den Begriff des „Graphen“ einer reellen Funktion wird

$$G(T) := \{(x, Tx) \in X \times Y : x \in D(T)\}$$

der *Graph* des Operators T genannt. Offensichtlich ist $G(T)$ ein Teilraum von $X \times Y$ (vgl. auch Aufgabe 7.6).

Der Operator heißt *abgeschlossen*, wenn $G(T)$ als Teilraum von $(X \times Y, \|\cdot\|)$ abgeschlossen ist. Offensichtlich gilt:

Bemerkung 8.9 Ist T ein injektiver Operator von X nach Y, so gilt: T ist genau dann abgeschlossen, wenn T^{-1} abgeschlossen ist.

Satz 8.10 *Die folgenden Aussagen sind äquivalent:*

(i) *T ist abgeschlossen.*

(ii) *Ist (x_n) eine Folge aus $D(T)$ mit $x_n \to x \in X$ und $Tx_n \to y \in Y$, so gilt $x \in D(T)$ und $y = Tx$.*

Sind X und Y Banachräume, so sind diese Aussagen äquivalent zu

(iii) *$(D(T), \|\cdot\|_T)$ mit $\|x\|_T = \|x\| + \|Tx\|$ ist vollständig.*

Beweis. **(i) $\Longleftrightarrow$ (ii)**: ist klar, da (ii) nur die Charakterisierung der Abgeschlossenheit von $G(T)$ mit Hilfe von Folgen ist (vgl. Satz 1.10).

(i) $\Longleftrightarrow$ (iii): Wegen $\|x\|_T = \|(x, Tx)\|$ (wobei $\|\cdot\|$ die obige Norm auf $X \times Y$ ist), bedeutet (iii) die Vollständigkeit von $G(T)$ bezüglich $\|\cdot\|$, also nach Satz 1.16 b die Abgeschlossenheit als Teilraum von $X \times Y$. ■

Aus Satz 7.10 ((i) $\Longleftrightarrow$ (ii)) ergibt sich sofort die

Folgerung: Jedes $T \in B(X, Y)$ ist abgeschlossen.

Satz 8.11 (Satz vom abgeschlossenen Graphen) *Seien X und Y Banachräume, T ein linearer Operator von X in Y mit $D(T) = X$. Ist T abgeschlossen, so ist T beschränkt. (Die gleiche Aussage gilt, wenn statt $D(T) = X$ nur vorausgesetzt wird, daß $D(T)$ abgeschlossen ist; im Beweis ersetzt man dann X durch den Banachraum $D(T)$.)*

Beweis. $G(T)$ ist ein abgeschlossener Teilraum von $X \times Y$. Die Abbildung

$$P : G(T) \to X, \quad P(x, Tx) = x$$

ist offenbar
linear: Sie ist eine Einschränkung der Projektion von $X \times Y$ auf die erste Komponente,
stetig: $\|P(x,Tx)\| = \|x\| \leq \|x\| + \|Tx\| = \|(x,Tx)\|$,
bijektiv: $P(x,Tx) = 0 \Rightarrow x = 0 \Rightarrow (x,Tx) = (0,0)$ (injektiv),
$x = P(x,Tx)$ für jedes $x \in X$ (surjektiv).

Nach dem Satz von der stetigen Inversen (Satz 7.8) ist also P^{-1} stetig und somit auch

$$T = QP^{-1} \qquad (x \mapsto (x,Tx) \mapsto Tx),$$

wobei Q die ebenfalls stetige Projektion von $X \times Y$ auf die zweite Komponente ist. ■

Satz 8.12 *Seien X, Y normierte Räume, T ein linearer Operator von X nach Y.*

(a) Ist T beschränkt und $D(T)$ abgeschlossen, so ist T abgeschlossen.

(b) Sei Y ein Banachraum. Ist T beschränkt und abgeschlossen, so ist $D(T)$ abgeschlossen.

Beweis. Sei (x_n) aus $D(T)$ mit $x_n \to x \in X$ und $Tx_n \to y \in Y$. Da $D(T)$ abgeschlossen ist, gilt $x \in D(T)$. Da T stetig ist, gilt $Tx = \lim Tx_n = y$. Also gilt $x \in D(T),\ y = Tx$.

b) Sei $x \in \overline{D(T)}$. Dann existiert eine Folge x_n aus $D(T)$ mit $x_n \to x$. Da T beschränkt ist, ist (Tx_n) eine Cauchyfolge und somit konvergent $Tx_n \to y \in Y$ (da Y ein Banachraum ist). Da T abgeschlossen ist, gilt $x \in D(T)$. ■

Als wichtige Klasse von abgeschlossenen Operatoren betrachten wir noch Multiplikationsoperatoren in $L_p(X, \mu)$. Sei μ ein Maß auf X, $t : X \to \mathbb{C}$ sei μ–meßbar.

Der durch t erzeugte *Maximale Multiplikationsoperator* M_t in $L_p(X,\mu)$ ist definiert durch

$$\begin{aligned} D(M_t) &:= \{f \in L_p(X,\mu) : tf \in L_p(X,\mu)\}, \\ M_t f &:= tf \text{ für } f \in D(M_t). \end{aligned}$$

(Eigentlich sollte M_t auch mit einem Index p versehen werden, worauf wir verzichten wollen.)

Satz 8.13 *(a) M_t ist abgeschlossen für alle $p \in [1,\infty]$.*

(b) $D(M_t)$ ist dicht in $L_p(X,\mu)$ für $1 \le p \le \infty$.

(c) Für $p = \infty$ gilt: Entweder ist $D(M_t) = L_\infty(X,\mu)$ [wenn t wesentlich beschränkt ist], oder $D(M_t)$ ist nicht dicht *in $L_\infty(X;\mu)$ [wenn t nicht wesentlich beschränkt ist].*

Beweis. a) Gilt $f_n \to f$ und $M_t f_n = tf_n \to g$ in $L_p(X,\mu)$, so existiert (Satz 6.3) eine Teilfolge (f_{n_k}) mit

$$f_{n_k}(x) \to f(x), \quad t(x)f_{n_k}(x) \to t(x)f(x) \quad \mu\text{–f. ü.}$$

Daraus folgt

$$g(x) = t(x)f(x) \quad \mu\text{–f. ü.},$$

also ist $tf \in L_p(X,\mu)$, d. h.

$$f \in D(M_t) \text{ und } g = tf = M_t f.$$

b) Für $n \in \mathbb{N}$ sei $X_n := \{x \in X : |t(x)| \le n\}$; es gilt also $X = \cup X_n$. Für $f \in L_p(X,\mu)$ sei

$$f_n(x) := \begin{cases} f(x) & \text{für } x \in X_n, \\ 0 & \text{sonst.} \end{cases}$$

Dann gilt $f_n \in D(M_t)$ und $f_n \to f$.

c) (i) Ist t wesentlich beschränkt, so ist $tf \in L_\infty(X,\mu)$ für jedes $f \in L_\infty(X,\mu)$, also $D(M_t) = L_\infty(X,\mu)$.

(ii) Sei t nicht wesentlich beschränkt. Wir zeigen, daß die konstante Funktion 1 nicht in $\overline{D(M_t)}$ liegt (d. h. $D(M_t)$ ist nicht dicht). Ist (f_n) eine Folge aus $L_\infty(X,\mu)$ mit $\|f_n - 1\|_\infty \to 0$, so existiert ein $n_0 \in \mathbb{N}$ und eine μ–Nullmenge $N \subset X$ mit

$$|f_n(x)| \geq \frac{1}{2} \text{ für alle } x \in X \backslash N \text{ und } n \geq n_0.$$

Da t nicht wesentlich beschränkt ist, ist dann tF_n nicht in $L_\infty(X,\mu)$ für $n \geq n_0$, also ist die Folge (f_n) nicht in $D(M_t)$. ■

8.6 Übungsaufgaben

8.1 Sei X ein normierter Raum. Zwei Teilräume X_1 und X_2 von X heißen *topologisch komplementär*, wenn gilt:

$$\begin{array}{lll} X_1 \cap X_2 = \{0\}, & X = X_1 + X_2, & (\textit{algebraisch komplementär}) \\ P_j : X \to X_j, & P_j(x_1 + x_2) = x_j & \text{sind stetig für } 1 = 1,2. \end{array}$$

(a) Seien X_1, X_2 topologisch komplementär. Man zeige: $P_j^2 = P_j$ für $j = 1,2$ (d. h. P_1 und P_2 sind Projektoren), $R(P_1) = N(P_2), R(P_2) = N(P_1), X_1$ und X_2 sind abgeschlossen.

(b) Ist X_1 ein Teilraum von X und P ein stetiger Projektor mit $R(P) = X_1$, so ist $X_2 := (I - P)X$ ein topologischer Komplementärraum zu X_1.

(c) Ist X_1 ein endlich dimensionaler Teilraum von X, so gibt es einen topologischen Komplementärraum zu X_1.
Anleitung: Sei $\{e_1, \ldots, e_n\}$ eine Basis von X_1, $F_j \in X^*$ mit $F_j(e_k) = \delta_{jk}$, $P_1 x := \sum\limits_j F_j(x) e_j$.

8.2 Man erkläre eine Topologie $\mathcal{T}$ (Nullumgebungsbasis) auf $B(X,Y)$ so, daß die Konvergenz bezüglich $\mathcal{T}$ gleich der starken Konvergenz ist.

8.3 Mit Hilfe von Beispielen in $L_p(\mathbb{R})$ zeige man

$$T_n \xrightarrow{w} T \not\Rightarrow T_n T_n \xrightarrow{s} T \not\Rightarrow T_n \to T.$$

$$\textit{Vorschläge} : T_n f(x) = f(x-n),\ \ S_n f(x) = \begin{cases} 0 & \text{für } |x| \leq n, \\ f(x) & \text{für } |x| > n. \end{cases}$$

8.4 Seien (X, μ) und (Y, ν) σ–endliche Maßräume, $k : X \times Y \to \mathbb{C}$ sei $(\mu \times \nu)$–meßbar.

(a) Ist $\displaystyle\int |k(x,y)|\, \mathrm{d}\mu(x) \leq C$ für ν–fast alle $y \in Y$, so wird durch

$$(*) \qquad Kf(x) = \int k(x,y) f(y)\, \mathrm{d}\nu(y)$$

ein Operator $K \in B(L_1(Y,\nu), L_1(X,\mu))$ erklärt mit $\|K\| \leq C$.

(b) Ist $\displaystyle\int |k(x,y)|\, \mathrm{d}\,\nu(y) \leq C$ für μ–fast alle $x \in X$, so wird durch $(*)$ ein Operator $K \in B(L_\infty(Y,\nu))$ erklärt mit $\|K\| \leq C$.

8.5 Sei $f : \mathbb{R} \to \mathbb{C}$ absolut stetig, d. h. $f(x) = f(0) + \displaystyle\int_0^x g(t)\, \mathrm{d}t$ mit einer lokal integrierbaren Funktion g, $g_h(x) := \dfrac{1}{h}(f(x+h) - f(x))$ für $h \neq 0$. Dann gilt

$$\|g_h - g\|_1 \to 0 \text{ für } h \to 0.$$

Anleitung: Man vergleiche die bei der „Glättung“ benutzte Methode.
Anmerkung: Es gilt sogar $g_h(x) \to g(x)$ für μ–f. a. x, d. h. f ist f. ü. differenzierbar (vgl. W. Rudin: Real and complex analysis, Theorem 8.17).

8.6 Seien X und Y Vektorräume. Ein Teilraum G von $X \times Y$ ist genau dann Graph eines Operators T von X nach Y, wenn aus $(0, y) \in G$ folgt $y = 0$ (oder äquivalent dazu: wenn aus $(x, y), (x, y') \in G$ folgt $y = y'$).

8.7 In jedem unendlich–dimensionalen normierten Raum X gibt es Operatoren T mit

$$D(T) = X; \quad \textit{nicht} \text{ abgeschlossen (also } \textit{nicht} \text{ stetig)}$$

Anleitung: Aus einer Basis $\{x_\alpha : \alpha \in A\}$ wähle man eine Folge (x_{α_n}) mit $\alpha_n \neq \alpha_m$ für $n \neq m$ aus und definiere

$$T\Big(\sum c_\alpha x_\alpha\Big) = \sum n c_{\alpha_n} x_{\alpha_n}.$$

8.8 Sei μ ein Maß auf X, $k : X \times X \to \mathbb{C}$ $\mu \times \mu$–meßbar. Für jedes $f \in L_p(X, \mu)$ sei $k(x, \cdot) f(\cdot) \in L_1(X, \mu)$ für μ–f. a. $x \in X$, und mit

$$Kf(x) = \begin{cases} \displaystyle\int k(x,y) f(y)\, \mathrm{d}\mu(y) & \text{falls } k(x,\cdot)f(\cdot) \in L_1, \\ 0 & \text{sonst} \end{cases}$$

sei $Kf \in L_q$. Dann ist $K \in B(L_p(x,\mu), L_q(X,\mu))$.
Anleitung: Man zeige die Abgeschlossenheit von K; dabei benutze man die bei der Vollständigkeit von L_p mitbewiesene Tatsache, daß zu jeder konvergenten Folge (f_n) aus L_p eine Teilfolge (f_{n_k}) und ein $g \in L_p$ existieren mit $|f_{n_k}(x)| \leq g(x)$ μ–f. ü. für alle $k \in \mathbb{N}$.

8.9 Sei μ ein Maß auf X, $1 \leq p, q \leq \infty$, $T : X \to \mathbb{C}$ sei μ–meßbar,

$$\begin{aligned} D(M_t) &= \Big\{ f \in L_p(X,\mu) : tf \in L_q(X,\mu) \Big\}, \\ M_t f &= tf \text{ für } f \in D(M_t); \end{aligned}$$

M_t ist der durch t erzeugte *maximale Multiplikationsoperator* von $L_p(X,\mu)$ nach $L_q(X,\mu)$.

(a) M_t ist abgeschlossen.

(b) Ist $q \leq q$, so ist M_t genau dann aus $B(L_p(X,\mu), L_q(X,\mu))$, wenn $t \in L_r(X,\mu)$ gilt mit $\dfrac{1}{p} + \dfrac{1}{r} = \dfrac{1}{q}$; es gilt dann $\|M_t\| = \|t\|_r$. (Man verwende die verallgemeinerte Hölder'sche Ungleichung, Aufgabe 6.2)

(c) Ist $q > p$ und $\mu(\{x\}) = 0$ für jedes $x \in X$, so ist

$$\begin{aligned} D(M_t) &= \{ f \in L_p(X,\mu) : tf = 0 \ \mu\text{–f. ü.} \}, \\ M_t f &= 0 \text{ für alle } f \in D(M_t). \end{aligned}$$

8.10 Sei μ ein Maß auf X, $f : X \to \mathbb{C}$ sei μ–meßbar. Ist $fg \in L_1(X,\mu)$ für jedes $g \in L_q(X,\mu)$, so ist $f \in L_p(X,\mu)$ mit $\dfrac{1}{p} + \dfrac{1}{q} = 1$.
Anleitung: Man zeige, daß das Funktional $F : L_q(X,\mu) \to \mathbb{C}$, F(g)=$\displaystyle\int fg \, \mathrm{d}\mu$ abgeschlossen ist.

8.11 Sei (X,μ) ein σ–endlicher Maßraum, $t_n : X \to \mathbb{C}$ μ–meßbar, M_{t_n} der maximale durch t_n erzeugte Multiplikationsoperator in $L_p(X,\mu)$.

(a) Gilt $M_{t_n} \xrightarrow{s} T$, so gibt es eine μ–meßbare Funktion $t : X \to \mathbb{C}$ mit $T = M_t$.

(b) Gilt $\|t_n\|_\infty \leq C > \infty$ für alle $n \in \mathbb{N}$ und $t_n(x) \to t(x)$ μ–f. ü., so folgt $M_{t_n} \xrightarrow{s} M_t$.

8.12 Sei $k \in L_1(\mathbb{R})$.

(a) Durch $Kf(x) = \int\limits_{\mathbb{R}} k(x-y)f(y)\,\mathrm{d}y$ wird ein Operator $K \in B(L_2(\mathbb{R}))$ definiert mit $\|K\| \leq \|k\|_1$.

(b) Ist $k \geq 0$ für f. a. x, so gilt $\|K\| = \|k\|_1$.
Anleitung: Man betrachte zunächst k mit kompaktem Träger.

9 Grundlagen der Spektraltheorie

9.1 Begriffsbildung, Resolventengleichungen

Die folgenden Begriffsbildungen sind motiviert durch die Untersuchung der Lösbarkeitseigenschaften von Gleichungen der Form

$$(T - zI)x = y, \tag{1}$$

wobei T ein linearer Operator in einem normierten oder Banachraum X ist (wobei $D(T)$ nicht notwendig der ganze Raum X ist), $z \in \mathbb{K}$, $y \in X$. Gesucht ist die Lösung $x \in X$, d. h. formal $x = (T - zI)^{-1}y$. Statt $T - zI$ werden wir im folgenden meist nur $T - z$ schreiben.

