Skript zur Vorlesung

Funktionalanalysis

Sommersemester 2017

Robert Denk

Universität Konstanz

Fachbereich Mathematik und Statistik

Stand: 03.08.2017

Inhaltsverzeichnis

1. Topologische und metrische Räume, Kompaktheit

1.1 Worum geht's? Aus der Analysis sind viele Begriffe bekannt, welche mit Konvergenz und Stetigkeit zusammenhängen. Auftretende Strukturen sind hier z.B. metrische Räume, normierte Räume oder Hilberträume. Der gemeinsame Begriff für diese Räume sind topologische Räume. Die Topologie ist eines Raums beschreibt das System aller offenen Mengen und ist axiomatisch ähnlich wie eine σ-Algebra definiert. Einer der wichtigsten Begriffe bei topologischen Räumen ist die Kompaktheit. Bekannte Sätze wie „Bilder kompakter Mengen bei stetigen Funktionen sind kompakt“ übertragen sich ohne Schwierigkeiten auf allgemeine topologische Räume.

a) Topologie und Metrik

Wir starten mit der Definition einiger grundlegender Begriffe der Topologie. Eine Topologie ist ähnlich wie eine σ-Algebra ein Mengensystem.

1.2 Definition. Sei $X \neq \emptyset$ eine Menge; $\mathscr{P}(X)$ bezeichne die Potenzmenge von X.

a) Ein Mengensystem $\mathcal{T} \subset \mathscr{P}(X)$ heißt eine Topologie auf X, falls gilt

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

(ii) Falls $A, B \in \mathcal{T}$, so ist auch $A \cap B \in \mathcal{T}$,

(iii) Falls I eine Indexmenge ist und $A_i \in \mathcal{T} \quad (i \in I)$, so ist auch $\bigcup_{i\in I} A_i \in \mathcal{T}$.

In diesem Falls heißt $(X, \mathcal{T})$ ein topologischer Raum.

b) Seien $(X_1, \mathcal{T}_1)$ und $(X_2, \mathcal{T}_2)$ topologische Räume. Dann heißt eine Abbildung $f\colon X_1 \to X_2$ stetig, falls $f^{-1}(\mathcal{T}_2) \subset \mathcal{T}_1$. Die Abbildung f heißt offen, falls für alle $U \in \mathcal{T}_1$ gilt $f(U) \in \mathcal{T}_2$. Die Abbildung f heißt ein Homöomorphismus, falls f bijektiv ist und f und f^{-1} beide stetig sind.

c) Sei $\mathscr{U} \subset \mathscr{P}(X)$. Dann heißt die kleinste Topologie, die $\mathscr{U}$ enthält, die von $\mathscr{U}$ erzeugte Topologie $\mathcal{T}(\mathscr{U})$. In diesem Fall heißt $\mathscr{U}$ eine Subbasis der Topologie.

d) Sei $(X, \mathcal{T})$ ein topologischer Raum. Dann heißt ein Mengensystem $\mathscr{U} \subset \mathcal{T}$ eine Basis der Topologie $\mathcal{T}$, falls sich jedes Element von $\mathcal{T}$ als Vereinigung von Elementen aus $\mathscr{U}$ schreiben lässt.

e) Sei I eine Menge und $(Y_i, \mathcal{T}_i)$ topologischer Raum für $i \in I$. Sei $F = \{f_i\colon X \to Y_i\}_{i\in I}$ eine Familie von Abbildungen. Dann heißt die kleinste (gröbste) Topologie auf X, für die alle $f \in F$ stetig sind, die F-schwache Topologie $\mathcal{T}(F)$ auf X.

f) Sei $X = \prod_{i\in I} X_i$ das kartesische Produkt, wobei $(X_i, \mathcal{T}_i)$ ein topologischer Raum für $i \in I$ ist. Sei $\mathrm{pr}_i\colon X \to X_i$ die Projektion auf die i-te Komponente. Dann heißt $\mathcal{T}(\{\mathrm{pr}_i\colon i \in I\})$ die Produkttopologie auf X.

g) Sei $(X, \mathcal{T})$ ein topologischer Raum, und sei $Y \subset X$ eine nichtleere Teilmenge. Das Mengensystem $\mathcal{T}_Y := \{U \cap Y : U \in \mathcal{T}\} \subset \mathscr{P}(Y)$ heißt Spurtopologie auf Y (auch Teilraumtopologie oder relative Topologie oder induzierte Topologie genannt). Dies ist die gröbste Topologie auf Y, für die die Inklusionsabbildung $Y \to X$, $y \mapsto y$, stetig ist.

1.3 Bemerkung. a) Die kleinste (gröbste) Topologie auf einer Menge X ist gegeben durch $\mathcal{T}_1 := \{\emptyset, X\}$. Die größte (feinste) Topologie ist gegeben durch $\mathcal{T}_2 := \mathcal{P}(X)$.

b) In der Situation von 1.2 d) gilt

$$\mathcal{T}(F) = \mathcal{T}\Big(\big\{f_i^{-1}(U_i) : U_i \in \mathcal{T}_i,\ i \in I\big\}\Big).$$

[[Sei $\mathscr{U} := \{f_i^{-1}(U_i) : U_i \in \mathcal{T}_i,\ i \in I\}$.

$\mathcal{T}(F) \supset \mathcal{T}(\mathscr{U})$: Da jedes f_i stetig ist, gilt $f_i^{-1}(U_i) \in \mathcal{T}(F)$ $(U_i \in \mathcal{T}_i, i \in I)$, d.h. $\mathscr{U} \subset \mathcal{T}(F)$. Da $\mathcal{T}(F)$ eine Topologie ist, welche $\mathscr{U}$ enthält, gilt $\mathcal{T}(F) \supset \mathcal{T}(\mathscr{U})$.

$\mathcal{T}(F) \subset \mathcal{T}(\mathscr{U})$: Wegen $f_i^{-1}(U_i) \in \mathcal{T}(\mathscr{U})$ ist jedes f_i stetig. Damit ist $\mathcal{T}(\mathscr{U})$ eine Obermenge der kleinsten Topologie, in welcher jedes f_i stetig ist.]]

c) Die erzeugte Topologie $\mathcal{T}(\mathscr{U})$ ist das System aller Mengen der Form $\bigcup_{i\in I}\bigcap_{n=1}^{N_i} U_{in}$ mit $U_{in} \in \mathscr{U}$, $N_i \in \mathbb{N}_0$, I eine Indexmenge. Dabei ist $\bigcap_{n=1}^{0} U_{in} = \bigcap_{n\in\emptyset} U_{in} = X$. Insbesondere ist eine Subbasis einer Topologie bereits eine Basis, falls sie durchschnittstabil ist, d.h. falls der Durchschnitt endlich vieler Mengen aus der Subbasis wieder Element der Subbasis ist.

[[Jede Menge dieser Form muss nach Definition einer Topologie in $\mathcal{T}(\mathscr{U})$ enthalten sein. Zu zeigen ist daher, dass das System all dieser Mengen eine Topologie ist. Sei $\mathscr{B} := \{\bigcap_{n=1}^{N} U_n : U_n \in \mathscr{U}, N \in \mathbb{N}_0\}$. Dann gilt $X \in \mathscr{B}$ und zu $V_1, V_2 \in \mathscr{B}$ und $x \in V_1 \cap V_2$ existiert ein $V \in \mathscr{B}$ mit $x \in V \subset V_1 \cap V_2$ (denn man kann $V := V_1 \cap V_2$ wählen). Dann ist $\mathscr{B}$ Basis einer Topologie, und (nach Definition einer Basis) das oben angegebene System ist gleich $\mathcal{T}(\mathscr{B})$, also insbesondere eine Topologie.]]

d) Zur Erinnerung: das kartesische Produkt X einer (eventuell überabzählbaren) Familie $(X_i)_{i\in I}$ von Mengen X_i ist definiert als die Menge aller Abbildungen $f\colon I \to \cup_{i\in I} X_i$ mit $f(i) \in X_i$ für jedes $i \in I$. Falls jedes X_i nichtleer ist, dann ist auch das kartesische Produkt nichtleer — dies ist äquivalent zum Auswahlaxiom, welches wir in dieser Vorlesung immer annehmen wollen.

e) Man beachte die Ähnlichkeit zur Definition einer σ-Algebra. Bei einer σ-Algebra $\mathcal{A}$ gilt statt a) (iii) nur

$$A_n \in \mathcal{A} \quad (n \in \mathbb{N}) \Longrightarrow \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A},$$

dafür ist mit einer Menge A auch $X \setminus A$ in der σ-Algebra. Viele Begriffe wie die erzeugte σ-Algebra und die Produkt-σ-Algebra sind analog definiert. Es gibt allerdings keine so einfache Darstellung der erzeugten σ-Algebra wie in a).