Wünschenswert wären die folgenden Eigenschaften:

(i) Die Gleichung (1) ist für jedes $y \in X$ lösbar; d. h. $T - z$ ist *surjektiv*,

(ii) die Lösung x von (1) ist eindeutig bestimmt, d. h. $T - z$ ist *injektiv* (z ist *nicht* Eigenwert),

(iii) die Lösung x von (1) hängt stetig von y ab, d. h. $(T - z)^{-1}$ ist *beschränkt* (stetig),

(iv) in dem Bereich der $z \in \mathbb{K}$, für die (i), (ii), und (iii) gilt, hängt die Lösung x von (1) stetig von z ab.

Die Menge $\varrho(T)$ der $z \in \mathbb{K}$, für die (i), (ii) und (iii) gelten, heißt *Resolventenmenge* von T, also

$$\varrho(T) := \{z \in \mathbb{K} : \ T - z \text{ [als Abbildung von } D(T) \text{ nach } X] \text{ ist bijektiv,} \\ (T - z)^{-1} \text{ ist stetig}\}.$$

Insbesondere ist also $(T - z)^{-1}$ (und somit auch $T - z$) abeschlossen für jedes $z \in \varrho(T)$.

Bemerkung 9.1 Ist T *nicht abgeschlossen*, so gilt $\varrho(T) = \emptyset$, denn es gilt dann für jedes $z \in \mathbb{K} : T - z$ ist nicht abgeschlossen. Also kann z nicht in $\varrho(T)$ sein.

Wir setzen deshalb im folgenden in der Regel voraus, daß T abgeschlossen ist.

Ist T ein *abgeschlossener Operator in einem Banachraum* X, so gilt:

$$\varrho(T) = \Big\{z \in \mathbb{K}, T - z \text{ ist bijektiv [als Abbildung von } D(T) \text{ nach } X\,]\Big\},$$

da die Stetigkeit von $(T-z)^{-1}$ mit Hilfe der Abgeschlossenheit von $(T-z)^{-1}$ automatisch folgt (Satz vom abgeschlossenen Graphen, Satz 7.12). Wir haben hier zweimal benutzt, daß mit T auch $T-z$ abgeschlossen ist; dies ist sehr leicht nachzuprüfen, folgt aber auch als Spezialfall aus Satz 8.4.

Die Funktion

$$\varrho(T) \to B(X), \quad z \mapsto R(T,z) := (T-z)^{-1}$$

heißt die *Resolvente* (oder Resolventenfunktion) von T, $R(T,z)$ heißt die *Resolvente* von T *im Punkt* z.

Nach Definition ist die Resolventenmenge von T gleich der Menge der $z \in \mathbb{K}$, für die Gleichung (1) die Eigenschaften (i), (ii) und (iii) erfüllt. Wir werden in Abschnitt 8.3 sehen, daß dann die Eigenschaft (iv) automatisch erfüllt ist, d. h. daß $R(T,\cdot)$ in $\varrho(T)$ stetig ist.

Das *Spektrum* $\varrho(T)$ von T ist definiert als das Komplement der Resolventenmenge

$$\sigma(T) := \mathbb{K} \setminus \varrho(T) = \complement\, \varrho(T).$$

Für das folgende müssen wir erklären, wie nicht notwendig überall definierte Operatoren zusammengesetzt werden. Für Operatoren S und T von X nach Y U von Y nach Z und $a \in \mathbb{K}$ definiert man

aS :	$D(aS) = D(S)$,	$(aS)x = a(Sx)$,
S + T :	$D(S+T) = D(S) \cap D(T)$,	$(S+T)x = Sx + Tx$,
UT :	$D(UT) = \{x \in D(T) : Tx \in D(U)\}$,	$(UT)x = U(Tx)$.

Satz 9.2 *Seien S und T bijektive Operatoren von X nach Y (als Abb. von $D(S)$ bzw. $D(T)$ nach Y). Ist $D(S) \subset D(T)$, so gilt*

$$T^{-1} - S^{-1} = T^{-1}(S-T)S^{-1}.$$

Ist $D(S) = D(T)$, so gilt

$$T^{-1} - D^{-1} = T^{-1}(S-T)S^{-1} = S^{-1}(S-T)T^{-1}.$$

Anmerkung. Diese Identität entspricht formal der Elemetaren Gleichung $\frac{1}{x} - \frac{1}{y} = \frac{y-x}{xy}$ für $x, y \in \mathbb{K}$.

Beweis. Es genügt offenbar, die erste Aussage zu beweisen. Dazu ist zu zeigen:

$$T^{-1}y = S^{-1}y + T^{-1}(S-T)S^{-1}y \text{ für alle } y \in Y.$$

Tatsächlich gilt für alle $y \in Y$

$$\begin{aligned}
T^{-1}y &= T^{-1}SS^{-1}y && (\text{da } SS^{-1}y = y \text{ für alle } y \in Y)\\
&= T^{-1}[T + (S-T)]S^{-1}y && (\text{da } T + (S-T) = S)\\
&= T^{-1}TS^{-1}y + T^{-1}(ST)S^{-1}y\\
&= S^{-1}y + T^{-1}(S-T)S^{-1}y.
\end{aligned}$$

■

Satz 9.3 (Resolventengleichungen) *Seien S und T lineare Operatoren im normierten Raum X.*

(a) Erste Resolventengleichung*: Für $z, z' \in \varrho(T)$ gilt*

$$R(T,z) - R(T,z') = (z-z')R(T,z)R(T,z') = (z-z')R(T,z')R(T,z).$$

(b) Ist $D(S) \subset D(T)$, so gilt für $z \in \varrho(S) \cap \varrho(T)$

$$R(T,z) - R(S,z) = R(T,z)(S-T)R(S,z).$$

(c) Zweite Resolventengleichung*: Ist $D(S) = D(T)$, so gilt für $z \in \varrho(S) \cap \varrho(T)$*

$$R(T,z) - R(S,z) = R(T,z)(S-T)R(S,z) = R(S,z)(S-T)R(T,z).$$

Der ***Beweis*** folgt unmittelbar aus Satz 8.2.

9.2 Stabilität der Abgeschlossenheit und der stetigen Invertierbarkeit

Seien S und T lineare Operatoren von X nach Y. S heißt T–*beschränkt*, wenn gilt

(i) $D(S) \supset D(T)$,

(ii) es gibt $a, b \geq 0$ mit $\|Sx\| \leq a\|x\| + c\|Tx\|$ für alle $x \in D(T)$.

Das Infimum c aller $b \leq 0$ für die diese Ungleichung mit einem geeigneten $a \geq 0$ gilt, heißt die T–*Schranke* von S. (Im allgemeinen gibt es kein a mit $\|Sx\| \leq a\|x\| + c\|Tx\|$ für alle $x \in D(T)$, wie das Beispiel in Aufgabe 8.1 beweist.)

Ist $S \in B(X;Y)$, so gilt offenbar für *jeden* Operator T von X nach Y: S ist T–beschränkt mit T–Schranke 0. (In diesem Fall kann tatsächlich $b = 0$ gewählt werden, $a = \|S\|$.)

Satz 9.4 (Stabilität der Abgeschlossenheit) *Seien X und Y Banachräume, S und T lineare Operatoren von X nach Y. S sei T–beschränkt mit T–Schranke < 1. Dann ist $T+S$ genau dann abgeschlossen, wenn T abgeschlossen ist.*

Beweis. Da S die T–Schranke < 1 hat, gibt es ein $b < 1$ (zwischen der T–Schranke von S und 1) und ein $a \geq 0$ mit

$$\|Sx\| \leq a\|x\| + b\|Tx\| \text{ für alle } x \in D(T).$$

Daraus folgt für alle $x \in D(T)$

$$\begin{aligned}
&\|Tx\| \leq \|(T+S)x\| + \|Sx\| \leq \|(T+S)x\| + a\|x\| + b\|Tx\|,\\
&(1-b)\|Tx\| \leq a\|x\| + \|(T+S)x\|,\\
&\|x\|_T = \|x\| + \|Tx\| \leq C_1\Big(\|x\| + \|(T+S)x\|\Big) = C_1\|x\|_{T+S},\\
&\|x\|_{T+S} = \|x\| + \|(T+S)x\| \leq \|x\| + \|Tx\| + \|Sx\|\\
&\qquad \leq (1+a)\|x\| + (1+b)\|Tx\| \leq C_2\Big(\|x\| + \|Tx\|\Big) = C_2\|x\|_T.
\end{aligned}$$

Die Normen $\|\cdot\|_T$ und $\|\cdot\|_{T+S}$ auf $D(T) = D(T+S)$ sind also äquivalent, d. h. $(D(T), \|\cdot\|_T)$ ist genau dann vollständig, wenn $(D(T), \|\cdot\|_{T+S}) = (D(T+$

$S), \|\cdot\|_{T+S})$ vollständig ist. Mit Satz 7.10 folgt die Behauptung. ∎

Satz 9.5 (Stabilität der stetigen Invertierbarkeit) *Seien X und Y Banachräume, S und T seien lineare Operatoren von X nach Y, T sei abgeschlossen und bijektiv (also $T^{-1} \in B(X,Y)$), uns es gelte $D(S) \supset D(T)$ und $\|ST^{-1}\| < 1$. Dann ist auch $T+S$ abgeschlossen und bijektiv; es gilt*

$$(T+S)^{-1} = \sum_{n=0}^{\infty}(-1)^n T^{-1}(ST^{-1})^n = \sum_{n=0}^{\infty}(-1)^n (T^{-1}S)^n T^{-1},$$

wobei die Reihen bezüglich der Norm in $B(Y,X)$ konvergieren.

Anmerkung. Die Reihen sind offenbar identisch, *formal* erhält man sie, indem man in

$$\begin{aligned}(T+S)^{-1} &= [T(I+T^{-1}S)]^{-1} = (I+T^{-1}S)^{-1}T^{-1}\\ \text{bzw.} &= [(I+ST^{-1})T]^{-1} = T^{-1}(I+ST^{-1})^{-1}\end{aligned}$$

die Terme $(I+T^{-1}S)^{-1}$ bzw. $(IST^{-1})^{-1}$ als geometrische Reihen schreibt.

Beweis. Einige Teile des Beweises werden wesentlich einfacher, wenn T (und damit auch S) in $B(X,Y)$ ist.

Nach Voraussetzung gilt $D(S) \supset D(T)$. Für jedes $x \in D(T)$ gilt

$$\|Sx\| = \|ST^{-1}Tx\| \le \|ST^{-1}\|\|Tx\|,$$

d. h. S ist T–beschränkt mit T–Schranke $\le \|ST^{-1}\| < 1$. Nach Satz 8.4 ist also auch $T+S$ abgeschlossen.

Für alle $x \in D(T+S) \setminus \{0\} = D(T) \setminus \{0\}$ gilt

$$\|(T+S)x\| \ge \|Tx\| - \|Sx\| = \|Tx\| - \|(ST^{-1})Tx\| \ge (1 - \|ST^{-1}\|)\|Tx\| > 0,$$

d. h. auch $T+S$ ist injektiv.

Sei

$$A_p := \sum_{n=0}^{p}(-1)^n T^{-1}(ST^{-1})^n \text{ für } p \in \mathbb{N}.$$

Dann gilt für $q > p$

$$\begin{aligned}\|A_q - A_p\| &= \Big\| \sum_{n=p+1}^{q} (-1)^n T^{-1}(ST^{-1})^n \Big\| \\ &\leq \|T^{-1}\| \|ST^{-1}\|^{p+1} \sum_{n=0}^{q-p-1} \|ST^{-1}\|^n \\ &\leq \|T^{-1}\| \|ST^{-1}\|^{p+1} \sum_{n=0}^{\infty} \|ST^{-1}\|^n \leq C \|ST^{-1}\|^{p+1}.\end{aligned}$$

also ist (A_p) eine Cauchyfolge in $B(Y, X)$. Da $B(Y, X)$ vollständig ist, existiert ein $A \in B(Y, X)$ mit $\|A - A_p\| \to 0$ für $p \to \infty$.

$(\mathbf{T} + \mathbf{S})\mathbf{A} = \mathbf{I}$: Offensichtlich gilt für $p \to \infty$

$$(T + S)A_p = \sum_{n=0}^{p} (-1)^n (T + S)T^{-1}(ST^{-1})^n = I + (-1)^p (ST^{-1})^{p+1} \to I.$$

Für *jedes* $y \in Y$ gilt also

$$A_p y \to Ay \text{ und } (T + S)A_p y \to y \text{ für } p \to \infty.$$

Da $T + S$ abgeschlossen ist, folgt hieraus für jedes $y \in Y$

$$Ay \in D(T + S) \text{ und } (T + S)Ay = y.$$

Das bedeutet $(T + S)A = I$, wie behauptet.

Hieraus folgt, daß $T + S$ surjektiv (also bijektiv) ist und $A = (T + S)^{-1}$. ■

9.3 Die Resolvente

Sei T ein Operator im normierten Raum X. Nach Abschnitt 8.1 ist die Resolvente von T die Funktion

$$R(T, \cdot) : \varrho(T) \to B(X); \quad R(T, z) = (T - z)^{-1}.$$

Satz 9.6 *Sei X ein Banachraum, T ein abgeschlossener Operator in X.*

(a) $\varrho(T)$ ist eine offene Teilmenge von $\mathbb{K}$ ($\sigma(T)$ ist also abgeschlossen).

(b) Die Resolvente von T ist analytisch *in $\varrho(T)$, d. h. sie ist um jeden Punkt z_0 von $\varrho(T)$ in eine normkonvergente Potenzreihe entwickelbar.*

Beweis. a) Sei $z_0 \in \varrho(T)$, $r := \|(T-z_0)^{-1}\|^{-1}$. Dann gilt

$$\|(z-z_0)(T-z_0)^{-1}\| < 1 \text{ für } z \in \mathbb{K} \text{ mit } |z_0 - z| < r.$$

Nach Satz 8.5 ist also

$$T - z = (T - z_0) + (z_0 - z) \text{ bijektiv für } z \in \mathbb{K} \text{ mit } |z - z_0| < r.$$

Also liegt auch jedes $z \in \mathbb{K}$ mit $|z - z_0| < r$ in $\varrho(T)$, d. h. $\varrho(T)$ ist offen.

b) Seien z_0 und r wie in Teil a. Nach Satz 8.5 gilt für $z \in \mathbb{K}$ mit $|z - z_0| < r$

$$\begin{aligned} R(T,z) &= \Big([T-z_0] + [z_0 - z]\Big)^{-1} = \sum_{n=0}^{\infty}(T-z_0)^{-1}\Big\{(z-z_0)(T-z_0)^{-1}\Big\}^n \\ &= \sum_{n=0}^{\infty}(z-z_0)^n R(T,z_0)^{n+1}. \end{aligned}$$

Wegen $|z-z_0|\|R(T,z_0)\| < r\|(T-z_0)^{-1}\| = 1$ ist diese Reihe normkonvergent. Das ist die behauptete Analytizität von $(T-z)^{-1}$. ∎

Korollar 9.7 *Die Resolvente $R(T,\cdot)$ ist normstetig in $\varrho(T)$, d. h.*

$$\|R(T,z') - R(T,z)\| \to 0 \text{ für } z' \to z.$$

Beweis. Aus der letzten Identität im Beweis von Satz 8.6 folgt

$$\begin{aligned} &R(T,z') - R(T,z) = \sum_{n=1}^{\infty}(z'-z)^n R(T,z)^{n+1}, \\ &\|R(T,z') - R(T,z)\| \le |z'-z|\|R(T,z)\|^2 \sum_{n=0}^{\infty}\Big\{|z'-z|\|R(T,z)\|\Big\}^n \to 0 \end{aligned}$$

für $z' \to z$. ∎

Satz 9.8 *Sei X ein Banachraum, $T \in B(X)$.*

(a) Es gilt

$$\sigma(T) \subset \Big\{z \in \mathbb{K} : |z| \le \|T\|\Big\}, \varrho(T) \supset \Big\{z \in \mathbb{K} : |z| > \|T\|\Big\}$$

Für $|z| > \|T\|$ (genauer: für $|z| > \limsup \|T^n\|^{1/n}$) gilt

$$(T-z)^{-1} = -\sum_{n=0}^{\infty} z^{-n-1}T^n \quad (\text{Neumann'sche Reihe}).$$

(b) Ist $\mathbb{K} = \mathbb{C}$, so ist $\sigma(T) \neq \emptyset$ und $\max\{|z| : z \in \sigma(T)\} = \limsup \|T^n\|^{1/n}$; diese Zahl heißt der Spektralradius *von T.*

Beweis. a) Für $|z| > \|T\|$ ist (natürlich) zI bijektiv und es gilt $\|Tz^{-1}\| = |z|^{-1}\|T\| < 1$. Nach Satz 8.5 (mit zI statt T und $-T$ statt S) ist also auch $T - z = -(zI - T)$ bijektiv und es gilt

$$(T-z)^{-1} = -(zI-T)^{-1} = -\sum_{n=0}^{\infty}(-1)^n z^{-1}(-Tz^{-1} = -\sum_{n=0}^{\infty} z^{-n-1}T^n.$$

Die Neumann'sche Reihe konvergiert offensichtlich auch für $|z| > \limsup \|T^n\|^{1/n}$ gegen ein $A \in B(X)$ mit $(T-z)A = A(T-z) = I$ (beim Ausmultiplizieren fallen alle Terme bis auf I weg), d. h. auch für diese z gilt $z \in \varrho(T)$.

b) Wir nehmen an, daß $\sigma(T)$ leer ist, d. h. es gilt $\varrho(T) = \mathbb{C}$; $R(T,\cdot)$ ist also in ganz $\mathbb{C}$ analytisch. Für jedes $x \in X$ und $F \in X^*$ ist deshalb die Funktion

$$f_{x,F} : \mathbb{C} \to \mathbb{C}, \quad f_{x,F}(z) = F(R(T,z)x)$$

eine ganze Funktion (ist nämlich in einer Umgebung von z_0

$$R(T,z) = \sum_{n=0}^{\infty}(z-z_0)^n A_n \quad (A_n = R(T,z_0)^{n+1} \in B(X))$$

mit einer normkonvergenten Potenzreihe, so gilt dort

$$f_{x,F}(z) = \sum_{n=0}^{\infty} F(A_n x)(z-z_0)^n,$$

d. h. $f_{x,F}$ ist bei z_0 holomorph).