1.4 Definition. Sei $(X, \mathcal{T})$ topologischer Raum.

a) Eine Menge $U \subset X$ heißt genau dann offen, falls $U \in \mathcal{T}$, und genau dann abgeschlossen, falls $X \setminus U \in \mathcal{T}$. Das Innere $\mathring{A}$ einer Menge $A \subset X$ ist definiert als die größte offene Menge U mit $U \subset A$. Der Abschluss $\overline{A}$ ist definiert als die kleinste abgeschlossene Menge V mit $A \subset V$.

b) Seien $A \subset B \subset X$. Dann heißt A dicht in B, falls $\overline{A} \supset B$ gilt.

c) Der topologische Raum $(X, \mathcal{T})$ heißt separabel, falls eine abzählbare Teilmenge $A \subset X$ existiert, welche dicht in X ist.

d) Eine (nicht notwendig offene) Menge $V \subset X$ heißt eine Umgebung eines Punktes $x \in X$, falls eine offene Menge $U \in \mathcal{T}$ existiert mit $x \in U \subset V$. Eine Familie $\mathscr{N}$ von Teilmengen von X heißt eine Umgebungsbasis des Punktes $x \in X$, wenn jedes $N \in \mathscr{N}$ eine Umgebung von x ist und für jede Umgebung M von x ein $N \in \mathscr{N}$ existiert mit $N \subset M$.

e) $(X, \mathcal{T})$ heißt Hausdorff-Raum (oder T_2-Raum), falls für jedes $x, y \in X$ mit $x \neq y$ offene Mengen $U_x, U_y \in \mathcal{T}$ existieren mit $x \in U_x,\ y \in U_y$ und $U_x \cap U_y = \emptyset$.

f) $(X, \mathcal{T})$ heißt normal (oder T_4-Raum), falls $(X, \mathcal{T})$ Hausdorffsch ist und für alle abgeschlossenen Mengen A_1, A_2 mit $A_1 \cap A_2 = \emptyset$ offene Mengen $U_1 \supset A_1$, $U_2 \supset A_2$ existieren mit $U_1 \cap U_2 = \emptyset$.

Die beiden folgenden Sätze werden hier nicht bewiesen.

1.5 Satz (Lemma von Urysohn)**.** *Sei $(X, \mathcal{T})$ normaler topologischer Raum, und seien $A_1, A_2 \subset X$ abgeschlossen mit $A_1 \cap A_2 = \emptyset$. Dann existiert ein $f \in C(X; \mathbb{R})$ mit $0 \leq f \leq 1$ und $f|_{A_1} = 0,\ f|_{A_2} = 1$.*

1.6 Satz (Erweiterungslemma von Tietze)**.** *Sei (X, τ) ein normaler topologischer Raum. Sei $M \subset X$ abgeschlossen, $a, b \in \mathbb{R}$ mit $a < b$ und $f \colon M \to [a, b]$ stetig. Dann existiert eine stetige Fortsetzung $F \colon X \to [a, b]$ von f.*

Um hierbei die Stetigkeit der Abbildung f erklären zu können, wird M mit der Spurtopologie ausgestattet.

Bei topologischen Räumen kann man einige Begriffe definieren, die schon aus dem ersten Studienjahr bekannt sind. Hier einige Beispiele:

1.7 Definition. Seien $(X, \mathcal{T})$ ein topologischer Raum, $A \subset X$ eine Teilmenge und

$(x_n)_{n\in\mathbb{N}} \subset X$ eine Folge.

a) Ein Element $x \in X$ heißt Häufungspunkt der Menge A, falls für jede Umgebung U von x ein $a \in A$ existiert mit $a \in U$ und $a \neq x$.

Ein Element $x \in X$ heißt Häufungspunkt der Folge $(x_n)_{n\in\mathbb{N}}$, falls in jeder Umgebung von x unendlich viele Folgenglieder liegen.

b) Ein Element $x \in X$ heißt Grenzwert oder Limes der Folge $(x_n)_{n\in\mathbb{N}}$, falls in jeder offenen Menge, die x enthält, alle bis auf endlich viele Folgenglieder liegen. Man schreibt $\lim_{n\to\infty} x_n := x$, oder $x_n \to x$ $(n \to \infty)$.

Und auch für die Stetigkeit einer Abbildung $f\colon X \to Y$ gibt es eine Beschreibung, die dem bereits bekannten ε–δ–Zugang entspricht:

1.8 Lemma. *Seien $(X, \mathcal{T}_X)$ und $(Y, \mathcal{T}_Y)$ topologische Räume. Dann ist eine Abbildung $f\colon X \to Y$ stetig genau dann, wenn zu jedem $x_0 \in X$ und jeder Umgebung V_Y von $y_0 := f(x_0)$ eine Umgebung U_X von x_0 existiert mit $f(U_X) \subset V_Y$.*

[[(i) Sei f stetig. Wähle $W \subset Y$ offen mit $y_0 \in W \subset V_Y$ (Definition Umgebung). Dann ist $U_X := f^{-1}(W) \in \mathcal{T}_X$ mit $f(U_X) \subset W \subset V_Y$.

(ii) Es gelte die Bedingung des Lemmas, und sei $B \subset Y$ offen. Für jedes $x \in f^{-1}(B)$ ist B eine Umgebung von $f(x)$, und nach Voraussetzung existiert eine Umgebung V_x von x mit $f(V_x) \subset B$. Damit existiert eine offene Menge U_x mit $x \in U_x \subset V_x$ und $f(U_x) \subset B$, also ist $f^{-1}(B) = \bigcup_{x\in f^{-1}(B)} U_x$ offen.]]

1.9 Definition. Sei X eine Menge und $d\colon X \times X \to \mathbb{R}$ eine Abbildung. Dann heißt d eine Metrik und (X, d) ein metrischer Raum, falls für alle $x, y, z \in X$ gilt:

(i) $d(x, y) = d(y, x)$

(ii) $d(x, y) = 0 \iff x = y$

(iii) $d(x, z) \leq d(x, y) + d(y, z)$

1.10 Beispiele. Einige Beispiele von metrischen Räumen sind:

a) Sei X beliebig und

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

(diskrete Metrik).

b) Die Standardmetrik auf $X = \mathbb{R}$ oder $\mathbb{C}$ ist gegeben durch $d(x, y) = |x - y|$.

c) Sei $X = C([0,1];\mathbb{R})$. Dann sind

$$d_1(f,g) := \|f-g\|_\infty := \sup_{x\in[0,1]} |f(x)-g(x)|,$$
$$d_2(f,g) := \|f-g\|_{L^1} := \int_0^1 |f(x)-g(x)|\,\mathrm{d}x$$

zwei Metriken auf X.

1.11 Definition. Sei (X,d) ein metrischer Raum.

a) Die (offene) Kugel um x_0 mit Radius $r>0$ ist definiert als $B(x_0,r) := \{x \in X \colon d(x_0,x) < r\}$.

b) Eine Teilmenge $U \subset X$ heißt offen, falls für alle $u \in U$ ein $\varepsilon > 0$ existiert mit $B(u,\varepsilon) \subset U$. Dies definiert die Topologie $\mathcal{T}_d$, d.h. jeder metrische Raum ist damit ein topologischer Raum $(X, \mathcal{T}_d)$.

c) Ein topologischer Raum $(X,\mathcal{T})$ heißt metrisierbar, falls es eine Metrik d gibt mit $\mathcal{T}_d = \mathcal{T}$.

[[Zu b): Man sieht direkt aus der Definition, dass

$$\mathcal{T}_d := \{U \subset X : \forall\, u \in U\ \exists\, \varepsilon > 0 : B(u,\varepsilon) \subset U\}$$

eine Topologie ist: Offensichtlich ist $\emptyset, X \in \mathcal{T}_d$. Für $U_1, U_2 \in \mathcal{T}_d$ und $u \in U_1 \cap U_2$ existieren $\varepsilon_1, \varepsilon_2 > 0$ mit $B(u,\varepsilon_1) \subset U_1$ und $B(u,\varepsilon_2) \subset U_2$. Damit ist $B(u,\varepsilon) \subset U_1 \cap U_2$ für $\varepsilon := \min\{\varepsilon_1,\varepsilon_2\}$.

Falls $U_\lambda \in \mathcal{T}_d$ für $\lambda \in \Lambda$ und $u \in \bigcup_\lambda U_\lambda$, so existiert ein λ_0 mit $u \in U_{\lambda_0}$. Damit existiert ein $\varepsilon > 0$ mit $B(u,\varepsilon) \subset U_{\lambda_0}$ und damit $B(u,\varepsilon) \subset \bigcup_\lambda U_\lambda$.]]

1.12 Beispiel. Das folgende Beispiel zeigt, dass nicht jeder topologischer Raum metrisierbar ist: Sei X ein Menge mit mindestens zwei Elementen und $\mathcal{T} := \{\emptyset, X\}$. Falls $(X,\mathcal{T})$ metrisierbar mit Metrik d wäre, so wäre für zwei Elemente $x \neq y \in X$ der Abstand $\gamma := d(x,y)$ positiv und es folgt $B(x,\frac{\gamma}{2}) \in \mathcal{T}_d$. Wegen $\emptyset \subsetneq B(x,\frac{\gamma}{2}) \subsetneq X$ erhalten wir einen Widerspruch.

Metrische Räume sind zwar weniger allgemein als topologische Räume, aber sie verfügen über einige nützliche Eigenschaften, wie in den beiden folgenden Bemerkungen ausgeführt:

1.13 Bemerkung. Sei X ein metrischer Raum und $A \subset X$. Dann sind äquivalent:

(i) Die Menge A enthält alle ihre Häufungspunkte.

(ii) Es gilt $A = \overline{A}$.

(iii) Die Menge A ist abgeschlossen.

1.14 Bemerkung. Jeder metrische Raum ist ein Hausdorffraum: Betrachte

$$U_x = B\Big(x, \frac{d(x,y)}{2}\Big), \quad U_y = B\Big(y, \frac{d(x,y)}{2}\Big).$$

1.15 Beispiel. Wir greifen das Beispiel 1.10 c) noch einmal auf. Die identische Abbildung $\mathrm{id}_X \colon f \mapsto f$ ist zwar bijektiv, aber die Stetigkeit hängt von der gewählten Metrik ab:

So ist $\mathrm{id}\colon (C([0,1]), d_1) \to (C([0,1]), d_2)$ stetig, denn es gilt

$$d_2(f,g) = \int_0^1 |f(x) - g(x)|\,\mathrm{d}x \le \sup_{x\in[0,1]} |f(x) - g(x)| = d_1(f,g).$$

Hingegen ist $\mathrm{id}\colon (C([0,1]), d_2) \to (C([0,1]), d_1)$ nicht stetig: Definiere zu $n \in \mathbb{N}$ die Funktion f_n wie in folgender Abbildung:

Dann gilt

$$d_2(f_n, 0) = \int_0^1 |f(x)|\,\mathrm{d}x = \frac{1}{2}\left(\frac{1}{2n-1} - \frac{1}{2n+1}\right) = \frac{1}{(2n-1)(2n+1)} \longrightarrow 0,$$

aber: $d_1(f_n, 0) = \sup_{x\in[0,1]} |f_n(x)| = 1$.

1.16 Definition. Seien (X, d_X), (Y, d_Y) metrische Räume. Dann heißt eine Abbildung $f\colon X \to Y$ eine Isometrie, falls

$$d_Y(f(x), f(y)) = d_X(x,y) \quad (x, y \in X) \tag{1-1}$$

gilt. Eine bijektive Isometrie heißt isometrischer Isomorphismus.

1.17 Beispiel. Versieht man $\mathbb{R}^n$ mit der euklidischen Metrik, so sind die durch orthogonale Matrizen gegebenen Abbildungen Isometrien von $\mathbb{R}^n$ nach $\mathbb{R}^n$.

1.18 Definition. Sei (X, d) ein metrischer Raum.

a) Eine Folge $(x_n)_{n\in\mathbb{N}} \subset X$ heißt eine Cauchyfolge, falls gilt

$$\forall\, \varepsilon > 0\ \exists\, n_0 \in \mathbb{N}\ \forall\, n, m \ge n_0\colon d(x_n, x_m) < \varepsilon.$$

Der Raum heißt vollständig, falls jede Cauchyfolge konvergent ist.

b) Eine Menge $A \subset X$ heißt beschränkt, falls ein $r > 0$ und ein $x_0 \in X$ existieren mit $A \subset B(x_0, r)$.

1.19 Beispiel. Sei (X, d) ein metrischer Raum mit diskreter Metrik, und sei $(x_n)_{n\in\mathbb{N}}$ eine Cauchyfolge. Dann existiert ein $x_0 \in X$ und ein $n_0 \in \mathbb{N}$ so, dass für alle $n \geq n_0$ gilt: $x_n = x_0$.

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

a) Es existiert ein beschränkter metrischer Raum (X, d'), welcher homöomorph zu (X, d) ist.

b) Es existiert ein vollständiger metrischer Raum (Y, d_Y) und ein dichter Teilraum $Y_0 \subset Y$ so, dass (X, d) und (Y_0, d_Y) isometrisch isomorph sind.

Beweis. a) Man wählt $d'(x, y) := \frac{d(x,y)}{1+d(x,y)}$ $(x, y \in X)$.

b) wird hier nicht bewiesen. □

Auch der folgende Satz, der aus der Analysis bekannt sein sollte, wird hier nicht bewiesen.

1.21 Satz (Banachscher Fixpunktsatz)**.** *Sei (X, d) ein vollständiger metrischer Raum, und $T: X \to X$ eine Kontraktion, d.h. es existiere ein $k \in [0, 1)$ mit*

$$d(Tx, Ty) \leq k \cdot d(x, y) \quad (x, y \in X).$$

Dann existiert genau ein Fixpunkt von T, d.h. genau ein $x^* \in X$ mit $Tx^* = x^*$. Für alle $x_0 \in X$ konvergiert die Folge $(T^n x_0)_{n\in\mathbb{N}}$ gegen x^*, und es gilt die Abschätzung

$$d(x^*, T^n x_0) \leq \frac{k^n}{1-k} \cdot d(Tx_0, x_0).$$

b) Kompaktheit

1.22 Definition. Sei $(X, \mathcal{T})$ ein topologischer Raum und $A \subset X$.

a) Eine offene Überdeckung von A ist eine Familie $(U_\lambda)_{\lambda\in\Lambda}$ (wobei Λ eine beliebige Indexmenge ist) mit $U_\lambda \in \mathcal{T}$ und $A \subset \bigcup_{\lambda\in\Lambda} U_\lambda$. Eine endliche Teilüberdeckung der offenen Überdeckung $(U_\lambda)_{\lambda\in\Lambda}$ ist eine Überdeckung der Form $A \subset \bigcup_{j=1}^n U_{\lambda_j}$ mit $n \in \mathbb{N}$ und $\lambda_j \in \Lambda$.

b) Die Menge $A \subset X$ heißt kompakt, wenn jede offene Überdeckung von A eine endliche Teilüberdeckung besitzt.

1.23 Satz. *Seien $(X, \mathcal{T}_X)$ und $(Y, \mathcal{T}_Y)$ topologische Räume. Falls X kompakt ist und $f: (X, \mathcal{T}_X) \to (Y, \mathcal{T}_Y)$ stetig, so ist der Wertebereich $f(X)$ kompakt.*

Beweis. Sei $f(X) \subset \bigcup_{\lambda\in\Lambda} V_\lambda$ mit $V_\lambda \in \mathcal{T}_Y$. Setze $U_\lambda := f^{-1}(V_\lambda)$. Da f stetig ist, folgt $U_\lambda \in \mathcal{T}_X$, und $X \subset \bigcup_{\lambda\in\Lambda} U_\lambda$ ist eine offene Überdeckung von X. Wegen der Kompaktheit von X existiert eine endliche Teilüberdeckung $X \subset \bigcup_{j=1}^n U_{\lambda_j}$. Dann ist aber $f(X) \subset \bigcup_{j=1}^n V_{\lambda_j}$ ebenfalls eine endliche Teilüberdeckung, und $f(X)$ ist kompakt. □

1.24 Bemerkung. In $(\mathbb{R}^n, |\cdot|)$ gilt: Eine Teilmenge A ist genau dann kompakt, wenn A beschränkt und abgeschlossen ist. In beliebigen metrischen Räumen (X, d) gilt nur eine Richtung: Jede kompakte Teilmenge A ist beschränkt und abgeschlossen, aber nicht umgekehrt, wie das unten stehende Beispiel 1.25 b) zeigt.

1.25 Beispiele. a) Sei X eine unendliche Menge mit diskreter Metrik. Dann ist X beschränkt und abgeschlossen, aber $X = \bigcup_{x\in X} B(x, \frac{1}{2})$ ist eine offene Überdeckung von X, zu welcher keine endliche Teilüberdeckung existiert.

b) Sei

$$X := \ell^2 := \left\{ x = (\xi_1, \xi_2, \dots) \colon \xi_j \in \mathbb{R},\ \sum_{j=1}^\infty |\xi_j|^2 < \infty \right\}.$$

Durch

$$d(x,y) := \Big(\sum_{j=1}^\infty |\xi_j - \eta_j|^2 \Big)^{\frac{1}{2}}$$

wird eine Metrik auf ℓ^2 definiert. Für die Einheitsvektoren

$$e_j = (0, 0, \dots, 0, \underbrace{1}_{j\text{-te Stelle}}, 0, \dots)$$

gilt $x = \sum_{j=1}^\infty \xi_j e_j$ und $d(e_i, e_j) = \sqrt{2}$ $(i \neq j)$. Die Einheitssphäre $S := \{x \colon d(x, 0) = 1\}$ ist beschränkt und abgeschlossen, aber nicht kompakt.

Denn $S \subset \bigcup_{x\in S} B(x, \frac{1}{2})$ ist eine offene Überdeckung von S. Angenommen, es gäbe eine endliche Teilüberdeckung $S \subset \bigcup_{j=1}^n B(x_j, \frac{1}{2})$. Dann kann jede Umgebung $B(x_j, \frac{1}{2})$ höchstens einen Einheitsvektor e_k enthalten, wegen $1 < \sqrt{2}$. Es gibt aber unendlich viele solche Einheitsvektoren. Widerspruch. Also ist S nicht kompakt.

c) Wäre in einem metrischen Raum jede abgeschlossene, beschränkte Menge automatisch kompakt, so wäre jeder metrische Raum kompakt, da jeder metrische Raum homöomorph zu einem beschränkten, metrischen Raum ist.