Außerdem gilt (Neumann'sche Reihe) für $|z| > \|T\|$

$$|f_{x,F}(z)| \le \|F\|\|x\||z|^{-1}\sum_{n=0}^{\infty}\Big\{\|T\||z|^{-1}\Big\}^n,$$

also $|f_{x,F}(z)| \to 0$ für $|z| \to \infty$.

Nach dem Ersten Satz von Liouville ist also

$$F(R(T,z)x) = f_{x,F}(z) \equiv 0 \text{ für alle } x \in X, F \in X^*.$$

Daraus folgt $R(T,z)x \equiv 0$ für alle $x \in X$; das ist unmöglich, da $R(T,z)$ als inverser Operator notwendig injektiv ist.

Ist $r_0 = \max\{|z| : z \in \sigma(T)\}$, so gilt $r_0 \leq \limsup \|T^n\|^{1/n}$ nach Teil a. $(T-z)^{-1}$ ist für $|z| > r_0$ holomorph und somit als Laurentreihe um 0 darstellbar. Die Laurentreihe (= Neumann'sche Reihe) ist aber für $|z| < \limsup \|T^n\|^{1/n}$ divergent, also ist $r_0 \geq \limsup \|T^n\|^{1/n}$. ■

(Die letzte Argumentation ist nicht ganz vollständig, da sie ohne Beweis Resultate der komplexen analytischen Funktionen für Banachraumwertige analytische Funktionen verwendet. Details findet man bei J. Weidmann, [11], Satz 5.9)

9.4 Beispiele

Wir betrachten zunächst den durch eine μ–meßbare Funktion $t : X \to \mathbb{K}$ erzeugten maximalen *Multiplikationsoperator* in $L_p(X,\mu)$. Wir wissen:

— M_t ist abgeschlossen (vgl. Satz 7.14 a),

— $D(M_t)$ ist dicht für $1 \leq p < \infty$ (vgl. Satz 7.14 b),

— M_t ist genau dann beschränkt, wenn t wesentlich beschränkt ist, $\|M_t\| =$ wes–sup $|t|$ (vgl. Aufgabe 6.4a).

Satz 9.9 *Es gilt*

$$\begin{aligned}
\varrho(M_t) &= \Big\{z \in \mathbb{K} : \exists \varepsilon > 0 \textit{ mit } \mu\{x \in X : |t(x) - z| < \varepsilon\} = 0\Big\} =: R \\
\sigma(M_t) &= \Big\{z \in \mathbb{K} : \forall \varepsilon > 0 \textit{ ist } \mu\{x \in X : |t(x) - z| < \varepsilon\} > 0\Big\} =: S \\
(M_t - z)^{-1} &= M_{1/(t-z)} \textit{ für } z \in \varrho(M_t).
\end{aligned}$$

Beweis. Ist $M_t - z$ injektiv, so ist jedenfalls $(M_t - z)^{-1}$ eine Einschränkung von $M_{1/(t-z)}$.

$\varrho(\mathbf{M_t}) \subset \mathbf{R}$: Ist $z \in \varrho(M_t)$, so ist $D((M_t - z)^{-1}) = L_p(X, \mu)$, also $(M_t - z)^{-1} = M_{1/(t-z)}$, und dieser Operator ist beschränkt, d. h. $1/(t-z)$ ist wesentlich beschränkt. Also gibt es ein $C \geq 0$ und eine Nullmenge $N \subset X$ mit

$$|t(x) - z|^{-1} \leq C, \quad |t(x) - z| \geq \frac{1}{C} \text{ für } x \in X \setminus N.$$

Das bedeutet $z \in R$.

$\mathbf{R} \subset \varrho(\mathbf{M_t})$: Ist $z \in R$, so existiert ein $\varepsilon > 0$ mit $|t(x) - z| > \varepsilon$ μ–f.u. Also ist $M_t - z = M_{t-z}$ injektiv. Für jedes $f \in L_p(X, \mu)$ gilt

$$f = (t - z)\{(t - z)^{-1} f\} \qquad mit \quad (t - z)^{-1} f \in L_p(X, \mu),$$

also $f \in R(M_t - z)$; d. h. $M_t - z$ ist surjektiv. Da M_t abgeschlossen ist, ist somit $z \in \varrho(M_t)$.

Wegen $S = \mathbb{K} \setminus R$ ist damit auch die zweite Aussage bewiesen; die Gleichung $M_t - z = M_{(t-z)^{-1}}$ für $z \in \varrho(M_t)$ wurde bereits im ersten Teil mitbewiesen. ■

Im folgenden betrachten wir noch einen einfachen Differentialoperator in $L_p(a, b)$. Nach Abschnitt 5.5 heißt eine Funktion : $[a, b] \to \mathbb{C}$ *absolut stetig*, wenn ein $g \in L_1(a, b)$ existiert mit

$$f(x) = f(a) + \int_a^x g(t)\,\mathrm{d}t \text{ für } x \in [a, b].$$

Die Funktion g heißt „*Ableitung*“von f, $g = f'$.

Wir definieren den Operator T in $L_p(a, b)$ durch

$$D(T) = \{f \in L_p(a, b) : f \text{ absolut stetig}, f' \text{ absolut stetig},$$
$$f(a) = f(b) = 0, f'' \in L_p(a, b).$$
$$Tf = -f'' \text{ für } f \in D(T).$$

Dieser Operator ist (insbesondere in $L_2(a, b)$) von Bedeutung bei der mathematischen Untersuchung der schwingenden Saite.

Satz 9.10 *T ist abgeschlossen für alle $p \in [1, \infty]$.*

Beweis. Sei (f_n) eine Folge aus $D(T)$ mit

$$f_n \to f,\ Tf_n \to g\ (\text{ d. h. } -f'' \to g) \text{ in } L_p(a,b).$$

Es gilt, da f_n und f_n' absolut stetig sind und $f_n(a) = 0$ gilt,

$$\begin{aligned} f_n'(x) &= f_n'(a) + \int_a^x f_n''(t)\,\mathrm{d}t, \\ f_n(x) &= f_n(a) + \int_a^x f_n'(s)\,\mathrm{d}s = \int_a^x f_n'(s)\,\mathrm{d}s \qquad (\text{da } f_n \in D(T), f_n(a) = 0) \qquad (*) \\ &= (x-a)f_n'(a) + \int_a^x \int_a^s f_n''(t)\,\mathrm{d}t\,\mathrm{d}s. \end{aligned}$$

Dabei ist der letzte Ausdruck wegen

$$\begin{aligned} &\Big| \int_a^x \int_a^s f_n''(t)\,\mathrm{d}t\,\mathrm{d}s + \int_a^x \int_a^s g(t)\,\mathrm{d}t\,\mathrm{d}s \Big| \\ &\leq \int_a^x \int_a^s |f_n''(t) + g(t)|\,\mathrm{d}t\,\mathrm{d}s \leq \int_a^b \int_a^b |f_n''(t) + g(t)|\,\mathrm{d}t\,\mathrm{d}s \\ &\leq \int_a^b (b-a)^{1/q} \Big\{ \int_a^b |f_n''(t) + g(t)|^p\,\mathrm{d}t \Big\}^{1/p} \mathrm{d}s = (b-a)^{1+1/q} \|f_n'' + g\|_p \\ &\to 0 \text{ für } n \to \infty \end{aligned}$$

gleichmäßig konvergent gegen

$$-\int_a^x \int_a^s g(t)\,\mathrm{d}t\,\mathrm{d}s.$$

Daraus folgt, daß die Folge $(f_n'(a))$ konvergiert (sonst könnte die rechte Seite von $(*)$ nicht in $L_p(a,b)$ konvergieren); sei $c := \lim f_n'(a)$. Die rechte Seite von $(*)$ ist also gleichmäßig konvergent. Da die linke Seite in $L_p(a,b)$ konvergiert, folgt

$$f(x) = c(x-a) - \int_a^x \int_a^s g(t)\,\mathrm{d}t\,\mathrm{d}s.$$

Dies gilt zunächst fast überall; der Repräsentant der linken Seite kann so abgeändert werden, daß die Gleichung für alle x gilt. Hieran erkennt man sofort, daß f und f' absolut stetig sind und $f'' = -g$ in $L_p(a,b)$ liegt. Also ist $f \in D(T)$ und $Tf = g$. ∎

Satz 9.11 *(a) Es gilt*

$$\sigma(T) = \Big\{\frac{n^2\pi^2}{(b+a)^2} : n \in \mathbb{N}\Big\}.$$

Jedes $\lambda_n = n^2\pi^2(b-a)^{-2}$ ist Eigenwert von T mit dem (nicht normierten) Eigenelement

$$e_n(x) = \sin\Big\{\frac{n\pi(x-a)}{b-a}\Big\}.$$

(b) Für $z \in \varrho(T) = \mathbb{C} \setminus \sigma(T), z \neq 0$ gilt (vgl. Aufgabe 8.5 für $z = 0$)

$$R(T,z)g(x) = \frac{1}{\sqrt{z}\sin\sqrt{z}(b-a)}\Big\{\sin\sqrt{z}(b-x)\int_a^x \sin\sqrt{z}(y-a)g(y)\,\mathrm{d}y$$
$$+ \sin\sqrt{z}(x-a)\int_x^b \sin\sqrt{z}(b-y)g(y)\,\mathrm{d}y\Big\}.$$

Anmerkung. Es ist bemerkenswert, daß das Spektrum, die Resolventenmenge und die Form der Resolvente von p unabhängig sind; dies ist ein Phänomen, das bei vielen (aber nicht allen) Operatoren auftritt; vgl. K. Jörgens, Lineare Integraloperatoren, Abschnitt 12.5, Aufgabe 12.11 für ein Gegenbeispiel.

Beweis. Für jedes $z \in \mathbb{C} \setminus \{0\}$ ist ein $\sin\sqrt{z}(x-a)$ eine Lösung von $-f'' = zf$ mit $f(a) = 0$. Damit diese Lösung auch die Randbedingung $f(b) = 0$ erfüllt, muß $z = n^2\pi^2(b-a)^{-2}$ sein mit einem $n \in \mathbb{N}$, also

$$f(x) = \sin\{n\pi(b-a)^{-1}(x-a)\}.$$

Für $z = 0$ sind $f(x) = c(x-a)$ die einzigen Lösungen mit $f(a) = 0$; nur für $c = 0$ erfüllt diese auch die Randbedingung bei b, dann ist aber $f \equiv 0$. Also ist $T - z$ injektiv für $z \neq n^2\pi^2(b-a)^{-2} (n \in \mathbb{N})$.

Ist $z \neq n^2\pi^2(b-a)^2$ $(n \in \mathbb{N})$ und $z \neq 0$, so prüft man leicht nach, daß für jedes $g \in L_p(a,b)$

$$f(x) = \frac{1}{\sqrt{z}\sin\sqrt{z}(b-a)}\Big\{\sin\sqrt{z}(b-x)\int\limits_a^x \sin\sqrt{z}(y-a)g(y)\,\mathrm{d}y + \sin\sqrt{z}(x-a)\int\limits_x^b \sin\sqrt{z}(b-y)g(y)\,\mathrm{d}y\Big\}$$

eine Lösung von $-f''-zf=g$ ist mit $f(a)=f(b)=0$. (man gewinnt diesen Ausdruck mit Hilfe der Variation der Konstanten.) Also ist $(T-z)$ bijektiv, $z \in \varrho(T)$, und der angegebene Ausdruck stellt die Resolvente $R(T,z)$ dar. ■

9.5 Übungsaufgaben

9.1 In $L_pp(\mathbb{R})$ seien T und S definiert durch

$$\begin{aligned} D(T) &= \{f \in L_p : x^2 f \in L_p\}, \quad Tf(x) = x^2 f(x), \\ D(S) &= \{f \in L_p : xf \in L_p\}, \quad Sf(x) = xf(x). \end{aligned}$$

Man zeige: S ist T–beschränkt mit T–Schranke 0 (aber es gilt natürlich nicht $\|Sf\| \leq a\|f\| + 0\|Tf\|$).

9.2 (a) S_1 und S_2 seien T–beschränkt mit T–Schranken b_1 bzw. b_2. Man zeige: $S_1 + S_2$ ist T–beschränkt mit T–Schranke $\leq b_1 + b_2$.

(b) S sei T_1–beschränkt mit T_1–Schranke b_1, T_1 sei T–beschränkt mit T–Schranke b_2. Dann ist S T–beschränkt mit T–Schranke $\leq b_1 b_2$.

9.3 Aus Satz 8.8 wissen wir: Jeder Operator $T \in B(X)$ hat ein nicht–leeres Spektrum, falls X ein komplexer Banachraum ist.

(a) Dies gilt i.allg. nicht, wenn X ein reeller Banachraum ist.

(b) Es gibt abgeschlossene (natürlich unbeschränkte) Operatoren T in komplexen Banachräumen mit $\sigma(T) = \emptyset$. Beispiel in $L_p(0,1)$:

$$D(T) = \Big\{f \in l_p(0,1) : f \text{ absolut stetig, } f(0) = 0, f' \in L_p(0,1)\Big\}, \quad Tf = f'.$$

9.4 Sei X ein Banachraum. Ein Operator $T \in B(X)$ heißt *quasi-nilpotent*, wenn gilt $\|T^n\|^{1/n} \to 0$ für $n \to \infty$.

(a) Ist T quasi–nilpotent im komplexen Banachraum X, so gilt $\sigma(T) = \{0\}$. (Mit Hilfe der Komplexifizierung kann man das auch für reelle Banachräume beweisen.)

(b) In ℓ_2 sei $T(x_1, x_2, \ldots,) = \left(\frac{1}{2}x_2, \frac{1}{3}x_3, \ldots\right)$. Dann ist T quasi–nilpotent.

(c) In $C[0,1]$ oder $L_p(0,1)$ sei

$$Kf(x) = \int_0^x k(x,y)f(y)\,\mathrm{d}y$$

mit stetigem $k : [0,1] \times [0,1] \to \mathbb{C}$, $|k(x,y)| \leq M$. Dann ist K quasi–nilpotent.
Anleitung: Durch Induktion zeige man für den Kern $k^{(n)}(x,y)$ von K^n:

$$|k^{(n)}(x,y)| \leq \frac{M^n}{(n-1)!}|x-y|^{n-1}, \text{ also } \|K^n\| \leq \frac{M^n}{(n-1)!}.$$

9.5 Mit Hilfe der Variation der Konstanten gewinne man die in Satz 8.11 angegebene Resolvente und zeige, daß für $z = 0$ gilt

$$R(T,0)g(x) = (a-b)^{-1}\left\{(x-a)\int_a^x (t-b)g(t)\,\mathrm{d}t + (x-b)\int_x^b (t-a)g(t)\,\mathrm{d}t\right\}.$$

9.6 (a) Der in Satz 8.10 und Satz 8.11 studierte Operator hat für $p = \infty$ einen *nicht* dichten Definitionsbereich in $L_\infty(a,b)$.

(b) Dieser Operator kann entsprechend in $(C[a,b], \|\cdot\|_\infty)$ untersucht werden. Er hat auch in diesem Raum das gleiche Spektrum und formal die gleiche Resolvente wie in $L_p(a,b)$.