1.26 Definition. Sei X ein topologischer Raum.

a) Der Raum X heißt folgenkompakt, wenn jede Folge in X eine konvergente Teilfolge besitzt.

b) Eine Menge $A \subset X$ heißt relativ kompakt, wenn $\overline{A}$ kompakt ist.

c) Sei nun X sogar ein metrischer Raum mit Metrik d. Der Raum X heißt totalbeschränkt oder präkompakt, wenn es zu jedem $\varepsilon > 0$ eine endliche Menge $E \subset X$ gibt mit:

$$X = \bigcup_{x \in E} B(x, \varepsilon). \tag{1-2}$$

E heißt (endliches) ε-Netz für X.

1.27 Bemerkung. a) Sei (X, d) ein metrischer Raum. Dann ist eine Teilmenge $A \subset X$ (mit der induzierten Metrik) genau dann präkompakt, wenn eine endliche Teilmenge $E \subset X$ existiert mit (1-2). Denn mit Hilfe der Dreiecksungleichung kann man sich dann ein ε-Netz in A konstruieren.

b) Sei (X, d) ein metrischer Raum. Dann kann man leicht zeigen, dass X genau dann kompakt ist, wenn X folgenkompakt ist.

Der folgende Satz ist einer der wichtigsten Sätze der Topologie; der Beweis (der etwa mit Hilfe von Ultrafiltern geführt werden kann) ist aber zu aufwändig für diese Vorlesung.

1.28 Satz (Tychonov)**.** *Sei I eine Menge und $X = \prod_{i \in I} X_i$ das kartesische Produkt der topologischen Räume $(X_i, \mathcal{T}_i)$. Falls jedes $(X_i, \mathcal{T}_i)$ kompakt ist, dann ist auch X in der Produkttopologie kompakt.*

2. Normierte Räume und Hilberträume

2.1 Worum geht's? Unter den topologischen Räumen haben die Hilberträume die meiste Struktur: Sie besitzen ein Skalarprodukt, d.h. man kann von Orthogonalität und auch von Winkeln sprechen. In Kombination mit der Vollständigkeit ergeben sich eine Vielzahl von in Anwendungen wichtigen Sätzen, wie etwa die Existenz des orthogonalen Komplements und eine einfache Darstellung des Dualraums. Jeder nichttriviale Hilbertraum besitzt zudem eine Orthonormalbasis, bezüglich welcher alle Elemente des Hilbertraums als Reihe geschrieben werden können. Dies verallgemeinert den Begriff eines linearen Erzeugendensystems, wie er aus der linearen Algebra für endlich-dimensionale Vektorräume bekannt ist.

Im folgenden sei X ein $\mathbb{K}$-VR, wobei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

a) Normierte Räume und Banachräume

2.2 Definition. Eine Norm ist eine Abbildung $\|\cdot\| : X \to [0, \infty)$ mit folgenden Eigenschaften:

(i) $\|x\| = 0 \iff x = 0$ $(x \in X)$,

(ii) $\|\alpha x\| = |\alpha| \cdot \|x\|$ $(\alpha \in \mathbb{K},\ x \in X)$

(iii) $\|x + y\| \leq \|x\| + \|y\|$ $(x, y \in X)$

In diesem Fall heißt $(X, \|\cdot\|)$ normierter Raum. Falls in (i) nur die Richtung „$\Leftarrow$" gilt, so heißt $\|\cdot\|$ eine Halbnorm oder Seminorm.

2.3 Bemerkung. a) Ein normierter Raum $(X, \|\cdot\|)$ wird mit $d(x, y) := \|x - y\|$ zu einem metrischen Raum (X, d). Die Umkehrung gilt i.A. nicht.

b) Sei (X, d) ein metrischer Vektorraum, so ist $\widehat{d}\colon X \to \mathbb{R}$, $\widehat{d}(x) := d(x, 0)$ eine Norm genau dann, wenn gilt

(i) $d(x + z, y + z) = d(x, y)$, $\forall x, y, z \in X$

(ii) $d(\alpha x, \alpha y) = |\alpha| d(x, y)$, $\forall x, y \in X, \alpha \in \mathbb{K}$

Wenn dies gilt, dann ist die von der Norm $\widehat{d}$ induzierte Metrik wieder d.

2.4 Definition. Ein Banachraum ist ein vollständiger normierter Raum.

2.5 Beispiele. a) Der Raum $(\mathbb{R}^n, \|\cdot\|)$, $\|\cdot\| = |\cdot|$ ist ein Banachraum.

b) Sei A kompakter, metrischer Raum und Y Banachraum. Dann ist der Raum $C(A; Y)$ aller stetigen Funktionen von A nach Y, versehen mit der Supremumsnorm $\|f\|_\infty := \sup_{a \in A} \|f(a)\|_Y$, ein Banachraum.

2.6 Definition und Satz. *a) Sei $p\colon X \to [0,\infty)$ eine Seminorm. Dann ist das System*

$$\Big\{U \subset X : \forall\, x \in U\; \exists\, \varepsilon > 0\colon\; B(x,\varepsilon) \subset U\Big\}$$

eine Topologie auf X und heißt die von p erzeugte Topologie (vgl. Definition 1.11). Dabei ist $B(x,\varepsilon) := \{y \in X : p(x-y) < \varepsilon\}$.

b) Diese von einer Seminorm p erzeugte Topologie ist genau dann Hausdorffsch, wenn die Seminorm sogar eine Norm ist. Die Menge $\{B(x,\varepsilon) : x \in X,\, \varepsilon > 0\}$ bildet eine Basis der Topologie. Die von p erzeugte Topologie ist die gröbste Topologie auf X, bezüglich der alle Abbildungen $x \mapsto p(x-x_0)$ mit $x_0 \in X$ stetig sind.

[[Das das System $\mathcal{T}_p := \{U \subset X : \forall\, x \in U\; \exists\, \varepsilon > 0\colon\; B(x,\varepsilon) \subset U\}$ eine Topologie ist, sieht man sofort wie in 1.11. Falls die Seminorm eine Norm ist, ist X Hausdorffsch nach Bem. 1.14. Ansonsten existiert ein $x \in X \setminus \{0\}$ mit $p(x) = 0$, und für die Punkte 0 und x ist die Bedingung eines Hausdorff-Raums verletzt.

Aufgrund der Dreiecksungleichung ist $B(x,\varepsilon) \in \mathcal{T}_p$ für alle $x \in X$ und $\varepsilon > 0$. Andererseits lässt sich jedes $U \in \mathcal{T}_p$ in der Form $U = \bigcup_{x\in U} B(x,\varepsilon_x)$ schreiben, d.h. die Menge aller $B(x,\varepsilon)$ bildet eine Basis der Topologie.

Sei $\mathcal{T}_1$ die von allen Abbildungen $p_{x_0} : x \mapsto p(x-x_0)$ erzeugte Topologie auf X. Dann gilt $B(x_0,\varepsilon) = p_{x_0}^{-1}((-\varepsilon,\varepsilon)) \in \mathcal{T}_1$ und damit $\mathcal{T}_p \subset \mathcal{T}_1$.

Für die Richtung $\mathcal{T}_p \supset \mathcal{T}_1$ ist zu zeigen, dass jedes $p_{x_0}\colon (X,\mathcal{T}_p) \to \mathbb{R}$ stetig ist. Dazu verwendet man Lemma 1.8: Sei $x \in X$ und $W \subset \mathbb{R}$ eine Umgebung von $p_{x_0}(x) = p(x-x_0)$. Dann existiert ein $\varepsilon > 0$ mit $V := (p_{x_0}(x) - \varepsilon, p_{x_0}(x) + \varepsilon) \subset W$. Mit Hilfe der Dreiecksungleichung sieht man, dass $U := B(x,\varepsilon)$ eine Umgebung von x ist mit $p_{x_0}(U) \subset V \subset W$.]]

Der Beweis des folgenden Satzes sollte aus der Analysis bekannt sein.

2.7 Satz. *Seien X, Y normierte Räume, $T\colon X \to Y$ linear. Dann sind äquivalent:*

(i) *T ist beschränkt, d.h. $\exists\, c > 0\colon\; \|Tx\|_Y \le c\,\|x\|_X \quad (x \in X)$.*

(ii) *$T\colon X \to Y$ ist stetig.*

(iii) *$T\colon X \to Y$ ist stetig an der Stelle $0 \in X$.*

2.8 Definition. Seien X, Y Vektorräume, ausgestattet mit Topologien. Der Raum

$$L(X,Y) := \{T\colon X \to Y \;:\; T \text{ linear, stetig}\}$$

heißt der Raum der linearen stetigen Operatoren von X nach Y (im Falle von normierten Räumen X und Y sind diese Operatoren dann auch beschränkt). Wir setzen $L(X) := L(X,X)$. Der Raum $X' := L(X,\mathbb{K})$ heißt der (topologische) Dualraum von X. Oft schreibt man Tx statt $T(x)$.

Für $T \in L(X,Y)$ (mit normierten Räumen X und Y) definiert man

$$\|T\| := \sup_{x\in X\setminus\{0\}} \frac{\|Tx\|_Y}{\|x\|_X} = \sup_{x\in X,\; \|x\|_X \le 1} \|Tx\|_Y\,.$$

$\|T\|$ heißt die Operatornorm von T. Sei

$$\begin{aligned} \ker T &:= N(T) := \{x \in X \colon Tx = 0\}, \\ \operatorname{Im} T &:= R(T) := T(X) := \{Tx \colon x \in X\} \end{aligned}$$

der Kern (englisch „null space“) bzw. der Wertebereich (englisch „range“) von T.