10 Hilberträume, der Satz von Radon–Nikodym

10.1 Grundbegriffe der Hilbertraumtheorie

Sei X ein Vektorraum über $\mathbb{K}$. Eine Abbildung $s : X \times X \to \mathbb{K}$ heißt eine *Sesquilinearform* auf X, wenn gilt

$$\left.\begin{aligned} s(x, ay + bz) &= as(x, y) + bs(x, z) \\ s(ax + by, z) &= \overline{a}s(x, z) + \overline{b}s(y, z) \end{aligned}\right\} \textit{ für } x, y, z \in X,\ a, b \in \mathbb{K}.$$

(In der Literatur werden Sesquilinearformen häufig so definiert, daß sie nicht in der ersten, sondern in der zweiten Variablen *konjugiert linear* sind. Dies hat natürlich Auswirkungen auf manche Formeln, z. B. Satz 9.1 und Satz 9.7)

Die durch s erzeugte *quadratische Form* $q = q_s : X \to \mathbb{K}$ ist definiert durch

$$q(x) := s(x, x) \text{ für } x \in X.$$

Es gilt also

$$q(ax) = |a|^2 q(x) \quad (\text{speziell: } q(ax) = q(x) \text{ für } |a| = 1)$$

Die Sesquilinearform s heißt *hermitesch*, wenn gilt

$$s(x, y) = \overline{s(y, x)} \text{ für alle } x, y \in X.$$

Für eine hermitesche Sesquilinearform s ist die zugehörige quadratische Form *reell*:

$$\overline{q(x)} = \overline{s(x, x)} = q(x) \text{ für alle } x \in X.$$

Satz 10.1 *Sei s ein Sesquilinearform, q die durch s erzeugte quadratische Form. Dann gilt*

(a) Die Polarisierungsidentität

$$s(x, y) = \begin{cases} \frac{1}{4}\Big\{q(x+y) - q(x-y) + iq(x-iy) - q(x+iy)\Big\} & (\mathbb{K} = \mathbb{C}), \\ \frac{1}{4}\Big\{q(x+y) - q(x-y)\Big\} & (\mathbb{K} = \mathbb{R},\ s \textit{ hermitesch}). \end{cases}$$

(b) Die Parallelogrammidentität

$$q(x+y) + q(x-y) = 2\{q(x) + q(y)\}$$

(eine geometrische Deutung dieser Identität kann in ihrer speziellen Version für eine durch ein Skalarprodukt erzeugte Norm gegeben werden.,Satz 9.7).

Der ***Beweis*** ergibt sicht leicht durch Ausrechnen der rechten Seite im Fall a) bzw. der linken Seite im Fall b).

Ist $\mathbb{K} = \mathbb{C}$ und q reellwertig, so folgt aus der Polarisierungsidentität leicht, daß s hermitesch ist. (Eine hermitesche Sesquilinearform auf einem reellen Vektorraum [der Querstrich ist dann bedeutungslos] heißt *symmetrisch*.)

Eine hermitesche Sesquilinearform heißt *positiv* (bzw. *nicht–negativ*), wenn gilt

$$s(x,x) > 0 \text{ für alle } x \in X \backslash \{0\} \text{ (bzw. } s(x,x) \geq 0 \text{ für alle } x \in X \text{).}$$

Satz 10.2 (Schwarzsche Ungleichung) *(a) Ist s eine nicht–negative Sesqulinearform und q die zugehörige quadratische Form, so gilt die* Schwarzsche Ungleichung

$$|s(x,y)| \leq (q(x)q(y)]^{1/2} \textit{ für alle } x, y \in X.$$

(b) Sei jetzt s positiv. *In der Schwarzschen Ungleichung gilt „=“ genau dann, wenn x und y linear abhängig sind; die Gleichung $s(x,y) = (q(x)q(y))^{1/2}$ gilt genau dann, wenn $ax = by$ gilt mit $a, b \geq 0$ und $(a,b) \neq (0,0)$ (d. h. wenn x und y* positiv linear abhängig *sind).*

Beweis. Wir beweisen hier nur Teil a; für Teil b vgl. J. Weidmann [11]
Für alle $x, y \in X$ und $t \in \mathbb{R}$ gilt, da s nicht–negativ ist,

$$0 \leq q(x+ty) = q(x) + 2t \operatorname{Re} s(x,y) + t^2 q(y).$$

Da ein reelles Polynom $at^2 + 2bt + c$ mit $a \geq 0$ genau dann für alle t nicht–negativ ist, wenn $b^2 - ac \leq 0$ gilt, folgt

$$(\operatorname{Re} s(x,y))^2 \leq q(x)q(y). \tag{$*$}$$

Wählt man nun $a \in \mathbb{K}$ so, daß $|a| = 1$ und $a \cdot s(x,y) = |s(x,y)|$ gilt, so folgt aus $(*)$

$$\begin{aligned}|s(x,y)|^2 &= (\text{Re } as(x,y))^2 = (\text{Re } s(x,ay))^2 \leq q(x)q(ay) \\ &= |a|^2 q(x)q(y) = q(x)q(y).\end{aligned}$$

■

Eine positive (nicht–negative) Sequilinearform heißt ein *Skalarprodukt (Semiskalarprodukt)*. Skalarprodukte schreibt man meist in der Form $\langle\cdot,\cdot\rangle$ statt $s(\cdot,\cdot)$. Ein Skalarprodukt $\langle\cdot,\cdot\rangle$ hat also die Eigenschaften

(S 1) $\langle x, y+z\rangle = \langle x,y\rangle + \langle x,z\rangle$,

(S 2) $\langle x, ay\rangle = a\langle x,y\rangle$,

(S 3) $\langle x,y\rangle = \overline{\langle y,x\rangle}$,

(S 4) $\langle x,x\rangle \geq 0$,

(S 5) aus $x \neq 0$ folgt $\langle x,x\rangle > 0$.

Gelten nur die Eigenschaften (S 1) – (S 4), so ist $\langle\cdot,\cdot\rangle$ ein Semi–Skalarprodukt. Aus (S 2) und (S 3) folgt insbesondere

(S 6) $\langle ax,y\rangle = \overline{a}\langle x,y\rangle$.

Satz 10.3 *Ist $\langle\cdot,\cdot\rangle$ ein Skalarprodukt auf X, so ist durch*

$$\|x\| := \langle x,x\rangle^{1/2} \text{ für } x \in X$$

eine Norm *auf X erklärt. Ist $\langle\cdot,\cdot\rangle$ ein Semi–Skalarprodukt, so ist $\|\cdot\|$ eine* Halbnorm.

Beweis. Offensichtlich ist $\|x\| \geq 0$ für alle $x \in X$.

(N 1) $\|ax\| = \langle ax,ax\rangle^{1/2} = |a|\langle x,x\rangle^{1/2} = |a|\,\|x\|$.

(N 2) Mit der Schwarzschen Ungleichung folgt

$$\begin{aligned}&\|x+y\|^2 = \|x\|^2 + 2\,\text{Re}\,\langle x,y\rangle + \|y\|^2 \\ &\quad\leq \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.\end{aligned}$$

(N 3) Falls $\langle\cdot,\cdot\rangle$ ein Skalarprodukt ist, gilt $\|x\| = \langle x,x\rangle^{1/2} > 0$ für $x \neq 0$. ■

Ein normierter Raum $(X, \|\cdot\|)$, dessen Norm durch ein Skalarprodukt $\langle\cdot,\cdot\rangle$ erzeugt wird, heißt ein *Prä–Hilbertraum*; soll die Tatsache, daß es sich um einen *Raum* *mit Skalarprodukt* handelt, betont werden, so schreibt man auch $(X, \langle\cdot,\cdot\rangle)$. Ein vollständiger Prä–Hilbertraum heißt ein *Hilbertraum.*

Der Satz 4.17 läßt sich leicht so modifizieren, daß er besagt: Jeder Prä–Hilbertraum X kann als dichter Teilraum eines Hilbertraumes $\tilde{X}$ aufgefaßt werden. Ein solcher Hilbertraum heißt eine *Vervollständigung* des Prä–Hilbertraumes X.

Beispiel 10.4 Der Raum $C[a,b]$ mit dem Skalarprodukt

$$\langle f,g\rangle = \int\limits_a^b \overline{f(x)}g(x)\,\mathrm{d}x \text{ für } f,g \in C[a,b]$$

und der zugehörigen Norm

$$\|f\| = \left\{\int\limits_a^b |f(x)|^2\,\mathrm{d}x\right\}^{1/2} \text{ für } f \in C[a,b]$$

ist ein *Prä–Hilbertraum* (aber *kein* Hilbertraum, vgl. Beispiel 4.7) □

Beispiel 10.5 Der Folgeraum ℓ_2 mit dem Skalarprodukt

$$\langle f,g\rangle = \sum \overline{f_n}g_n \text{ für } f=(f_n),\ (g_n) \in \ell_2$$

und der zugehörigen Norm

$$\|f\|_2 = \left\{\sum |f_n|^2\right\}^{1/2} \text{ für } f=(f_n) \in \ell_2$$

ist ein *Hilbertraum*, der *Hilbertsche Folgenraum* (vgl. Abschnitt 4.3). □

Beispiel 10.6 Sei μ ein Maß auf X. Der Raum $L_2(X,\mu)$ mit dem Skalarprodukt

$$\langle f,g\rangle = \int \overline{f(x)}g(x)\,\mathrm{d}\mu(x) \text{ für } f,g \in L_2(X,\mu)$$

und der hierdurch erzeugten Norm

$$\|f\|_2 = \left\{\int |f(x)|^2\,\mathrm{d}\mu(x)\right\}^{1/2} \text{ für } f \in L_2(X,\mu)$$

ist ein *Hilbertraum* (vgl. Satz 6.3). □

Satz 10.7 *Sei $\langle\cdot,\cdot\rangle$ ein Skalarprodukt auf dem Vektorraum $X\|\cdot\|$ die hierdurch erzeugte Norm. Dann gilt:*

(a) *Die* Polarisierungsidentität

$$\langle x, y\rangle = \begin{cases} \frac{1}{4}\Big\{\|x+y\|^2 - \|x-y\|^2 + i\|x-iy\|^2 - i\|x+iy\|^2\Big\} & \mathbb{K}=\mathbb{C} \\ \frac{1}{4}\Big\{\|x+y\|^2 - \|x-y\|^2\Big\} & \mathbb{K}=\mathbb{R}. \end{cases}$$

(b) *Die* Parallelogrammidentität

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

(Anschaulich besagt die Parallelogrammidentität, daß die Summe der Quadrate aller vier Seiten eines Parallelogramms gleich der Summe der Quadrate der Diagonalen ist).

Der ***Beweis*** folgt unmittelbar aus Satz 9.1.

Satz 10.8 *Sei $(X,\langle\cdot,\cdot\rangle)$ ein Prä–Hilbertraum. Aus $x_n \to x$ und $y_n \to y$ folgt $\langle x_n, y_n\rangle \to \langle x, y\rangle$; d. h. das Skalarprodukt ist stetig bezüglich beider Variablen.*

Beweis. Es gibt ein C mit $\|y_n\| \leq C$ für alle n. Damit folgt

$$\begin{aligned} \Big|\langle x_n, y_n\rangle - \langle x, y\rangle\Big| &\leq |\langle x_n - x, y_n\rangle| + |\langle x, y_n - y\rangle| \\ \leq\ & \|x_n - x\|\,\|y_n\| + \|x\|\,\|y_n - y\| \leq \|x_n - x\|C + \|x\|\,\|y_n - y\| \to 0. \end{aligned}$$

■

Satz 10.9 (Satz von Jordan und von Neumann) *Eine Norm auf einem Vektorraum wird genau dann durch ein Skalarprodukt erzeugt, wenn sie die Parallelogrammidentität erfüllt. Das Skalarprodukt wird dann mit Hilfe der Polarisierungsidentität gegeben.*

Für den Beweis dieses Satzes, den wir hier nicht benötigen, sei auf J. Weidmann [11], Satz 1.6, verwiesen.

10.2 Orthogonalität, der Projektsatz

Sei $(X, \langle\cdot,\cdot\rangle)$ ein Prä–Hilbertraum. Zwei Elemente $x, y \in X$ heißen *orthogonal*, $x \perp y$, wenn gilt $\langle x, y\rangle = 0$. (In $\mathbb{R}^2$ und $\mathbb{R}^3$ mit dem üblichen Skalarprodukt bedeutet dies, daß der Winkel zwischen x und y ein rechter ist [oder, daß einer der Vektoren Null ist].)

Offensichtlich gilt:

$$\text{aus } x \perp y \text{ folgt } \|x+y\|^2 = \|x\|^2 + \|y\|^2,$$

der sogenannte „*Satz des Pythagoras*“.

Man definiert weiter:

Ist $x \in X$ und $A \subset X$, so heißt x *orthogonal* zu A, $x \perp A$, wenn $x \perp y$ für jedes $y \in A$ gilt.

Sind $A, B \subset X$, so heißt A *orthogonal* zu $B, A \perp B$, wenn $x \perp y$ für alle $x \in A$, $y \in B$ gilt.

Für $A \subset X$ ist $A^\perp$ definiert durch $A^\perp := \{x \in X : x \perp A\}$.

Satz 10.10 *Sei $(X, \langle\cdot,\cdot\rangle)$ ein Prä–Hilbertraum. Dann gilt:*

(a) $\{0\}^\perp = X, \quad X^\perp = \{0\}$.

(b) Für jede Teilmenge $A \subset X$ ist $A^\perp$ ein abgeschlossener Teilraum von X; man nennt deshalb $A^\perp$ den Orthogonalraum *zu A.*

(c) Aus $A \subset B$ folgt $B^\perp \subset A^\perp$.

(d) Für jede Teilmenge $A \subset X$ gilt $A^\perp = L(A)^\perp = \overline{L(A)}^\perp$.

Beweis. a) Es gilt $\langle 0, x\rangle = \langle 0 \cdot 0, x\rangle = 0\langle 0, x\rangle = 0$ für alle $x \in X$. Für jedes $x \in X\setminus\{0\}$ gilt $\langle x, x\rangle \neq 0$.

b) Auf Grund der Linearitätseigenschaften des Skalarprodukts ist klar, daß $A^\perp$ ein Vektorraum ist. Sei $x \in \overline{A^\perp}$, (x_n) eine Folge aus $A^\perp$ mit $x_n \to x$. Dann gilt (vgl. Satz 9.8)

$$\langle x, y\rangle = \lim \langle x_n, y\rangle = 0 \text{ für alle } y \in A.$$

c) Ist offensichtlich.

d) Aus $A \subset L(A) \subset \overline{L(A)}$ und Teil c) folgt

$$\overline{L(A)}^{\perp} \subset L(A)^{\perp} \subset A^{\perp}.$$

Sei $x \in A^{\perp}$, $y \in \overline{L(A)}$. Dann existiert eine Folge

$$y_n \in L(A), \quad y_n = \sum_{j=1}^{N(n)} c_{n,j} a_{n,j} \quad (c_{n,j} \in \mathbb{K}, a_{n,j} \in A)$$

mit $y_n \to y$. Damit folgt (vgl. Satz 9.8)

$$\begin{aligned} \langle x, y \rangle &= \lim_n \langle x, y_n \rangle = \lim_n \langle x, \sum_{j=1}^{N(n)} c_{n,j} a_{n,j} \rangle \\ &= \lim_n \sum_{j=1}^{n(n)} c_{n,j} \langle x, a_{n,j} \rangle = 0, \end{aligned}$$

also $x \in \overline{L(A)}^{\perp}$. ■

Satz 10.11 (Approximationssatz) *Sei X ein Hilbertraum, $A \neq \emptyset$ eine abgeschlossene* konvexe *Teilmenge von X (d. h. mit $x, y \in A$ liegt auch die Verbindungsstrecke $\{ax + (1-a)y : a \in [0,1]\}$ in A), z. B. ein abgeschlossener Teilraum. Dann gilt: Für jedes $x \in X$ existiert genau ein $y \in A$ mit*

$$\|x - y\| = d = d(x, A) := \inf \Big\{ \|x - z\| : z \in A \Big\}.$$

Beweis. Existenz: Sei (y_n) eine Folge aus A mit $\|x - y_n\| \to d(x, A)$. Ersetzt man in der Parallelogrammidentität x durch $y_n - x, y$ durch $y_m - x$, so folgt

$$\Big\| 2x - (y_n + y_m) \Big\|^2 + \|y_n - y_m\|^2 = 2\Big\{ \|y_n - x\|^2 + \|y_m - x\|^2 \Big\},$$

also, wegen $\frac{1}{2}(y_n + y_m) \in A$ und $\|y_n - x\| \to d$,

$$\begin{aligned} \|y_n - y_m\|^2 &= 2\Big\{ \|y_n - x\|^2 + \|y_m - x\|^2 \Big\} - 4\Big\| x - \frac{1}{2}(y_n + y_m) \Big\|^2 \\ &\leq 2\Big\{ \|y_n - x\|^2 + \|y_m - x\|^2 \Big\} - 4d^2 \leq \varepsilon \text{ für } n, m \geq N(\varepsilon). \end{aligned}$$

Also ist (y_n) eine Cauchyfolge, d. h. es gibt ein $y \in A$ mit $y_n \to y$ (da A abgeschlossene Teilmenge eines vollständigen Raumes ist) und

$$\|x - y\| = \lim_{n\to\infty} \|x - y_n\| = d.$$

Es bleibt die *Eindeutigkeit* von y zu beweisen. Sind y_1, y_2 Punkte dieser Art, so zeigt die obige Überlegung, daß $(x_n) = (y_1, y_2, y_1, y_2, \ldots)$ eine Cauchyfolge ist, also $y_1 = y_2$. ■

Satz 10.12 (Projektionssatz) *Sei X ein Hilbertraum, T ein abgeschlossener Teilraum von X.*

(a) Jedes $x \in X$ läßt sich eindeutig in der Form $x = y + z$ darstellen mit $y \in T$, $z \in T^\perp$.