2.9 Bemerkung. Es gilt

$$\|Tx\|_Y \le \|T\| \cdot \|x\|_X \quad (x \in X)$$

und

$$\|T\| = \inf \{C > 0 \colon\ \forall\, x \in X \colon\ \|Tx\|_Y \le C\,\|x\|_X\}.$$

2.10 Definition und Satz. *Sei X ein $\mathbb{K}$-Vektorraum, und sei $L = \{p_\lambda \colon \lambda \in \Lambda\}$ eine Familie von Seminormen $p_\lambda \colon X \to [0,\infty)$. Definiere $\mathcal{T}$ als das System aller Mengen $U \subset X$, für welche gilt*

$$\forall\, x \in U\ \exists\, r \in \mathbb{N}\ \exists\, \lambda_1, \ldots, \lambda_r \in \Lambda\ \exists\, \varepsilon > 0 \colon\ B^{(\lambda_1)}(x,\varepsilon) \cap \cdots \cap B^{(\lambda_r)}(x,\varepsilon) \subset U.$$

Dabei sei $B^{(\lambda_j)}(x,\varepsilon) := \{y \in X : p_{\lambda_j}(x-y) < \varepsilon\}$ die Kugel um x mit Radius ε bzgl. der Seminorm p_{λ_j}. Dann ist $\mathcal{T}$ eine Topologie auf X, die sog. lokalkonvexe Topologie zu L auf X. Dies ist die gröbste Topologie auf X, für die jede der Abbildungen $p_\lambda(\cdot - x_0)$ mit $\lambda \in \Lambda$ und $x_0 \in X$ stetig von X nach $\mathbb{R}$ ist, vgl. Definition 1.2. Eine Subbasis von $\mathcal{T}$ ist gegeben durch

$$\{B^{(\lambda)}(x,r) \colon \lambda \in \Lambda,\ x \in X,\ r > 0\}.$$

[[Das zeigt man analog zum Beweis von Satz 2.6.]]

2.11 Definition. Sei X $\mathbb{K}$-VR und $\mathcal{T}$ eine Topologie auf X. Dann heißt $(X,\mathcal{T})$ ein topologischer Vektorraum , falls die Abbildungen

$$\begin{aligned} s &\colon X \times X \to X, & (x_1,x_2) &\mapsto x_1 + x_2 \\ m &\colon \mathbb{K} \times X \to X, & (\alpha,x) &\mapsto \alpha x \end{aligned}$$

beide stetig sind.

2.12 Bemerkung. Normierte Räume sind topologische Vektorräume.

2.13 Definition. a) Sei X ein $\mathbb{K}$-Vektorraum und $X^* := \{f \colon X \to \mathbb{K},\ f \text{ linear }\}$ der algebraische Dualraum von X. Für $Y \subset X^*$ bezeichnet man die Y-schwache Topologie $\mathcal{T}(Y)$ (vgl. Definition 1.2) auf X auch mit $\sigma(X,Y)$.

b) Sei X topologischer Vektorraum und $X' \subset X^*$ der topologische Dualraum. Dann heißt $\sigma(X, X')$ die schwache Topologie auf X und $\sigma(X', X)$ die schwach-*-Topologie auf X'. (Wir werden später sehen, dass man X als Teilmenge von X'' auffassen kann.) Für die schwache Konvergenz schreibt man $x_n \rightharpoonup x$ oder $x_n \xrightarrow{w} x$, für die schwach-*-Konvergenz schreibt man $f_n \overset{*}{\rightharpoonup} f$ oder $f_n \xrightarrow{w^*} f$ in X'.

Als Warnung vermerken wir, dass die schwache Topologie eines unendlichdimensionalen Raumes im Allgemeinen nicht metrisierbar ist. Wir können in Zukunft also nicht davon ausgehen, dass jeder von uns benötigte Raum eine Metrik besitzt, die zu seiner Topologie passt.

2.14 Bemerkung. Sei X ein topologischer Vektorraum. Dann gilt:

a) Die Verschiebung einer in X offenen Menge um einen konstanten Vektor ergibt in X wieder eine offene Menge.

b) Wenn $A \subset X$ eine offene Teilmenge ist und $B \subset X$ eine beliebige Teilmenge, dann ist $A + B$ offen in X.

2.15 Definition (Quotientenraum für topologische Vektorräume)**.** Sei $(X, \mathcal{T})$ ein topologischer Vektorraum (nicht unbedingt Hausdorffsch), und $M \subset X$ ein Untervektorraum, sowie $X/M := \{[x] = x+M : x \in X\}$ der algebraische Quotientenraum. Wir schreiben Φ für die (lineare) Quotientenabbildung,

$$\Phi: x \mapsto [x], \qquad \Phi: X \to X/M,$$

und wir statten X/M mit der feinsten Topologie aus, für die $\Phi: X \to X/M$ stetig ist.

2.16 Bemerkung. Die oben definierte Topologie auf X/M ist gegeben durch

$$\mathcal{T}_M := \{V \subset X/M : \Phi^{-1}(V) \subset X \text{ offen}\}.$$

Somit besteht diese Topologie aus genau jenen Mengen $H + M \subset X/M$, für die $H + M$ eine in X offene Menge ist. Und aus Bemerkung 2.14 bekommen wir dann: wenn $H \subset X$ offen in X ist, dann ist $\Phi(H)$ offen in X/M, d.h. die Abbildung Φ ist offen.

[[Man sieht direkt aus der Definition, dass $\mathcal{T}_M$ eine Topologie auf X/M ist und dass Φ bzgl. dieser Topologie stetig ist. Falls andererseits $\Phi : (X, \mathcal{T}) \to (X/M, \mathcal{T}')$ für eine beliebige Topologie $\mathcal{T}'$ stetig ist, so folgt für alle $V \in \mathcal{T}'$ schon $\Phi^{-1}(V) \in \mathcal{T}$ und damit $V \in \mathcal{T}_M$, d.h. $\mathcal{T}_M$ ist die maximale Topologie mit dieser Eigenschaft.]]

Wie benötigen noch eine Eigenschaft dieser Quotientenräume, welche hier nicht bewiesen wird.

2.17 Bemerkung. Sei X ein Vektorraum mit einer Seminorm $p(\cdot)$, ausgestattet mit der davon gemäß Definition 2.6 erzeugten Topologie; und es sei $M \subset X$ ein Untervektorraum. Dann ist der Quotientenraum X/M Hausdorffsch genau dann, wenn M eine abgeschlossene Teilmenge von X ist.

2.18 Definition und Satz (Quotientenraum für normierte Räume)**.** *Sei X normierter Raum, $M \subset X$ ein Untervektorraum, und X/M der Quotientenraum. Dann ist*

$$\|[x]\| := \operatorname{dist}(x, M) := \inf_{y \in M} \|x - y\|$$

eine Seminorm auf X/M. Falls M abgeschlossen ist, so ist $\|[\cdot]\|$ eine Norm und $(X/M, \|[\cdot]\|)$ ein normierter Raum. Falls X Banachraum ist und M abgeschlossen ist, so ist auch X/M Banachraum.

Beweis. Nur die letzte Aussage folgt nicht durch direktes Nachrechnen. Sei $([x_n])_{n\in\mathbb{N}}$ Cauchyfolge in X/M.

(i) Übergang zur Teilfolge: Da $\|([x_n]) - ([x_m])\| \longrightarrow 0 \quad (n, m \to \infty)$, existiert eine Teilfolge $([x_{n_k}])_{k\in\mathbb{N}}$ von $([x_n])_{n\in\mathbb{N}}$, so dass $\|[x_{n_k}] - [x_{n_{k+1}}]\| < 2^{-(k+1)}$. Schreibe wieder $[x_k]$ statt $[x_{n_k}]$.

(ii) Wahl einer Cauchyfolge in X: Wir wählen induktiv eine Folge von Repräsentanten $z_k \in [x_k]$, beginnend mit $z_1 := x_1$. Sei $z_k = x_k + m_k$ mit $m_k \in M$ bereits gewählt. Dann wählt man $\widetilde{m}_{k+1} \in M$ so, dass

$$\|x_{k+1} - x_k + \widetilde{m}_{k+1}\| \le \|[x_{k+1}] - [x_k]\| + 2^{-k}$$

gilt. Setzt man $z_{k+1} := x_{k+1} + m_{k+1} := x_{k+1} + \widetilde{m}_{k+1} + m_k$, so erhält man

$$\|z_{k+1} - z_k\| = \|x_{k+1} - x_k + \widetilde{m}_{k+1}\| \le \|[x_{k+1}] - [x_k]\|_{X/M} + 2^{-k} \quad (k \in \mathbb{N}).$$

Damit

$$\begin{aligned}
\|z_{k+m} - z_k\|_X &\le \sum_{\ell=k+1}^{k+m} \|z_\ell - z_{\ell-1}\|_X \\
&\le \sum_{\ell=k+1}^{k+m} \Big(\|[x_\ell] - [x_{\ell-1}]\|_{X/M} + 2^{-\ell} \Big) \\
&\le \sum_{\ell=k+1}^{k+m} 2^{-\ell+1} \le 2^{-k+1},
\end{aligned}$$

d.h. $(z_k)_k \subset X$ ist eine Cauchyfolge. Setze $z := \lim_k z_k$. Wegen

$$\|[x_k] - [z]\|_{X/M} = \|[z_k] - [z]\|_{X/M} \le \|z_k - z\|_X \to 0$$

gilt $[x_k] \to [z]$ in X/M. □

2.19 Definition und Satz (direkte Summe). *Seien $(X_1, \|\cdot\|_1)$, $(X_2, \|\cdot\|_2)$ normierte Räume. Dann wird durch*

$$\|(x_1, x_2)\| = \|x_1\|_1 + \|x_2\|_2 \quad ((x_1, x_2) \in X_1 \times X_2)$$

eine Norm auf $X = X_1 \times X_2$ definiert. Mit dieser Norm heißt X die direkte Summe von X_1 und X_2, Schreibweise $X = X_1 \oplus X_2$. Falls X_1 und X_2 beide Banachräume sind, so ist auch $X_1 \oplus X_2$ ein Banachraum.