(b) Es gilt $T^{\perp\perp} = T$.

Die Abbildung $P : X \to T$, $x \mapsto y$ (mit $x = y + z$, $y \in T$, $z \in T^\perp$) heißt die *orthogonale Projektion auf T*, y heißt die *Projektion* von x auf T. Es ist leicht zu sehen, daß P ein linearer Operator ist mit $\|P\| \leq 1$ ($\|P\| = 1$, falls $T \neq \{0\}$). Teil a des Satzes besagt, daß $X = T + T^\perp$ gilt.

Beweis. a) Da T abgeschlossen und konvex ist, existiert nach Satz 9.11 ein $y \in T$ mit $\|x - y\| = d(x, T)$. Sei $z := x - y$.

$\mathbf{z} \in \mathbf{T}^\perp$: Es ist $\langle u, z\rangle = 0$ zu zeigen für alle $u \in T$. Für $u = 0$ ist dies klar. Sei also $u \in T\backslash\{0\}$. Wegen $y + au \in T$ für alle $a \in \mathbb{K}$ gilt

$$\begin{aligned} d^2 = d(x,T)^2 &\leq \|x - (y + au)\|^2 = \|z - au\|^2 \\ &= \|z\|^2 - 2\,\mathrm{Re}\,\{a\langle z, u\rangle\} + |a|^2\|u\|^2 \\ &= d^2 - 2\,\mathrm{Re}\,\{a\langle z, u\rangle\} + |a|^2\|u\|^2. \end{aligned}$$

Mit $a = \|u\|^{-2}\langle u, z\rangle$ folgt hieraus

$$\|u\|^{-2}|\langle u, z\rangle|^2 \leq 0, \text{ d. h. } \langle u, z\rangle = 0.$$

Damit ist die Existenz einer Darstellung der gewünschten Form gezeigt.

Eindeutigkeit: Sei

$$x = y + z = y' + z' \text{ mit } y, y' \in T,\ z, z' \in T^\perp.$$

Dann gilt

$$y - y' = z' - z \in T \cap T^{\perp} = \{0\},$$

also $y = y'$, $z = z'$.

b) $\mathbf{T} \subset \mathbf{T}^{\perp\perp}$: Ist $x \in T$, so gilt $\langle x, z\rangle = 0$ für alle $z \in T^{\perp}$, also $x \in T^{\perp\perp}$.

$\mathbf{T}^{\perp\perp} \subset \mathbf{T}$: Sei $x \in T^{\perp\perp}$. Nach Teil a gilt $x = y + z$ mit $y \in T$, $z \in T^{\perp}$, also (wegen $y \in T \subset T^{\perp\perp}$)

$$z = x - y \in T^{\perp} \cap T^{\perp\perp} = \{0\},$$

und somit $x = y \in T$. ■

Satz 10.13 *Sei X ein Hilbertraum, A eine Teilung von X.*

(a) Es gilt $A^{\perp\perp} = \overline{L(A)}$.

(b) A ist genau dann total, wenn gilt $A^{\perp} = \{0\}$.

Beweis. a) Aus $A^{\perp} = \overline{L(A)}^{\perp}$ folgt mit Teil b des Projektionssatzes $A^{\perp\perp} = \overline{L(A)}^{\perp\perp} = \overline{L(A)}$.

b) Aus $A^{\perp} = \{0\}$ und Teil a folgt $\overline{L(A)} = A^{\perp\perp} = \{0\}^{\perp} = X$. — Umgekehrt folgt aus $\overline{L(A)} = X$ zunächst $A^{\perp\perp} = X$ und wegen der Abgeschlossenheit des Teilraumes $A^{\perp}$ folgt hieraus $A^{\perp} = (A^{\perp})^{\perp\perp} = X^{\perp} = \{0\}$. ■

10.3 Orthonormalsysteme und Orthonormalbasen

Sei $(X, \langle\cdot,\cdot\rangle)$ ein Prä–Hilbertraum. Eine Familie $\{x_\alpha : \alpha \in A\}$ heißt ein *Orthonormalsystem* (ONS), wenn für alle $\alpha, \beta \in A$ gilt $\langle x_\alpha, x_\beta\rangle = \delta_{\alpha\beta}$. Ein Orthonormalsystem heißt eine *Orthonormalbasis* (ONB), wenn es in X total ist.

Satz 10.14 *Sei X ein Prä–Hilbertraum.*

(a) Jedes Orthonormalsystem ist linear unabhängig.

(b) Jede Orthonormalbasis ist ein maximales Orthonormalsystem.

(c) Ist X ein Hilbertraum, so ist jedes maximale Orthonormalsystem eine Orthonormalbasis.

Beweis. a) Aus $0 = \sum c_\alpha x_\alpha$ folgt

$$0 = \| \sum c_\alpha x_\alpha \|^2 = \sum |c_\alpha|^2, \text{ also } c_\alpha = 0 \text{ für alle } \alpha.$$

b) Da $\{x_\alpha : \alpha \in A\}$ total ist, gilt $\{x_\alpha : \alpha \in A\}^\perp = \{0\}$, d. h. es gibt kein Element der Norm 1, das zu allen x_α orthogonal ist.

c) Sei $\{x_\alpha : \alpha \in A\}$ ein maximales ONS. Wir nehmen an, daß es keine ONB ist, also *nicht* total. Dann ist $\overline{L\{x_\alpha : \alpha \in A\}} \neq X$ und somit nach dem Projektionssatz $(L\{x_\alpha : \alpha \in A\}^\perp \neq \{0\})$. Es existiert also ein $(e \in L\{x_\alpha : \alpha \in A\}^\perp)$ mit $\|e\| = 1$, d. h. $\{x_\alpha : \alpha \in A\}$ ist *nicht* maximal. ■

Beispiel 10.15 In ℓ_2 sei $e_n := (\delta_{nj})_{j \in \mathbb{N}}$ der *n–te Einheitsvektor*. Dann ist offenbar $\{e_n : n \in \mathbb{N}\}$ eine Orthonormalbasis. □

Beispiel 10.16 In $L_2(0,1)$ sei $e_n(x) := \exp(2\pi i n x) (n \in \mathbb{Z})$, dann ist $\{e_n : n \in \mathbb{Z}\}$ eine Orthonormalbasis. Das kann man z. B. folgendermaßen einsehen: Der Raum $\tilde{C}[0,1]$ der auf $[0,1]$ stetigen Funktionen mit $f(0) = f(1)$ ist dicht in $L_2(0,1)$. Es ist $\overline{L\{e_n : n \in \mathbb{Z}\}} \supset \tilde{C}[0,1]$ (vgl. z. B. O. Forster, Analysis I, §23 oder Korollar 11.3 in diesem Skript) und somit

$$\overline{L\{e_n : n \in \mathbb{Z}\}} = \overline{\tilde{C}[0,1]} = L_2(0,1).$$

□

Der folgende Satz zeigt, wie man auf einfache Weise Orthonormalsysteme und Orthonormalbasen konstruieren kann.

Satz 10.17 (Schmidtsches Orthonormalisierungsverfahren) *Sei X ein Prä–Hilbertraum,*

$$F = \{x_n : n \in \mathbb{N}\} \textit{ oder } F = \{x_n : n \in \{1, \ldots, m\}\}.$$

Dann existiert ein Orthonormalsystem $M = \{e_n\}$ mit $L(F) = L(M)$. Ist F linear unabhängig, so kann M so gewählt werden, daß gilt

$$L\{x_1, \ldots, x_n\} = L\{e_1, \ldots, e_n\} \textit{ für alle } n.$$

Beweis. Es genügt, die zweite Behauptung zu zeigen. Es ist leicht nachzuprüfen, daß diese gilt mit:

$$\begin{aligned} e_1 &:= \|x_1\|^{-1}x_1, \\ e_2 &:= \|x_2 - \langle e_1, x_2\rangle e_1\|^{-1}(x_2 - \langle e_1, x_2\rangle e_1), \\ &\vdots \\ e_{n+1} &:= \|x_{n+1} - \sum_{j=1}^{n}\langle e_j, x_{n+1}\rangle e_j\|^{-1}(x_{n+1} - \sum_{j=1}^{n}\langle e_j, x_{n+1}\rangle e_j). \end{aligned}$$

Details können dem Leser überlassen werden. ■

Satz 10.18 (Existenz von Orthonormalbasen) *(a) Jeder separable Prä–Hilbertraum besitzt eine abzählbare Orthonormalbasis.*

(b) Jeder Hilbertraum besitzt eine Orthonormalbasis.

Beweis. a) Da X separabel ist, gibt es eine abzählbare totale Menge $\{x_n\}$ in X; o. E. können wir annehmen, daß $\{x_n\}$ linear unabhängig ist. Anwendung des Schmidt'schen Orthonormalisierungsverfahrens auf $\{x_n\}$ liefert eine ONB $\{e_n\}$.

Teil b) soll hier nicht bewiesen werden; vgl. z. B. J. Weidmann, [11] Satz 1.57. Man benötigt hier wieder das Zorn'sche Lemma. ■

Satz 10.19 (Entwicklungssatz) *(a) Sei X ein Prä–Hilbertraum, $\{e_n\}$ ein Orthonormalsystem in X. Dann gilt für jedes $x \in X$*

$$\sum_n |\langle e_n, x\rangle|^2 \le \|x\|^2 \qquad \text{(Besselsche Ungleichung).}$$

(b) Ist $\{e_n\}$ eine Orthonormalbasis, so gilt für jedes $x \in X$

$$\sum_n |\langle e_n, x\rangle|^2 = \|x\|^2 \qquad \text{(Parsevalsche Gleichung),}$$

$$x = \sum_n \langle e_n, x\rangle e_n = \lim_m \sum_{n=1}^{m} \langle e_n, x\rangle e_n.$$

(c) Ist X ein Hilbertraum (Prä–Hilbertraum) und $\{e_n\}$ ein Orthonormalsystem in X, so ist $\sum c_n e_n$ genau dann konvergent (eine Cauchy–Folge), wenn $(c_n) \in \ell_2$ gilt. Ist $f = \sum c_n e_n$, so gilt $c_n = \langle e_n, f\rangle$.

(d) Sei X ein Hilbertraum, $\{e_n\}$ ein Orthonormalsystem in X. Dann ist

$$\sum_n \langle e_n, x\rangle e_n = \textit{ Projektion von } x \textit{ auf } \overline{L\{e_n\}}.$$

Beweis. a) Für jedes $y = \sum_{n=1}^m c_n e_n \in L\{e_n\}$ gilt

$$\begin{aligned} \|x-y\|^2 &= \|x\|^2 - 2\,\mathrm{Re} \sum_{n=1}^m c_n \langle x, e_n\rangle + \sum_{n=1}^m |c_n|^2 \\ &= \|x\|^2 - \sum_{n=1}^m |\langle e_n, x\rangle|^2 + \sum_{n=1}^m |c_n - \langle e_n, x\rangle|^2 \\ &\geq \|x\|^2 - \sum_{n=1}^m |\langle e_n, x\rangle|^2. \end{aligned} \qquad (*)$$

Bei festem m ist also $\|x-y\|$ am kleinsten für $c_n = \langle e_n, x\rangle$,

$$\|x-y\|^2 = \|x\|^2 - \sum_{n=1}^m |\langle e_n, x\rangle|^2,$$
$$\sum_{n=1}^m |\langle e_n, x\rangle|^2 = \|x\|^2 - \|x-y\|^2 \leq \|x\|^2.$$

Für $m \to \infty$ folgt die Bessel'sche Ungleichung.

b) Ist $\{e_n\}$ eine ONB, so gibt es zu jedem $\varepsilon > 0$ ein $m(\varepsilon)$ und $c_{\varepsilon,1}, \ldots, c_{\varepsilon,m(\varepsilon)}$ mit

$$\|x-y\| \leq \varepsilon \text{ für } y = \sum_{n=1}^{m(\varepsilon)} c_{\varepsilon,n} e_n.$$

Damit folgt aus $(*)$

$$\|x\|^2 - \sum_{n=1}^{m(\varepsilon)} |\langle e_n, x\rangle|^2 \leq \|x-y\|^2 \leq \varepsilon^2,$$

und somit die Parseval'sche Gleichung.

Für $y_m = \sum_{n=1}^m \langle e_n, x\rangle e_n$ folgt wegen der Parseval'schen Gleichung aus $(*)$

$$\|x-y_m\|^2 = \|x\|^2 - \sum_{n=1}^m \||\langle e_n, x\rangle|^2 \to 0, \text{ also } x = \sum_n \langle e_n, x\rangle e_n.$$

c) Ist offensichtlich wegen

$$\|\sum_{n=1}^m c_n e_n - \sum_{n=1}^p c_n e_n\|^2 = \sum_{n=p+1}^m |c_n|^2 \text{ für } p < m.$$

d) Mit $y = \sum_{n=1}^{\infty} \langle e_n, x\rangle e_n$ erhält man aus $(*)$ für jedes $m \in \mathbb{N}$ und beliebige $c_j \in \mathbb{K}$

$$\begin{aligned}\|x-y\|^2 &= \|x\|^2 - \sum_{n=1}^{\infty} |\langle e_n, x\rangle|^2 \le \|x\|^2 - \sum_{n=1}^{m} |\langle e_n, x\rangle|^2 \\ &\le \|x - \sum_{n=1}^{m} c_n e_n\|^2.\end{aligned}$$

Also ist

$$\|x-y\|^2 = \inf\left\{\|x-z\|^2 : z \in L\{e_n\}\right\} = \inf\left\{\|x-z\|^2 : z \in \overline{L\{e_n\}}\right\},$$

d. h. y ist die Projektion von x auf $\overline{L\{e_n\}}$. ■

10.4 Der Rieszsche Darstellungssatz und der Satz von Radon-Nikodym

Für Hilberträume kann der topologische Dualraum besonders leicht beschrieben werden:

Satz 10.20 (Rieszscher Darstellungssatz) *Sei $(X, \langle\cdot,\cdot\rangle)$ ein Hilberraum. Jedes $y \in X$ erzeugt durch $F_y(x) := \langle y, x\rangle$ für $x \in X$ ein stetiges lineares Funktional F_y mit $\|F_y\| = \|y\|$. Diese Abbildung $y \mapsto F_y$ von X nach X^* ist* anti-linear *(d. h. $F_{ay+bz} = \overline{a}F_y + \overline{b}F_z$) und bijektiv, d. h. jedes stetige lineare Funktional läßt sich eindeutig auf diese Weise darstellen.*

Beweis. Offensichtlich ist F_y für jedes $y \in X$ ein stetiges lineares Funktional mit $|F_y(x)| = |\langle y, x\rangle| \le \|y\|\,\|x\|$ und somit $\|F_y\| \le \|y\|$; wegen $F_y(y) = \|y\|^2$ ist $\|F_y\| = \|y\|$. Also ist $y \mapsto F_y$ injektiv.

Die Antilinearität der Abbildung $y \mapsto F_y$ folgt aus

$$\begin{aligned}F_{ay+bz}(x) &= \langle ay+bz, x\rangle = \overline{a}\langle y, x\rangle + \overline{b}\langle z, x\rangle \\ &= \overline{a}F_y(x) + \overline{b}F_z(x) = (\overline{a}F_y + \overline{b}F_z)(x).\end{aligned}$$

Es bleibt die Surjektivität zu beweisen: Sei $F \in X^*$; es ist ein $y \in X$ zu finden mit $F_y = F$. Für $F = 0$ kann $y = 0$ ewählt werden. Sei also o. E. $F \ne 0$, d. h. der Nullraum

$$N(F) = \{x \in X : F(x) = 0\}$$

ist ein echter abgeschlossener Teilraum von X, also $N(F)^\perp \neq \{0\}$. Sei $z \in N(F)^\perp$ mit $\|z\| = 1$, $a := \overline{F(z)}$. Dann ist (nachrechnen!)

$$F(x)z - F(z)x \in N(F) \text{ für jedes } x \in X,$$

also (wegen $z \in N(F)^\perp$)

$$0 = \langle z, F(x)z - F(z)x \rangle = F(x) - \overline{a}\langle z, x \rangle$$

und somit für alle $x \in X$

$$F(x) = \langle az, x \rangle = F_{az}(x) = F_y(x) \text{ mit } y := az.$$

■

Anmerkung. Es scheint zunächst überraschend, daß dieses beliebig aus $N(F)^\perp$ gewählte Element z bis auf einen Faktor das ist, durch das F erzeugt wird. Nachträglich wird klar, daß $N(F)^\perp = \{z\}^{\perp\perp} = L\{z\}$ eindimensional ist. Das kann man auch aus Aufgabe 3.7a ableiten.

Als Anwendung des Rieszschen Darstellungssatzes beweisen wir nun noch einen außerordentlich wichtigen Satz aus der Maßtheorie. Seien μ und ν Maße auf X. ν heißt *absolut stetig bezüglich* μ (oder: μ*–absolut stetig*, $\nu \ll \mu$), wenn gilt:

jede μ–meßbare Menge ist auch ν–meßbar,
jede μ–Nullmenge ist auch ν–Nullmenge.

Satz 10.21 (Satz von Radon–Nikodym) *Seien μ und ν Maße auf X, μ sei σ–endlich, ν sei μ–absolut stetig mit $\nu(X) < \infty$. Dann existiert ein $h \in L_1(X, \mu)$ mit*

$$\nu(A) = \int \chi_A h \,\mathrm{d}\mu \text{ für jede } \mu\text{–meßbare Teilmenge } A;$$

h heißt die Dichte *von ν bezüglich μ.*

Beweis. Wir betrachten nur den Fall $\mu(X) < \infty$ (der allgemeine Fall ergibt sich leicht auf Grund der σ–Endlichkeit von μ durch Zusammensetzen).

Wir betrachten $L_2(X, \tau)$ mit $\tau := \mu + \nu$. Die Abbildung

$$L_2(X, \tau) \to \mathbb{K}, \quad f \mapsto \int f \, \mathrm{d}\nu$$

ist ein stetiges lineares Funktional, denn

$$\Big| \int f \, \mathrm{d}\nu \Big| \leq \int |f| \, \mathrm{d}\tau = \int 1 \cdot |f| \, \mathrm{d}\tau \leq \tau(X)^{1/2} \|f\|_2.$$

Nach dem Riesz'schen Darstellungssatz existiert also ein $g \in L_2(X, \tau)$ mit

$$\int f \, \mathrm{d}\nu = \int gf \, \mathrm{d}\tau \text{ für alle } f \in L_2(X, \tau). \tag{1}$$

Insbesondere gilt wegen $\chi_A = \nu(A) \leq \tau(A) \in L_2(X, \tau)$ für jede μ–meßbare Teilmenge A von X

$$\nu(A) = \int \chi_A \, \mathrm{d}\nu = \int g\chi_A \, \mathrm{d}\tau = \int_A g \, \mathrm{d}\tau.$$

Aus $0 \leq \nu(A) \leq \tau(A)$ folgt $0 \leq \displaystyle\int_A g \, \mathrm{d}\tau$ für jede μ–meßbare Teilmenge A von X. Daraus ergibt sich

$$0 \leq g(x) \leq 1 \text{ für } \tau\text{–f. a.} \quad x \in X.$$

(Wäre z. B. $g(x) < 0$ auf einer Menge M mit $\tau(M) > 0$, so wäre $\nu(M) = \int_M g \, \mathrm{d}\tau < 0$. Wäre $g(x) > 1$ auf eine Menge M mit $\tau(M) > 0$, so wäre $\nu(M) = \int_M g \, \mathrm{d}\tau > \tau(M)$.)

Aus (1) folgt

$$\int (1-g)f \, \mathrm{d}\nu = \int gf \, \mathrm{d}\tau - \int gf \, \mathrm{d}\nu = \int gf \, \mathrm{d}\mu. \tag{2}$$

Sei

$$N := \{x \in X : g(x) = 1\}, \qquad L := X \backslash N.$$

Dann gilt wegen (2)

$$\mu(N) = \int \chi_N \, \mathrm{d}\mu = \int g\chi_N \, \mathrm{d}\mu = \int (1-g)\chi_N \, \mathrm{d}\nu = 0,$$

also auch $\nu(N) = 0$, da ν μ–absolut stetig ist.

Mit $f := (1 + g + g^2 + \ldots + g^n)\chi_A$ folgt aus (2)

$$\int_A (1 - g^{n+1})\,\mathrm{d}\nu = \int_A (1-g)\,(1 + g + \ldots + g^n)\,\mathrm{d}\nu = \int_A g(1 + g + \ldots + g^n)\,\mathrm{d}\mu.$$

Die Folgen der Integranden auf beiden Seiten sind nicht–fallend; die Integralfolgen sind beschränkt durch $\int_A 1\,\mathrm{d}\,\nu = \nu(A)$. Der linke Integrand konvergiert gegen χ_L; also konvergiert nach dem Satz von B. Levi (Satz 5.9) die linke Seite gegen $\nu(A \cap L)$. Für die rechte Seite liefert der Satz von B. Levi die Existenz einer Funktion $h \in L_1(X, \mu)$ mit

$$\begin{aligned} &g(1 + g + \ldots + g^n) \to h \qquad \mu\text{–f. ü.}, \\ &\int_A g(1 + g + \ldots + g^n)\,\mathrm{d}\mu \to \int_A h\,\mathrm{d}\mu \text{ für } A \subset X\ \mu\text{–meßbar}. \end{aligned}$$

Insgesamt folgt (da N eine ν–Nullmenge ist)

$$\nu(A) = \nu(L \cap A) + \nu(N \cap A) = \nu(L \cap A) = \int_A h\,\mathrm{d}\mu.$$

■

Bemerkung. Im Satz von Radon–Nikodym ist die Voraussetzung der σ–Endlichkeit von μ wichtig. Sei $X = [0, 1]$ und μ das Zählmaß eingeschränkt auf die σ–Algebra der Lebesgue–meßbaren Mengen; ν sei das Lebesgue–Maß auf $[0, 1]$. Da nur die leere Menge eine μ–Nullmenge ist, ist offenbar ν absolut stetig bezüglich μ. Wir zeigen, daß ν keine Dichte bezüglich μ besitzt. Für eine solche Dichte h müßte gelten: Entweder

(α) $h \equiv 0$, also $\nu(A) = 0$ für jede ν–meßbare Teilmenge von $[0, 1]$,

oder

(β) es gibt ein $x_0 \in [0, 1]$ mit $h(x_0) \neq 0$, also

$$\nu(\{x_0\}) = h(x_0) \neq 0.$$

Beide Möglichkeiten liefern einen Widerspruch.

10.5 Übungsaufgaben

10.1 Man beweise Teil b von Satz 9.2: In der Schwarz'schen Ungleichung gilt genau dann das Gleichheitszeichen, wenn x und y positiv linear abhängig sind.

10.2 Sei $W_{2,1}(0,1) := \Big\{f : [0,1] \to \mathbb{C},\ f \text{ absolut stetig},\ f' \in L_2(0,1)\Big\}$.

(a) Durch $\langle f,g\rangle_1 := \displaystyle\int_0^1 \{\overline{f}g + \overline{f'}g'\}\,\mathrm{d}x$ ist ein Skalarprodukt auf $W_{2,1}(0,1)$ erklärt.

(b) $(W_{2,1}(0,1),\ \langle\cdot,\cdot\rangle_1)$ ist ein Hilbertraum.

(c) Für jedes $x_0 \in [0,1]$ ist $F_{x_0}(f) = f(x_0)$ ein stetiges lineares Funktional auf $W_{2,1}(0,1)$; es gibt also genau ein $g \in W_{2,1}(0,1)$ mit $f(x_0) = \langle g,f\rangle_1$ für alle $f \in W_{2,1}(0,1)$.

10.3 Man zeige, daß in ℓ_p und $L_p(X,\mu)$ für $p \neq 2$ die Parallelogrammidentität nicht gilt (d. h. die Normen werden nicht durch Skalarprodukte erzeugt).

10.4 In $L_2(-1,1)$ sei

$$T := \Big\{f \in L_2(-1,1) :\ f(x) = f(-x) \text{ für f. a. } x\Big\},$$

der Teilraum der *geraden Funktionen*. Man bestimme $T^\perp$.

10.5 Man zeige: Die Eindeutigkeitsaussage des Approximationssatzes (Satz 9.11) gilt nicht in $\ell_1,\ \ell_\infty,\ L_1(X,\mu)$ und $L_\infty(X,\mu)$.

10.6 Man wende das Schmidt'sche Orthonormalisierungsverfahren auf die Elemente $1, x, x^2$ in $L_2(-1,1)$ an.

10.7 (a) Sei X ein Hilbertraum, Y ein Banachraum, T_0 ein beschränkter (nicht auf ganz X definierter) linearer Operaotor von X nach Y. Dann gibt es eine Fortsetzung $T \in B(X,Y)$ von T_0 mit $\|T\| \equiv \|T_0\|$.
Anleitung: Ist T' die eindeutige Fortsetzung auf $\overline{D(T_0)}$ (vgl. Satz 7.2) und P die orthogonale Projektion auf $\overline{D(T_0)}$, so hat $T = T_0'P$ die gewünschte Eigenschaft.

(b) Man zeige, daß es in Teil a genügt, wenn X ein Prä–Hilbertraum ist.

(c) Ist $\overline{D(T_0)} \neq X$, so ist die Fortsetzung aus Teil a genau dann eindeutig, wenn $\|T_0 x\| = \|T_0\|\,\|x\|$ für alle $x \in D(T_0)$ gilt.

10.8 (a) In einem Hilbertraum X enthält jede beschränkte Folge (x_n) eine schwach konvergente Teilfolge.
Anleitung: Mit Hilfe von Satz 7.4 c zeigt man, daß es genügt, eine Teilfolge (z_k) von (x_n) zu finden so, daß für jedes $n \in \mathbb{N}$ die Folge $(\langle z_k, x_n\rangle)_{k\in\mathbb{N}}$ konvergent ist; die Konstruktion einer solchen Teilfolge gelingt mit dem Diagonalfolgenverfahren.

(b) Die abgeschlossene Einheitskugel in X ist folgenkompakt bezüglich der schwachen Topologie (*schwach folgenkompakt*).

10.9 Sind $\{f_j : j \in A\}$ und $\{g_k : k \in B\}$ Orthonormalbasen in $L_2(X,\mu)$ bzw. $L_2(Y,\nu)$, so ist $\{f_j(x)g_k(y) : j \in A,\ k \in B\}$ eine Orthonormalbsis in $L_2(X \times Y,\ \mu \times \nu)$.

10.10 Sei μ ein σ–endliches Maß auf X, ν ein μ–absolut stetiges Maß, $\nu(A) = \int_A h(x)\,\mathrm{d}\,\mu(x)$ für jede μ–meßbare Teilmenge A von X. Ist $f \in L_1(X,\nu)$ μ–meßbar, so ist $fh \in L_1(X,\mu)$ und es gilt $\displaystyle\int f(x)\,\mathrm{d}\nu(x) = \int f(x)h(x)\,\mathrm{d}\mu(x)$.
Anleitung: Es genügt $f \geq 0$ zu betrachten. Man zeige die Aussage zunächst für

$$f_n = \sum_{k=1}^{2^{2n}} (k-1)2^{-n}\chi_{M_{n,k}} \text{ mit } M_{n,k} = \{x \in X : (k-1)2^{-n} \leq f(x) \leq k2^{-n}\}$$

und benutze den Satz von B. Levi.

10.11 Die Funktion $f : [a,b] \to \mathbb{R}$ sei *Lipschitz–stetig* d. h. es gibt ein L mit $|f(x) - f(y)| \leq L|x-y|$ für alle $x, y \in [a,b]$.
Man zeige: f ist absolut stetig, d. h. es gibt ein $g \in L_1(a,b)$ mit $f(x) = f(a) + \displaystyle\int_a^x g(t)\,\mathrm{d}t$.
Anleitung: a) $\varphi(x) := f(x) + Lx$ ist nicht–fallend und Lipschitz–stetig mit Konstante $2L$. b) Das durch φ gemäß Beispiel 5.2 erzeugte Lebesgue–Stieltjes–Maß ist absolut stetig bezüglich dem Lebesgue-Maß.

11 Die Dualräume von ℓ_p und $L_p(X,\mu)$ für $1 \leq p < \infty$

11.1 Die Dualräume von ℓ_p für $1 \leq p < \infty$

Wir wissen aus Satz 7.7, daß jedes $f = (f_n) \in \ell_q$ ($\frac{1}{p} + \frac{1}{q} = 1$ für $1 < p < \infty,\ q = \infty$ für $p = 1,\ q = 1$ für $p = \infty$) ein stetiges lineares Funktional auf ℓ_p

$$F_f(g) = \sum_n f_n g_n \text{ für } g = (g_n) \in \ell_p$$

erzeugt mit $\|F_f\| = \|f\|_p$. Der folgende Satz besagt, daß für $1 \leq p < \infty$ jedes $F \in \ell_p^*$ auf diese Weise erzeugt wird, d. h. ℓ_p^* und ℓ_q sind *isomorph*, $\ell_p^* \cong \ell_q$. Nach Abschnitt 7.4 kann man ℓ_p als $l_p(\mathbb{N},\nu)$ mit dem Zählmaß ν auf $\mathbb{N}$ auffassen. In diesem Sinn ist Satz 11.1 ein Spezialfall des wesentlich allgemeineren Satzen 11.7.

Satz 11.1 *Sei $1 \leq p < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$ ($q = \infty$ für $p = 1$). Dann existiert zu jedem stetigen linearen Funktional $F \in \ell_p^*$ genau ein $f \in \ell_q$ mit $F = F_f$. (Eine entsprechende Aussage für $p = \infty$ gilt nicht, vgl. Aufgabe 4.6.)*

Beweis. Die Eindeutigkeit folgt aus

$$\|F_{f_1} - F_{f_2}\| = \|f_1 - f_2\|_q.$$

Sei $F \in \ell_p^*$, $f_n := F(e_n)$ mit $e_n = (\delta_{n,j})_{j\in\mathbb{N}}$ für $n \in \mathbb{N}$. Dann gilt

$$F(g) = \sum_n f_n g_n \text{ für alle } g \in \ell_0.$$

wobei $\ell_0 \subset \ell_p$ der Teilraum der endlichen (*finiten*) Folgen ist. Für jedes $g \in \ell_p$

$$g^{(k)} := (g_1, \ldots, g_k,\ 0,\ 0, \ldots) \to g \text{ für } k \to \infty.$$

Da F stetig ist, folgt hieraus

$$F(g) = \lim_{k\to\infty} F(g^{(k)}) = \lim_{k\to\infty} \sum_{n=1}^{k} f_n g_n = \sum_{n=1}^{\infty} f_n g_n.$$

Es bleibt zu zeigen, daß $f = (f_n) \in \ell_q$ gilt.

$\mathbf{p = 1}$ $(q = \infty)$: Es gilt

$$|f_n| = |F(e_n)| \le \|F\|\,\|e_n\|_p = \|F\| \text{ für alle } n \in \mathbb{N},$$

also $f \in \ell_\infty$.

$\mathbf{1 < p < \infty}$ $(1 < q < \infty)$: Sei $f^{(k)} := (f_1, \ldots f_k,\, 0, \ldots)$. Dann gilt $f^{(k)} \in \ell_0 \subset \ell_q$ und jedes $f^{(k)}$ erzeugt nach Satz 7.7 ein Funktional $F_k \in \ell_p^*$ mit $\|F_k\| = \|f^{(k)}\|_q$. Offenbar gilt

$$F_k(g) = F(P_k g) \text{ für alle } g \in \ell_p$$

mit dem linearen Operator (*Projektor* auf die ersten k Komponenten)

$$P_k : \ell_p \to \ell_p, P(g_n) = (g_1, \ldots, g_k, 0, 0, \ldots),$$

für den offensichtlich $\|P_k\| \le 1$ gilt. Daraus folgt

$$\left\{ \sum_{j=1}^{k} |f_j|^q \right\}^{1/q} = \|F_k\| = \|F P_k\| \le \|F\| \text{ für alle } k \in \mathbb{N},$$

also $f \in \ell_q$. ■

Im obigen Beweis war es recht leicht, den Kandidaten f für das Element aus ℓ_q zu finden, das das Funktional F erzeugt. Die einzige Hürde war der Beweis von $f \in \ell_q$. Für $L_p(X,\mu)$ wird es viel schwieriger sein, das erzeugende $f \in L_q(X,\mu)$ zu finden (Abschnitt 10.3). Wir benötigen einige Hilfsmittel.