Beweis. Nur die letzte Aussage ergibt sich nicht durch Nachrechnen. Falls die Folge $(x^{(n)})_{n\in\mathbb{N}} = (x_1^{(n)}, x_2^{(n)})_{n\in\mathbb{N}}$ eine Cauchyfolge in $X_1 \oplus X_2$ ist, so sind auch die einzelnen Komponenten Cauchyfolgen und damit, falls X_1 und X_2 Banachräume sind, konvergent gegen x_1 bzw. x_2. Damit ist aber auch die Folge $(x^{(n)})_{n\in\mathbb{N}} \subset X_1 \oplus X_2$ konvergent gegen $x := (x_1, x_2)$. □

2.20 Definition ($\mathscr{L}^p$-Räume). Sei $(Z, \mathscr{A}, \mu)$ ein Maßraum.

a) Sei $1 \le p < \infty$. Definiere $\mathscr{L}^p(\mu)$ als die Menge aller messbaren Funktionen $f\colon Z \to \mathbb{C}$ mit

$$\|f\|_p := \Big(\int |f|^p \,\mathrm{d}\mu\Big)^{1/p} < \infty.$$

b) Für $p = \infty$ wird $\mathscr{L}^\infty(\mu)$ definiert als die Menge aller messbarer Funktionen $f\colon Z \to \mathbb{C}$, für welche es ein $C_f > 0$ gibt mit $\mu(\{z \in Z\colon |f(z)| > C_f\}) = 0$. Man spricht von μ-fast überall beschränkten Funktionen. Für $f \in \mathscr{L}^\infty(\mu)$ definiert man

$$\|f\|_\infty := \inf\Big\{C \in \mathbb{R}\colon \mu(\{z \in Z\colon |f(z)| > C\}) = 0\Big\}.$$

c) Für Funktionen $f\colon Z \to \mathbb{R}$ werden die entsprechenden Funktionenräume mit $\mathscr{L}^p(\mu; \mathbb{R})$ bezeichnet. Manchmal schreiben wir auch $\mathscr{L}^p(\mu; \mathbb{C})$ statt $\mathscr{L}^p(\mu)$.

Der folgende Satz aus der Analysis wird hier nicht bewiesen.

2.21 Satz. *a)* **(Höldersche Ungleichung)** *Für $1 \le p, q \le \infty$ mit $\frac{1}{p} + \frac{1}{q} = 1$ gilt*

$$\|f \cdot g\|_1 \le \|f\|_p \cdot \|g\|_q \quad (f \in \mathscr{L}^p(\mu),\ g \in \mathscr{L}^q(\mu)).$$

b) **(Minkowskische Ungleichung)** *Für $1 \le p \le \infty$ gilt*

$$\|f + g\|_p \le \|f\|_p + \|g\|_p \quad (f, g \in \mathscr{L}^p(\mu)).$$

c) Für $1 \le p \le \infty$ ist der Raum $\mathscr{L}^p(\mu)$ ein Vektorraum, und $\|\cdot\|_p$ definiert eine Seminorm (Halbnorm) auf $\mathscr{L}^p(\mu)$.

2.22 Definition und Satz (L^p-Räume)**.** *Sei $(Z, \mathscr{A}, \mu)$ ein Maßraum, und sei $1 \leq p \leq \infty$. Den Vektorraum $X := \mathscr{L}^p(\mu)$ versehen wir mit der Seminorm $\|\cdot\|_p$ und anschließend mit der sich daraus ergebenden Topologie. Definiere eine Menge $M \subset \mathscr{L}^p(\mu)$, bestehend aus all jenen Funktionen, für die die Seminorm den Wert Null annimmt. Dann ist M ein Untervektorraum, und auch eine abgeschlossene Teilmenge von $\mathscr{L}^p(\mu)$, denn $\{0\}$ ist abgeschlossen in $\mathbb{R}$ und $\|\cdot\|_p : \mathscr{L}^p(\mu) \to \mathbb{R}$ ist stetig per Definition.*

Dann definieren wir
$$L^p(\mu) := \mathscr{L}^p(\mu)/M$$
als Quotientenraum. Dieser besteht aus Äquivalenzklassen von Funktionen, die μ–fast überall übereinstimmen.

Mit dieser Konstruktion wird $(L^p(\mu), \|\cdot\|_p)$ zu einem Banachraum.

Falls das Maß durch das Lebesgue–Maß $\lambda|_\Omega$ auf einer messbaren Mengen $\Omega \subset \mathbb{R}^n$ gegeben ist, schreibt man auch $L^p(\Omega)$ statt $L^p(\lambda|_\Omega)$.

2.23 Beispiele. Die folgenden Beispiele sind Spezialfälle der oberen Definition.

a) Der Raum $\mathbb{R}^n$ oder $\mathbb{C}^n$ wird mit jeder der Normen

$$\|x\|_p := \left(\sum_{j=1}^{n} |\xi_j|^p\right)^{\frac{1}{p}} \quad (1 \leq p < \infty),$$
$$\|x\|_\infty := \max\Big\{|\xi_j|, j = 1, \dots, n\Big\}$$

zu einem Banachraum.

b) Definiere für $1 \leq p < \infty$ die ℓ^p-Räume durch

$$\ell^p := \Big\{x = (\xi_1, \xi_2, \dots) | \xi_j \in \mathbb{C}, \|x\|_p := \Big(\sum_{j=1}^{\infty} |\xi_j|^p\Big)^{1/p} < \infty\Big\},$$

und für $p = \infty$ den Raum ℓ^∞ durch

$$\ell^\infty := \Big\{x = (\xi_1, \xi_2, \dots) | \xi_j \in \mathbb{C}, \|x\|_\infty := \sup\{|\xi_j|, j \in \mathbb{N}\} < \infty\Big\}.$$

Dann ist ℓ^p für alle $1 \leq p \leq \infty$ ein Banachraum. Hier ist jeweils $Z = \mathbb{N}$ und μ das Zählmaß.

b) Hilberträume

Im folgenden sei stets $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$.

2.24 Definition. a) Ein $\mathbb{K}$-Vektorraum X, versehen mit einer Abbildung $\langle\cdot,\cdot\rangle : X \times X \to \mathbb{K}$, heißt ein Vektorraum mit Skalarprodukt oder ein Prähilbertraum, falls gilt:

(i) Für alle $y \in X$ ist die Abbildung $x \mapsto \langle x, y\rangle$ linear.

(ii) Für alle $x, y \in X$ gilt $\langle x, y\rangle = \overline{\langle y, x\rangle}$.

(iii) Für alle $x \in X$ gilt $\langle x, x\rangle \geq 0$. Es gilt $\langle x, x\rangle = 0$ genau dann, wenn $x = 0$.

b) Zwei Vektoren $x, y \in X$ heißen orthogonal (in Zeichen $x \perp y$), falls $\langle x, y\rangle = 0$ gilt. Eine Familie $\{x_i\}_{i\in I}$ von Vektoren heißt orthonormal, falls gilt:

$$\langle x_i, x_j\rangle = \delta_{ij} := \begin{cases} 1 & \text{falls } i = j, \\ 0 & \text{sonst.} \end{cases}$$

c) In einem Prähilbertraum $(X, \langle\cdot,\cdot\rangle)$ wird durch $\|x\| := \langle x, x\rangle^{1/2}$ die kanonische Norm definiert. Ein vollständiger Prähilbertraum heißt Hilbertraum.

2.25 Beispiele. a) Mit dem Skalarprodukt $\langle x, y\rangle := \sum_{j=1}^{n} x_j \overline{y_j}$ werden $\mathbb{R}^n$ und $\mathbb{C}^n$ zu einem Hilbertraum.

b) Der Raum $C([0, 1])$ wird mit dem Skalarprodukt

$$\langle f, g\rangle := \int_0^1 f(x)\overline{g(x)}\,\mathrm{d}x$$

zu einem Prähilbertraum.

c) Sei $(X, \mathcal{A}, \mu)$ ein Maßraum. Dann wird $L^2(\mu)$, versehen mit dem Skalarprodukt

$$\langle f, g\rangle := \int_X f(x)\overline{g(x)}\,\mathrm{d}\mu(x) \quad (f, g \in L^2(\mu))$$

zu einem Hilbertraum. Insbesondere ist $L^2(\Omega)$ für $\Omega \subset \mathbb{R}^n$ ein Hilbertraum.