11.2 Komplexe Maße

Sei $\mathcal{U}$ eine σ–Algebra in X (vgl. Abschnitt 5.5). Eine Abbildung $\nu : \mathcal{U} \to \mathbb{C}$ heißt ein *komplexes Maß* auf $\mathcal{U}$, wenn gilt:

$$\text{sind } A_n \in \mathcal{U}\, (n \in \mathbb{N}) \text{ disjunkt, so gilt } \nu(\cup A_n) = \sum \nu(A_n).$$

Bemerkung. Es sei darauf hingewiesen, daß nur Werte aus $\mathbb{C}$ zugelassen sind; das Maß ∞ gibt es hier nicht. In diesem Sinne sind die in §5 betrachteten Maße i. allg. keine komplexen Maße. Da die Anordnung der A_n in $\cup A_n$ keine Rolle spielt, darf sie auch in der Summe keine Rolle spielen; die Summen sind also stets absolut konvergent.

Mit ν sind natürlich auch Re ν und Im ν,

$$(\mathrm{Re}\,\nu)\,(A) := \mathrm{Re}\,(\nu(A)), \qquad (\mathrm{Im}\,\nu)\,(A) := \mathrm{Im}\,(\nu(A)),$$

komplexe Maße.

Sei ν ein komplexes Maß auf $\mathcal{U}$. Die *totale Variation* $|\nu|$ von ν ist definiert durch

$$|\nu|(A) := \sup\left\{\sum_{j=1}^{\infty}|\nu(B_j)| : B_j \in \mathcal{U}\text{disjunkt},\ A = \bigcup_{j=1}^{\infty} B_j\right\} \text{ für alle } A \in \mathcal{U}.$$

Satz 11.2 *Ist ν ein komplexes Maß auf der σ–Algebra $\mathcal{U}$, so ist $|\nu|$ ein (positives) Maß auf $\mathcal{U}$.*

Beweis. Sei $A = \cup A_n$ mit disjunkten Mengen $A_n \in \mathcal{U}$; es ist $|\nu|(A) = \sum |\nu|(A_n)$ zu beweisen.

Für jedes $t_n < |\nu|(A_n)$ gibt es nach Definition von $|\nu|$ eine Darstellung $A_n = \cup_j B_{n,j}$ mit disjunkten Mengen $B_{n,j}$ aus $\mathcal{U}$ so, daß

$$t_n < \sum_j |\nu(B_{n,j})|$$

gilt. Da die $B_{n,j}$ in der Darstellung $A = \cup_{n,j} B_{n,j}$ disjunkt sind, , folgt (wiederum aus der Definition von $|\nu|$)

$$\sum_n t_n < \sum_{n,j} |\nu(B_{n,j})| \leq |\nu|(A).$$

Da dies für jede Zahl der $t_n < |\nu|(A_n)$ gilt, folgt

$$\sum_n |\nu|(A_n) \leq |\nu|(A).$$

Sei $A = \cup_j C_j$ mit disjunkten $C_j \in \mathcal{U}$. Dann sind auch $A_n = \cup_j (C_j \cap A_n)$ und $C_j = \cup_n (C_j \cap A_n)$ Darstellungen mit disjunkten Mengen, also

$$\begin{aligned}\sum_j |\nu(C_j)| &= \sum_j \Big|\sum_n \nu(C_j \cap A_n)\Big| \leq \sum_j \sum_n |\nu(C_j \cap A_n)| \\ &= \sum_n \sum_j |\nu(C_j \cap A_n)| \leq \sum_n |\nu|(A_n).\end{aligned}$$

Geht man auf der linken Seite zum Spremum über, so folgt

$$|\nu|(A) \leq \sum_n |\nu|(A_n).$$

Insgesamt ist dann die Gleichheit gezeigt. ■

Satz 11.3 *Für jedes komplexe Maß ν gilt $|\nu|(X) < \infty$. (Im konkreten Fall des im Beweis von Satz 11.7 benutzten Maßes ν ist diese Eigenschaft leichter direkt zu sehen.)*

Beweis. (i) Ist $E \in \mathcal{U}$ mit $|\nu|(E) = \infty$, so gibt es A, $B \in \mathcal{U}$ mit

$$E = A \cup B, A \cap B = \emptyset, |\nu(A)| > 1 \text{ und } |\nu(B)| = \infty.$$

Beweis von (i): Nach Definition von $|\nu|$ existiert für jedes $t < \infty$ eine Darstellung $E = \cup_j E_j$ mit disjunkten E_j so, daß $\sum_j |\nu(E_j)| > t$ gilt. Wir wählen

$$t > 6(1 + |\nu(E)|)$$

und $n \in \mathbb{N}$ so, daß

$$\sum_{j=1}^{n} |\nu(E_j)| > t.$$

Es ist leicht einzusehen (vgl. Aufgabe 10.2), daß es dann eine Teilmenge $S \subset \{1, \ldots, n\}$ gibt mit

$$\Big| \sum_{j \in S} \nu(E_j) \Big| > \frac{1}{6} t.$$

Sei $A := \cup_{j \in S} E_j$. Dann gilt $A \subset E$ und $|\nu(A)| > \dfrac{t}{6} \geq 1$. Mit $B := E \backslash A$ gilt

$$|\nu(B)| = |\nu(E) - \nu(A)| \geq |\nu(A)| - |\nu(E)| > \frac{t}{6} - |\nu(E)| > 1.$$

Aus $|\nu(E)| = |\nu|(A) + |\nu|(B)$ und $|\nu(E) = \infty$ folgt $|\nu|(A) = \infty$ oder $|\nu|(B) = \infty$. Damit folgt (i), eventuell durch Vertauschen von A und B.

(ii) Nehmen wir jetzt an, daß $|\nu|(X) = \infty$ gilt. Sei $B_0 := X$; $B_0 \supset \ldots \supset B_n$ seien so gewählt, daß $|\nu|(B_j) = \infty$ und $|\nu|(B_{j-1} \backslash B_j) > 1$ gilt für $j = 1, \ldots, n$. Dann läßt sich B_n nach (i) darstellen in der Form

$$B_n = A_{n+1} \cup B_{n+1} \text{ (disjunkt)}$$

mit $|\nu(A_{n+1})| > 1$ und $|\nu|(B_{n+1}) = \infty$. Wir erhalten durch diese Induktion eine disjunkte Folge (A_j) mit $|\nu(A_j)| > 1$. Für $C := \cup_j A_j$ ist also in

$$\nu(C) = \sum_j \nu(A_j)$$

die Reihe *nicht* absolut konvergent, ein Widerspruch zu obiger Bemerkung. ■

Satz 11.4 *Sei ν ein komplexes Maß auf X, μ ein positives Maß. Ist ν absolut stetig bezüglich μ, so ist auch $|\nu|$ absolut stetig bezüglich μ. (Die absolute Stetigkeit eines komplexen Maßes ist wie für positive Maße definiert.)*

Beweis. Jede μ–meßbare Menge ist ν–meßbar und somit auch $|\nu|$–meßbar. Gilt $\mu(N) = 0$, so gilt $\mu(M) = 0$ für jede (meßbare) Teilmenge M von N, also auch $\nu(M) = 0$. Damit folgt $|\nu|(N) = 0$ aus der Definition von $|\nu|$. ■

Satz 11.5 (Komplexe Version des Satzes von Radon–Nikodym) *Sei μ ein positives Maß auf X mit $\mu(X) < \infty$, ν ein bezüglich μ absolut stetiges komplexes Maß. Dann existiert ein $h \in L_1(X, \mu)$ mit*

$$\nu(A) = \int_A h(x)\,\mathrm{d}\mu(x) \text{ für jede } \mu\text{–meßbare Menge } A \subset X.$$

Beweis. Wir zerlegen ν in Real– und Imaginärteil $\nu = \sigma + i\tau$, σ und τ sind ebenfalls μ–absolut stetig. Dann sind offenbar

$$\sigma_\pm := \frac{1}{2}(|\sigma| \pm \sigma) \text{ und } \tau_\pm := \frac{1}{2}(|\tau| \pm \tau)$$

ebenfalls μ–absolut stetige positive Maße, $\sigma = \sigma_+ - \sigma_-, \tau = \tau_+ - \tau_-$. Nach dem Satz von Radon–Nikodym (Satz 10.21) gibt es also $s_\pm, t_\pm \in L_1(X,\mu)$ mit

$$\sigma_\pm(A) = \int\limits_A s_\pm(x)\,\mathrm{d}\mu(x), \quad \tau_\pm(A) = \int\limits_A t_\pm(s)\,\mathrm{d}\mu(X),$$

also

$$\nu(A) = \int\limits_A h(x)\,\mathrm{d}\mu(x) \text{ mit } h = s_+ - s_- + i(t_+ - t_-) \in L_1(X;\mu).$$

■

11.3 Die Dualräume von $L_p(X,\mu)$ für $1 \leq p < \infty$

Im Beweis des Hauptresultates dieses Abschnitts benötigen wir das folgende Hilfsmittel:

Satz 11.6 *Sei μ ein σ–endliches Maß auf $X, X = \cup X_n$ mit $\mu(X_n) < \infty$,*

$$\begin{aligned} L_0((X_n), M) \quad &= \quad \{g : X \to \mathbb{C},\ \mu\text{–meßbar},\ g \text{ beschränkt.} \\ &\qquad \text{es gibt ein } n \in \mathbb{N} \text{ mit } g(x) = 0 \text{ in } X \backslash X_n\}. \end{aligned}$$

Sei $1 \leq p < \infty$, $\dfrac{1}{p} + \dfrac{1}{q} = 1$ und $f : X \to \mathbb{C}$ μ–meßbar mit

$$fg \in L_1(X,\mu) \text{ und } \left|\int f(x)g(x)\,\mathrm{d}\mu(x)\right| \leq C\|g\|_p \text{ für alle } g \in L_0((X_n),\mu).$$

Dann ist $f \in \mathcal{L}_q(X,\mu)$ mit $\|f\|_q \leq C$. (Der Satz gilt auch für $p = \infty$, vgl. Aufg. 10.5.)

Beweis. $\mathbf{p = 1}$: Nehmen wir an, daß $\sup|f(x)| > C$ gilt. Dann existiert ein $\varepsilon > 0$ so, daß für $X_\varepsilon := \{x \in X : |f(x)| \geq C + \varepsilon\}$ gilt $\mu(X_\varepsilon) > 0$. Also existiert ein $n \in \mathbb{N}$ mit $\mu(X_\varepsilon \cap X_n) > 0$. Wählen wir $g := \overline{\operatorname{sgn} f}(\mu(X_\varepsilon \cap X_n))^{-1}\chi_{X_\varepsilon \cap X_n}$, so folgt $g \in L_0((X_n),\mu)$, $\|g\|_1 = 1$ und

$$\left|\int f(x)g(x)\,\mathrm{d}\mu(x)\right| \geq C + \varepsilon = (C+\varepsilon)\|g\|_1 > C\|g\|_1,$$

ein Widerspruch.

$\mathbf{1 < p < \infty}$: Sei 0. E. $f \not\equiv 0$. Für jedes $n \in \mathbb{N}$ sei

$$f_n(x) := \begin{cases} f(x) & \text{für } x \in X_n \text{ mit } |f(x)| \le n, \\ 0 & \text{sonst.} \end{cases}$$

Dann gilt $f_n \in L_q(X,\mu)$ und $\|f_n\|_q \to \left\{ \int |f|^q \, \mathrm{d}\mu \right\}^{1/q} \le \infty$, $|f_n|_q \ne 0$ für große n. Mit

$$g_n(x) := |f_n|_q^{-q/p} \overline{\operatorname{sgn} f_n(x)} |f_n(x)|^{q/p}$$

gilt dann $g_n \in L_p((X_n),\mu)$ und $\|g_n\|_p = 1$ sowie

$$\begin{aligned} C &= C\|g_n\|_p \ge |\int f(x) g_n(x) \, \mathrm{d}\mu(x)| = \int f_n(x) g_n(x) \, \mathrm{d}\mu(x) \\ &= \int |f_n(x)|^q \, \mathrm{d}\mu(x) \|f_n\|_q^{-q/p} = \|f_n\|_q^{q-q/p} = \|f_n\|_q \text{ für alle } n \in \mathbb{N}. \end{aligned}$$

hieraus folgt $f \in L_q(X,\mu)$ und $\|f\|_q \le C$. ■

Aus Satz 7.6 wissen wir für $1 \le p \le \infty$, $\dfrac{1}{p} + \dfrac{1}{q} = 1$, daß jedes $f \in L_q(X,\mu)$ ein stetiges lineares Funktional $F_f \in L_q(X,\mu)^*$

$$F_g(g) = \int f(x) g(x) \, \mathrm{d}\mu(x) \text{ für } f \in L_p(X,\mu)$$

erzeugt mit $\|F_f\| = \|f\|_q$. Der folgende Satz besagt, daß für $1 \le p < \infty$ jedes $F \in L_p(X,\mu)^*$ auf diese Weise erzeugt wird, d. h. $L_p(X,\mu)^*$ und $L_q(X,\mu)$ sind *isomorph*, $L_p(X,\mu)^* \cong L_q(x,\mu)$.

Satz 11.7 *Sei ein σ–endliches Maß auf X, $1 \le p < \infty$, $\dfrac{1}{p} + \dfrac{1}{q} = 1$ ($q = \infty$ für $p = 1$). Dann existiert zu jedem stetigen linearen Funktional $F \in L_p(X,\mu)^*$ genau ein $f \in L_q(X,\mu)$ mit $F = F_f$. (Der Satz gilt in der gleichen Form auch für nicht σ–endliche μ; vgl. N. Danford – J.T. Schwartz, Linear Operators I, Theorem IV. 8.1.)*

Beweis. Wieder ergibt sich die Eindeutigkeit aus $\|F_{f_1} - F_{f_2}\| = \|f_1 - f_2\|_q$.

Sei $X = \cup X_n$ mit $X_n \subset X_{n+1}$ und $\mu(X_n) < \infty$, $n \in \mathbb{N}$ im folgenden fest. Für jede μ–meßbare Menge $A \subset X_n$ ist $\chi_A \in L_p(X,\mu)$. Man muß also definieren.

$$\nu(A) := F(\chi_A) \text{ für } A \subset X_n \ \mu\text{–meßbar.}$$

ν ist ein komplexes Maß auf X_n: Ist $A \subset X_n$ μ–meßbar, $A = \cup_j A_j$ mit μ–meßbaren disjunkten Teilmengen A_j, so gilt [2]

$$\begin{aligned}
\nu(A) &= \nu(\bigcup_j A_j) = F(\chi_{\cup A_j}) = F(\sum_j \chi_{a_j}) \\
&= F(\|\cdot\|_p - \lim_N \sum_{j=1}^N \chi_{A_j}) = \lim_N \sum_{j=1}^N F(\chi_{A_j}) \\
&= \sum_{j=1}^\infty F(\chi_{A_j}) = \sum_{j=1}^\infty \nu(A_j).
\end{aligned}$$

Daß $|\nu|(X_n) < \infty$ gilt (vgl. Satz 11.3) sieht man in diesem konkreten Fall wie folgt: Für jede Darstellung $X_n = \cup A_j$ mit disjunkten Mengen A_j gilt $\sum |\nu(A_j)| = \sum |F(\chi_{A_j})| = \sum F(\chi_{A_j}) \overline{\operatorname{sgn} F(\chi_{A_j})} = F(\sum \overline{\operatorname{sgn} F(\chi_{A_j})} \chi_{A_j}) \le \|F\|(\mu(X_n))^{1/p}$ also $|\nu|(X_n) \le \|F\|(\mu(X_n))^{1/p}$

ν ist μ–absolut stetig: Ist $N \subset X_n$ mit $\mu(N) = 0$, so ist $\chi_N = 0$ in $L_p(X,\mu)$, also $\nu(N) = F(\chi_N) = 0$.

Aus dem komplexen Satz von Radon–Nikodym (Satz 11.5) ergibt sich die Existenz eines $f \in L_1(X_n,\mu)$ mit

$$F(\chi_A) = \nu(A) = \int_A f(x)\,\mathrm{d}\mu(x) = \int f(x)\chi_A(x)\,\mathrm{d}\mu(x)$$

für jede μ–meßbare Teilmenge A von X_n. Damit folgt

$$F(g) = \int f(x)g(x)\,\mathrm{d}\mu(x)$$

für jede *einfache Funktion* auf X_n (das sind μ–meßbare Funktionen, die nur endlich viele Werte annehmen; Linearkombinationen charakteristischer Funktionen μ–meßbarer Teilmengen X_n). Da jedes $g \in L_\infty(X,\mu)$ gleichmäßiger Limes von

[2]Diese Überlegung gilt nicht für $p = \infty$; dann ist ν nicht σ–additiv, sondern nur *endlich additiv*; $L_\infty(X;\mu)^*$ ist der Raum der *endlich additiven Maße*.

einfachen Funktionen auf X_n ist (und damit auch L_p–Limes, wegen $\mu(X_n) < \infty$), folgt

$$F(g) = \int f(x)g(x)\,\mathrm{d}\mu(x) \text{ für alle } g \in L_\infty(X_n,\mu).$$

Für jedes $n \in \mathbb{N}$ gewinnt man so eine L_1–Funktion f_n auf X_n. Die letzte Identität impliziert (vgl. Aufgabe 10.3)

$$\text{für } n \leq m \text{ gilt } f_n(x) = f_m(x)\text{–f. ü. in } X_n.$$

Es gibt also eine μ–meßbare Funktion $f : X \to \mathbb{C}$ mit $f(x) = f_n(x)$ μ–f. ü. in X_n. Es gilt insbesondere

$$|\int f(x)g(x)\,\mathrm{d}\mu(x)| = |F(g)| \leq \|F\|\,\|g\|_p \text{ für alle } g \in L_0(X,\mu).$$

Nach Satz 11.6 ist somit $f \in L_q(X,\mu)$ mit $\|f\|_q \leq \|F\|$. Deshalb kann die Darstellung von F durch f auf ganz $L_p(X,\mu)$ ausgedehnt werden. ■

Sei X ein normierter Raum, X^* sein topologischer Dualraum. Jedes $x \in X$ erzeugt durch

$$G_x(F) = F(x) \text{ für } F \in X^*$$

ein stetiges lineares Funktional auf X^*, d. h. ein Element aus $X^{**} = (X^*)^*$:

$$|G_x(F)| = |F(x)| \leq \|F\|\,\|x\|, \text{ d. h. } \|G_x\| \leq \|x\|.$$

Da es zu jedem x ein stetiges lineares Funktional F gibt mit $\|F\| = 1$ und $|F(x)| = \|x\|$ (Hahn–Banach–Fortsetzung des Funktionals $F_0 : L\{x\} \to \mathbb{K}$, $F_0(ax) = a\|x\|$, folgt sogar

$$\|G_x\| = \|x\|.$$

Die Einbettung $X \to X^{**}$, $x \mapsto G_x$ ist also isometrisch.

Ein normierter Raum heißt *reflexiv*, wenn diese Einbettung surjektiv (also ein Isomorphismus) ist. Da X^{**} stets ein Banachraum ist, ist das nur möglich, wenn X selbst ein Banachraum ist.

Die obigen Resultate besagen, daß ℓ_p und $L_p(X,\mu)$ für $1 < p < \infty$ reflexiv sind.

11.4 Übungsaufgaben

11.1 Sei X eine beliebige Menge (insbesondere auch überabzählbar), $\ell_p(X)$ sei die Menge der Funktionen $f: X \to \mathbb{K}$ mit
$p = \infty$: $\sup\limits_{x\in X} |f(x)| < \infty$,
$1 \leq p < \infty$: $f(x) = 0$ für alle bis aufzählbar viele $x \in X$, $\sum_{x\in X} |f(x)|^p < \infty$,

$$\|f\|_\infty = \sup_{x\in X} |f(x)|,\ \|f\|_p = \Big\{\sum_{x\in X} |f(x)|^p\Big\}^{1/2} \text{ für } 1 \leq p < \infty.$$

(a) Es gilt $\ell_p(X) \cong L_p(X,\nu)$, wenn ν das Zählmaß auf X ist; $\ell_p(X)$ sind Banachräume, $\ell_2(X)$ ist ein Hilbertraum.

(b) ist X nicht abzählbar, so sind alle $\ell_p(X)$ nicht separabel.

(c) Für $1 \leq p < \infty$ gilt im oben beschriebenen Sinn $\ell_p(X)^* \cong \ell_q(X)$ mit $\frac{1}{p}+\frac{1}{q} = 1$. (Dies ist nicht Satz 11.1 oder Satz 11.7 enthalten, da ν i. allg. nicht σ–endlich ist.)

11.2 Sei $\sum_{n\in M} |z_n| = C$ mit gewissen $z_n \in \mathbb{C}$ und $M \subset \mathbb{N}$ (endlich oder unendlich). Dann gibt es eine Teilmenge S von M mit $|\sum_{n\in S} z_n| \geq C/(4\sqrt{2})$.
Anleitung: Man zerlege $\mathbb{C}$ in die 4 Quadranten $Q_1,\dots,Q_4$. Dann gibt es mindestens ein $m \in \{1,\dots,4\}$ mit $\sum\{|z_n| : z_n \in Q_m\} \geq \frac{1}{4}C$. Für Zahlen w_n aus einem Quadranten gilt aber $|\sum w_n| \geq \sum |w_n|/\sqrt{2}$.

11.3 Sei μ ein Maß auf X. Ist $f \in L_1(X,\mu)$ und $\int_X f(x)g(x)\,\mathrm{d}\mu(x) = 0$ für alle $g \in L_\infty(X,\mu)$, so ist $f = 0$.

11.4 Jeder endlich–dimensionale normierte Raum ist reflexiv. (Dies gilt insbesondere auch für $\mathbb{K}^m$ mit den Normen $\|\cdot\|_1$ und $\|\cdot\|_\infty$, obwohl es für ℓ_1 und ℓ_∞ nicht gilt.)

11.5 Man beweise Satz 11.6 für $p = \infty$.

12 Anhang: Der Satz von Stone–Weierstraß

Satz 12.1 (Stone–Weierstraß) *Sei X ein kompakter Hausdorff–Raum, $C(X)$ der Raum der stetigen komplexwertigen Funktionen auf X ausgestattet mit der Supremumsnorm $\|\cdot\|_\infty$. Sei F eine Teilmenge von $C(X)$ mit den Eigenschaften*

(a) F enthält eine konstante Funktion ungleich 0 (z. B. die Funktion 1),

(b) F trennt die Punkte von X, d. h. zu $x_1, x_2 \in X$ mit $x_1 \neq x_2$ gibt es ein $f \in F$ mit $f(x_1) \neq f(x_2)$,

(c) mit jedem $f \in F$ liegt auch $\overline{f}$ $\quad(\overline{f}(x) := \overline{f(x)})$ in F (auf diese Eigenschaft kann natürlich verzichtet werden, wenn $C(X)$ den Raum der stetigen reellwertigen Funktionen bezeichnet).

Dann ist die von F erzeugte Algebra dicht in $C(X)$.

Beweis. Sei $A = \mathcal{A}(F)$ die durch F erzeugte Algebra. Es ist $\overline{A} = C(X)$ zu beweisen. Wir betrachten zunächst den **reellen Fall**.

$\boldsymbol{f \in \overline{A} \Longrightarrow |f| \in \overline{A}}$: Nach dem klassischen Satz von Weierstraß gibt es zu jedem $n \in \mathbb{N}$ ein Polynom P_n mit

$$\Big||t| - P_n(t)\Big| \leq \frac{1}{n} \qquad \text{für} \quad -n \leq t \leq n.$$

(Tatsächlich braucht man hierfür den Weierstraß'schen Approximationssatz, der natürlich als Spezialfall aus dem Satz von Stone–Weierstraß folgt, nicht zu verwenden: Man kann elementar zeigen, daß für $|s| \leq 1$ die Folge $\Big(p_k(s)\Big)$ mit

$$p_1(s) := 0, \quad p_{k+1}(s) := p_k(s) + \frac{1}{2}(s^2 - p_k(s)^2)$$

gleichmäßig gegen $|s| = \sqrt{s^2}$ konvergiert; dies folgt aus der monotonen und punktweisen Konvergenz $p_k(s) \nearrow |s|$; [vgl. z. B. H. Schubert: Topologie, II.4.3 Hilfssatz 5]. Also konvergiert $np_k\Big(\dfrac{t}{n}\Big)$ für $|t| \leq n$ gleichmäßig gegen $|t|$.) Damit folgt für $n \geq \|f\|_\infty$

$$\Big||f(x)| - P_n(f)(x)\Big| = \Big||f(x)| - P_n(f(x))\Big| \leq \frac{1}{n} \qquad \text{für alle} \quad x \in X.$$

Wegen $P_n(f(\cdot)) = P_n \circ f \in \mathcal{A}(\overline{A}) = \overline{A}$ folgt hieraus diese Behauptung.

$\boldsymbol{f, g \in \overline{A} \Longrightarrow \max\{f, g\}, \{f, g\} \in \overline{A}}$: Dies folgt sofort aus

$$\begin{aligned}
\max\{f,g\}(x) &= \max\{f(x), g(x)\} = \frac{1}{2}\Big(f(x)+g(x)\Big) + \frac{1}{2}|f(x)-g(x)| \\
&= \frac{1}{2}(f+g)(x) + \frac{1}{2}|f-g|(x), \\
\min\{f,g\}(x) &= \frac{1}{2}(f+g)(x) - \frac{1}{2}|f-g|(x).
\end{aligned}$$

Sei nun $f \in C(X)$ vorgegeben. Es ist zu zeigen:
Zu jedem $\boldsymbol{\varepsilon > 0}$ gibt es ein $\boldsymbol{f_\varepsilon \in \overline{A}}$ mit $\boldsymbol{\|f - f_\varepsilon\|_\infty < \varepsilon}$.

Wir zeigen zunächst: zu beliebigen $y, z \in X$ gibt es ein $f_{y,z} \in A$ mit

$$f_{y,z}(y) = f(y), \quad f_{y,z}(z) = f(z).$$

Für $y = z$ folgt dies sofort aus Voraussetzung (i). Für $y \neq z$ wählt man ein $g \in F$ mit $g(y) \neq g(z)$. Dann gibt es (Beweis!) reelle Zahlen α, β mit

$$\alpha g(y) + \beta = f(y), \qquad \alpha g(z) + \beta = f(z).$$

Mit $f_{y,z} := \alpha g + \beta$ gilt also die Behauptung.

Sei $z \in X$. Dann gibt es zu jedem $y \in X$ eine Umgebung U_y von y mit

$$f_{y,z}(x) > f(x) - \varepsilon \qquad \text{für alle} \quad x \in U_y.$$

Sei $\{U_{y_1}, \ldots, U_{y_p}\}$ eine endliche Überdeckung von X durch solche Umgebungen,

$$f_z := \max\{f_{y_1,z}, \ldots, f_{y_p,z}\}.$$

Dann ist

$$f_z \in \overline{A} \quad \text{und} \quad f_z(x) > f(x) - \varepsilon \quad \text{für alle} \quad x \in X.$$

Wegen $f_{y_j,z}(z) = f(z)$ für alle j gilt $f_z(z) = f(z)$. Also gibt es zu jedem z eine Umgebung V_z von z mit

$$f_z(x) < f(x) + \varepsilon \qquad \text{für alle} \quad x \in V_z.$$

Sei $\{V_{z_1}, \ldots, V_{z_q}\}$ eine endliche Überdeckung von X und

$$f_\varepsilon := \min\{f_{z_1}, \ldots, f_{z_q}\}.$$

Wegen $f_{z_i} > f - \varepsilon$ für $i = 1, \ldots, q$ gilt dann

$$f_\varepsilon(x) > f(x) - \varepsilon \qquad \text{für alle} \quad x \in X.$$

Andererseits gibt es zu jedem $x \in X$ ein i mit $x \in V_{z_i}$, und somit

$$f_\varepsilon(x) \leq f_{z_i}(x) < f(x) + \varepsilon.$$

Insgesamt haben wir also

$$|f(x) - f_\varepsilon(x)| < \varepsilon \qquad \text{für alle} \quad x \in X.$$

Betrachten wir nun den **komplexen Fall**: Offenbar hat die Menge der reellen Funktionen

$$\tilde{F} := \Big\{ \mathrm{Re}\, f, \mathrm{Im}\, f : f \in F \Big\} \subset A$$

die Eigenschaften (i) und (ii), d. h. jede stetige reelle Funktion (also der Real– und der Imaginärteil jeder stetigen Funktion) auf X liegt im Abschluß der durch A erzeugten Algebra. Dies gilt dann auch für jede stetige komplexwertige Funktion. ■

Korollar 12.2 *(a) Sei $C^*(\mathbb{R})$ der Raum der stetigen Funktionen auf $\mathbb{R}$ mit $\lim\limits_{x\to-\infty} f(x) = \lim\limits_{x\to\infty} f(x)$ (endlich), $z \in \mathbb{C}\backslash\mathbb{R}$. Dann ist die Menge der Polynome in $(x-z)^{-1}$ und $(x-\overline{z})^{-1}$ dicht in $C^*(\mathbb{R})$; das gleiche gilt für $(x-z)^{-2}$ und $(x-\overline{z})^{-2}$, aber nicht für $(x-z)^{-n}$ und $(x-\overline{z})^{-n}$ mit $n \geq 3$.*

(b) Ist $C^*([\mu,\infty))$ der Raum der auf $[\mu,\infty)$ stetigen Funktionen für die $\lim\limits_{x\to\infty} f(x)$ existiert und $z \in (-\infty,\mu)$, so ist die Menge der Polynome in $(x-z)^{-n}$ dicht in $C^*([\mu,\infty))$ für jedes $n \in \mathbb{N}$.

Zum ***Beweis*** betrachte man die Ein–Punkt–Kompaktifizierung $\mathbb{R}^*$ von $\mathbb{R}$; dann ist nämlich $C^*(\mathbb{R}) = C(\mathbb{R}^*)$. Die Menge F der Polynome in den Funktionen $t \mapsto (t-z)^{-1}, t \mapsto (t-\overline{z})^{-1}$ bzw. $t \mapsto (t-z)^{-2}, t \mapsto (t-\overline{z})^{-2}$ im Fall a) (bzw. $t \mapsto (t-z)^{-n}$ im Fall b) hat offenbar die im Satz geforderten Eigenschaften. Im Fall a) ist dies für $t \mapsto (t-z)^{-n},\ t \mapsto (t-\overline{z})^{-n}$ mit $n \geq 3$ nicht erfüllt (Beweis?).

Korollar 12.3 *Sei* $\tilde{C}[0,1]$ *der Raum der stetigen Funktionen auf* $[0,1]$ *mit* $f(0) = f(1)$, $e_n(x) := \exp(2\pi i n x)$ *für* $x \in [0,1]$ *und* $n \in \mathbb{Z}$. *Dann ist die lineare Hülle* $L\{e_n :\in \mathbb{Z}\}$ *dicht in* $\Big(\tilde{C}[0,1], \|\cdot\|_\infty\Big)$ *(und somit in* $L_2(0,1)$*, vgl. Beispiel 9.16).*

Beweis. Sei $S^1 = \{z \in \mathbb{C} : |z| = 1\}$. Durch $Tf(x) = f(e^{2\pi i x})$für $x \in [0,1]$ und $f \in C(S^1)$ wird eine Isometrie von $\Big(C(S^1), \|\cdot\|_\infty\Big)$ auf $\Big(\tilde{C}[0,1], \|\cdot\|_\infty\Big)$ definiert. Die Polynome in z und $\overline{z}$ sind nach dem Satz von Stone–Weierstraß dicht in $\Big(C(S^1), \|\cdot\|_\infty\Big)$. Bei Anwendung der Transformation T gehen diese Polynome in z und $\overline{z}$ über in Polynome in $e^{2\pi i x}$ und $e^{-2\pi i x}$, d. h. in Linearkombinationen der $e_n(\cdot)$, $n \in \mathbb{Z}$. Also ist $L\{e_n : n \in \mathbb{Z}\}$ dicht in $\Big(\tilde{C}[0,1], \|\cdot\|_\infty\Big)$. ■

Literatur

[1] Alt, H. W.: Lineare Funktionalanalysis, eine anwendungsorientierte Einführung. Springer 1985

[2] Constantinescu, F.:

[3] Dunford, N. – Schwartz, J. T.: Linear Operators Part I: General theory, part II: Spectral theory, self–adjoint operators in Hilbert space. Interscience, New York/London/Sydney 1964/1963

[4] Heuser, H.: Funktionalanalysis. Mathematische Leitfäden. Teubner–Verlag, Stuttgart 1986

[5] Jantscher, L.:

[6] Jörgens, K.: Lineare Integraloperatoren. Mathematische Leitfäden. Teubner–Verlag, Stuttgart 1970

[7] Meise, R, Voigt, D.: Einführung in die Funktionalanalysis. Vieweg–.Studium 1992

[8] Reed, M., Simon, B.: Methods of Modern Mathematical Physics. I Functional Analysis. Academic Press, New York/San Francisco/London 1972

[9] Rudin, W.: Functional Analysis. McGraw–Hill Book Company, New York 1973

[10] Schubert, H.: Topologie, eine Einführung. Mathematische Leitfäden. Teubner–Verlag, Stuttgart 1964

[11] Weidmann, J.: Lineare Operatoren in Hilberträumen, Teil 1: Grundlagen. Mathematische Leitfäden, Teubner–Verlag, Stuttgart/Leipzig/Wiesbaden 2000

Index