Auch die folgenden elementaren Eigenschaften von Prähilberträumen werden als bekannt vorausgesetzt, sie können aber auch leicht direkt nachgerechnet werden.

2.26 Satz. *Sei X Prähilbertraum.*

a) (Satz von Pythagoras) *Seien $x, y \in X$ orthogonal. Dann gilt*

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 .$$

b) *Sei $\{x_n\}_{n=1}^N$ orthonormal. Dann gilt*

$$\|x\|^2 = \sum_{n=1}^{N} |\langle x, x_n\rangle|^2 + \Big\| x - \sum_{n=1}^{N} \langle x, x_n\rangle\, x_n \Big\|^2 .$$

c) (Besselsche Ungleichung). *Sei* $\{x_n\}_{n=1}^N$ *orthonormal. Dann ist*

$$\|x\|^2 \geq \sum_{n=1}^N |\langle x, x_n\rangle|^2 \quad (x \in X).$$

d) (Cauchy–Schwarz-Ungleichung). *Es gilt*

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

e) (Parallelogramm-Identität) *Für* $x, y \in X$ *gilt*

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 .$$

f) (Polarisationsformel) *Für* $x, y \in X$ *gilt*

$$\langle x, y\rangle = \begin{cases} \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2), & \textit{falls } \mathbb{K} = \mathbb{R}, \\ \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2 + \mathrm{i}\,\|x+\mathrm{i}y\|^2 - \mathrm{i}\,\|x-\mathrm{i}y\|^2), & \textit{falls } \mathbb{K} = \mathbb{C}. \end{cases}$$

Die Bedeutung der Polarisationsformel liegt daran, dass Identitäten nur für die Norm nachgerechnet werden müssen und dann automatisch für die Skalarprodukte gelten.

Die Summe von zwei Hilberträumen X und Y ist wieder ein Hilbertraum, wenn man das Skalarprodukt auf $X \oplus Y$ durch

$$\langle (x_1, y_1), (x_2, y_2)\rangle_{X\oplus Y} := \langle x_1, x_2\rangle_X + \langle y_1, y_2\rangle_Y$$

definiert. Man beachte, dass die zugehörige Norm gegeben ist durch

$$\|(x, y)\|_{X\oplus Y} = \Big(\|x\|_X^2 + \|y\|_Y^2 \Big)^{1/2}.$$

Diese Norm ist äquivalent zur Norm aus Satz 2.19.

c) Der Approximationssatz und der Satz von Riesz für Hilberträume

Im folgenden sei $(X, \langle\cdot,\cdot\rangle)$ ein $\mathbb{K}$-Hilbertraum.

2.27 Definition. a)$M \subset X$ heißt konvex, falls gilt

$$\forall\, x, y \in M \;\forall\, \alpha \in [0,1]\colon \alpha x + (1-\alpha) y \in M.$$

b) Zu $M \subset X$ heißt

$$M^\perp := \{x \in X : \forall\, y \in M\colon \langle x, y\rangle = 0\}$$

das orthogonale Komplement von M in X.

2.28 Lemma. *Es ist $M^\perp$ ein abgeschlossener linearer Teilraum von X.*

Beweis. Die Teilraumeigenschaft ist klar. Für die Abgeschlossenheit vermerken wir

$$M^\perp = \bigcap_{y^* \in M} \{y^*\}^\perp,$$

sodass es genügt nachzuweisen, dass jede Menge $\{y^*\}^\perp$ abgeschlossen ist. Wir definieren eine lineare Abbildung

$$T_{y^*}\colon X \to \mathbb{K}, \qquad T_{y^*}\colon x \mapsto \langle x, y^*\rangle\,.$$

Dann ist $\{y^*\}^\perp = \ker T_{y^*}$. Wegen der Ungleichung von Cauchy–Schwarz ist T_{y^*} eine beschränkte Abbildung vom normierten Raum X in den normierten Raum $\mathbb{K}^1$, und aufgrund von Satz 2.7 ist T_{y^*} stetig. Bei jeder stetigen Abbildung ist das Urbild einer abgeschlossenen Menge wieder abgeschlossen. Nun ist aber $\{0\}$ eine abgeschlossene Teilmenge von $\mathbb{K}^1$, also ist auch $\ker T_{y^*} = T_{y^*}^{-1}(\{0\})$ abgeschlossen in X. □

2.29 Bemerkung. a) Es ist $M^\perp = \overline{M}^\perp = (\operatorname{span} M)^\perp$.

b) Es gilt $M \cap M^\perp \subset \{0\}$. Denn sei $x \in M \cap M^\perp$. Dann ist $\langle x, x\rangle = 0$, d.h. $x = 0$. Insbesondere gilt $M \cap M^\perp = \{0\}$, falls $0 \in M$ (z.B. falls M ein Untervektorraum von X ist).

2.30 Satz (Approximationssatz). *Sei X ein Hilbertraum, $M \subset X$ nichtleer, konvex und abgeschlossen, sowie $z_0 \in X$. Dann existiert genau ein $x \in M$ mit $\|x - z_0\| = \operatorname{dist}(z_0, M) := \inf\{\|y - z_0\| : y \in M\}$.*

Beweis. Ohne Einschränkung sei $z_0 = 0$ (betrachte die verschobene Menge $M - z_0$). Sei $d := \inf\{\|y\| : y \in M\}$. Wähle eine Folge $(y_n)_{n\in\mathbb{N}} \subset M$ mit $\|y_n\| \longrightarrow d$.

Unter Verwendung der Parallelogrammgleichung folgt

$$\begin{aligned}\|y_n - y_m\|^2 &= 2\,\|y_n\|^2 + 2\,\|y_m\|^2 - 4 \cdot \Big\| \underbrace{\frac{y_n + y_m}{2}}_{\in M} \Big\|^2 \\ &\le 2\,\|y_n\|^2 + 2\,\|y_m\|^2 - 4d^2 \longrightarrow 0,\end{aligned}$$

da $\|y_n\| \longrightarrow d$. Daher ist $(y_n)_{n\in\mathbb{N}}$ Cauchyfolge. Da M vollständig ist, existiert $x := \lim_n y_n \in M$. Es gilt $\|x\| = \lim_n \|y_n\| = d$. Damit folgt $\|x\| \le \|y\|$ für alle $y \in M$.

Eindeutigkeit: Sei $\|x_1\| \le \|y\|, \|x_2\| \le \|y\|$ für jedes $y \in M$. Dann ist

$$\begin{aligned}\|x_1 - x_2\|^2 &= 2\|x_1\|^2 + 2\|x_2\|^2 - \|x_1 + x_2\|^2 \\ &= 2\underbrace{\Big(\|x_1\|^2 - \Big\|\frac{x_1+x_2}{2}\Big\|^2\Big)}_{\le 0} + 2\underbrace{\Big(\|x_2\|^2 - \Big\|\frac{x_1+x_2}{2}\Big\|^2\Big)}_{\le 0} \le 0.\end{aligned}$$

□

2.31 Satz (Projektionssatz)**.** *Seien X ein Hilbertraum und $M \subset X$ ein abgeschlossener linearer Teilraum von X. Dann existiert für alle $x \in X$ eine eindeutige Zerlegung $x = m + m'$ mit $m \in M, m' \in M^\perp$. Somit ist $X = M \oplus M^\perp$. Es ist $\|x - m\| = \min_{y\in M} \|x - y\|$.*

Beweis. Nach dem Approximationssatz existiert genau ein $m \in M$ mit $\|m - x\| = \inf\{\|y - x\| : y \in M\}$. Setze $m' := x - m$. Für alle $y \in M$ ist dann $\|m'\| \le \|y - x\| = \|m' - (y - m)\|$.

(i) Wir zeigen $m' \in M^\perp$. Es gilt $\|m'\|^2 \le \|m' + ty\|^2 \quad (t \in \mathbb{K}, y \in M)$. Andererseits ist

$$\|m' + ty\|^2 = \|m'\|^2 + 2\,\mathrm{Re}\,\langle m', ty\rangle + |t|^2 \|y\|^2.$$

Für alle $t \in \mathbb{R}$ gilt somit

$$|t|^2 \|y\|^2 + 2t\,\mathrm{Re}\,\langle m', y\rangle \ge 0,$$

also $\mathrm{Re}\,\langle m', y\rangle = 0$.

Im Falle $\mathbb{K} = \mathbb{C}$ ersetzt man t durch it und erhält

$$|t|^2 \|y\|^2 + 2t\,\mathrm{Im}\,\langle m', y\rangle \ge 0,$$

also $\mathrm{Im}\,\langle m', y\rangle = 0$.

(ii) Um die Eindeutigkeit zu zeigen, sei $x = m + m' = z + z'$ mit $m, z \in M, m', z' \in M^\perp$. Dann ist $m - z = z' - m' \in M \cap M^\perp = \{0\}$, d.h. $m = z, m' = z'$. □

2.32 Satz (von Riesz)**.** *Seien X ein Hilbertraum und $T \in X'$. Dann existiert genau ein $x_T \in X$ mit*

$$Tx = \langle x, x_T\rangle \quad (x \in X).$$

Es gilt $\|T\|_{X'} = \|x_T\|_X$. Die Abbildung $I_{\mathrm{Riesz}}\colon X' \to X,\ T \mapsto x_T$ ist bijektiv, isometrisch und konjugiert linear.

Beweis. **(i)** Konstruktion von x_T: Der Raum $M := \ker T = T^{-1}(\{0\})$ ist abgeschlossen als Urbild einer abgeschlossenen Menge unter der stetigen Abbildung T. Damit ist $X = M \oplus M^\perp$ nach Satz 2.31.

Falls $M = X$, so folgt $T = 0$, und wir wählen $x_T = I_{\text{Riesz}}(T) := 0$.

Sei jetzt $M \neq X$. Wähle $y \in M^\perp \setminus \{0\}$. Wegen $M \cap M^\perp = \{0\}$ ist dann $Ty \neq 0$. Setze $x_T = I_{\text{Riesz}}(T) := \frac{\overline{Ty}}{\|y\|^2} \cdot y$. Dann gilt $\|x_T\| = \frac{|Ty|}{\|y\|}$, und wir merken uns, dass

$$Tx_T = \frac{\overline{Ty}}{\|y\|^2} \cdot Ty = \|x_T\|^2 .$$

Um zu zeigen, dass $Tx = \langle x, x_T \rangle$ für jedes $x \in X$ gilt, zerlegen wir x gemäß $X = M^\perp \oplus M$:

$$x = \frac{Tx}{Tx_T} x_T + \left(x - \frac{Tx}{Tx_T} x_T \right) =: x_\perp + x_\|.$$

Dies ist tatsächlich die angestrebte Zerlegung von x, denn es ist $x_\perp$ ein Vielfaches von $x_T \in M^\perp$, und es ist

$$Tx_\| = T\left(x - \frac{Tx}{Tx_T} x_T \right) = Tx - \frac{Tx}{Tx_T} Tx_T = 0,$$

also $x_\| \in \ker T = M$.

Damit haben wir dann tatsächlich

$$\begin{aligned} \langle x, x_T \rangle &= \langle x_\perp, x_T \rangle + \langle x_\|, x_T \rangle = \frac{Tx}{Tx_T} \langle x_T, x_T \rangle + 0 = \frac{Tx}{Tx_T} \|x_T\|^2 \\ &= Tx, \end{aligned}$$

wie gewünscht.

(ii) I_{Riesz} ist Isometrie: Nach Cauchy–Schwarz ist $\|T\|_{X'} = \sup_{\|x\| \le 1} |\langle x, x_T \rangle| \le \|x_T\|$. Andererseits haben wir $\|T\|_{X'} \ge |T(\frac{x_T}{\|x_T\|})| = \|x_T\|$.

(iii) $I_{\text{Riesz}}(T)$ ist eindeutig: Angenommen, es gäbe zusätzlich zu x_T noch ein $\widetilde{x}_T$ mit $Tx = \langle x, x_T \rangle = \langle x, \widetilde{x}_T \rangle \quad (x \in X)$. Dann gilt

$$0 = \langle x, x_T - \widetilde{x}_T \rangle \quad (x \in X).$$

Wähle $x = x_T - \widetilde{x}_T$ und erhalte $\|x_T - \widetilde{x}_T\| = 0$.

(iv) I_{Riesz} ist konjugiert linear: Sei $T = \alpha_1 T_1 + \alpha_2 T_2$. Dann ist einerseits $Tx = \langle x, x_T \rangle$, und andererseits

$$\begin{aligned} Tx &= \alpha_1 T_1 x + \alpha_2 T_2 x = \alpha_1 \langle x, x_{T_1} \rangle + \alpha_2 \langle x, x_{T_2} \rangle \\ &= \langle x, \overline{\alpha}_1 x_{T_1} \rangle + \langle x, \overline{\alpha}_2 x_{T_2} \rangle = \langle x, \overline{\alpha}_1 x_{T_1} + \overline{\alpha}_2 x_{T_2} \rangle , \end{aligned}$$

also $I_{\text{Riesz}}(T) = x_T = \overline{\alpha_1} x_{T_1} + \overline{\alpha_2} x_{T_2} = \overline{\alpha_1} I_{\text{Riesz}}(T_1) + \overline{\alpha_2} I_{\text{Riesz}}(T_2)$.

(v) I_{Riesz} ist surjektiv: Zu $y \in X$ sei $T_y x := \langle x, y \rangle$. Dann ist $|T_y x| \le \|y\| \cdot \|x\|$, d.h. T_y stetig und damit $T_y \in X'$ und $I_{\text{Riesz}}(T_y) = y$.

(vi) I_{Riesz} ist injektiv: Aus $\|T\|_{X'} = \|x_T\|_X$ ergibt sich direkt, dass der Kern der linearen Abbildung $I_{\text{Riesz}} \colon T \mapsto x_T$ nur das Nullfunktional enthalten kann. $\square$

d) Orthonormalbasen

2.33 Definition. (i) Sei $\mathcal{M}$ eine Menge. Eine Relation $\prec$ auf $\mathcal{M}$ heißt Halbordnung, falls für A, B, $C \in \mathcal{M}$ gilt:

$$\begin{aligned} &A \prec A, \\ &A \prec B,\ B \prec A \Longrightarrow A = B, \\ &A \prec B,\ B \prec C \Longrightarrow A \prec C. \end{aligned}$$

Beachte, dass $A \prec B$ oder $B \prec A$ nicht für alle $A, B \in M$ gelten muss.

(ii) Eine Menge $Q \subset \mathcal{M}$ heißt total geordnet oder eine Kette, falls für alle $A, B \in Q$ gilt: $A \prec B$ oder $B \prec A$.

(iii) Ein Element $A \in \mathcal{M}$ heißt obere Schranke für $S \subset \mathcal{M}$, falls $B \prec A$ für alle $B \in S$.

(iv) Ein Element $M \in \mathcal{M}$ heißt maximal, falls aus $M \prec A$ folgt $M = A$.

2.34 Beispiel. Sei $X \neq \emptyset$ eine Menge, und $\mathcal{M}$ eine nichtleere Teilmenge der Potenzmenge von X. Dann definieren wir, dass $A \prec B$ genau dann gilt, wenn A, $B \in \mathcal{M}$ und $A \subset B$.

2.35 Lemma (von Zorn). *Sei $\mathcal{M}$ eine nichtleere Menge mit Halbordnung, für welche jede Kette eine obere Schranke in $\mathcal{M}$ besitzt. Dann besitzt $\mathcal{M}$ ein maximales Element.*

Dieses Lemma ist eigentlich ein Axiom und äquivalent zum Wohlordnungssatz, welcher wiederum äquivalent zum Auswahlaxiom ist. Die Formulierung dieser Axiome und der Beweis der Äquivalenz werden hier aber weggelassen.

2.36 Definition. Sei X ein Hilbertraum. Eine Teilmenge $S \subset X$ heißt Orthonormalbasis oder vollständiges orthonormales System, falls S eine maximale orthonormale Teilmenge von X ist (maximal bezüglich Mengeninklusion). Man spricht auch von Hilbertraumbasis.

2.37 Satz. *Jeder nichttriviale Hilbertraum besitzt eine Orthonormalbasis.*

Beweis. Sei $\mathscr{S}$ die Menge aller orthonormalen Teilmengen des Hilbertraumes X. Dann ist $\mathscr{S} \neq \emptyset$ da $\{\frac{x}{\|x\|}\} \in \mathscr{S}$ für jedes $x \in X \setminus \{0\}$.

Sei $\{S_\alpha\}_{\alpha\in\mathcal{A}}$ eine Kette in $\mathscr{S}$, d.h. für $\alpha, \beta \in \mathcal{A}$ gilt $S_\alpha \subset S_\beta$ oder $S_\beta \subset S_\alpha$. Setze $S_0 := \bigcup_{\alpha\in\mathcal{A}} S_\alpha \subset X$. Dann ist $S_0 \supset S_\alpha$ für alle $\alpha \in \mathcal{A}$. Zu $x, y \in S_0$ existiert ein $\alpha \in \mathcal{A}$ mit $x, y \in S_\alpha$, d.h. S_0 ist orthonormal, also $S_0 \in \mathscr{S}$. Damit ist S_0 eine obere Schranke zu $\{S_\alpha\}_{\alpha\in\mathcal{A}}$. Nach dem Lemma von Zorn existieren maximale Elemente in $\mathscr{S}$. □

2.38 Lemma. *Seien X ein Prähilbertraum und $\{e_n : n \in \mathbb{N}\}$ ein Orthonormalsystem. Dann gilt*

$$\sum_{n=1}^{\infty} |\langle x, e_n\rangle \overline{\langle y, e_n\rangle}| < \infty \quad (x, y \in X).$$

Beweis. Für alle $N \in \mathbb{N}$ gilt nach der Cauchy–Schwarz-Ungleichung in $\mathbb{R}^N$ und der Besselschen Ungleichung

$$\sum_{n=1}^{N} |\langle x, e_n\rangle \overline{\langle y, e_n\rangle}| \leq \Big(\sum_{n=1}^{N} |\langle x, e_n\rangle|^2\Big)^{1/2} \cdot \Big(\sum_{n=1}^{N} |\langle y, e_n\rangle|^2\Big)^{1/2} \leq \|x\| \cdot \|y\|\,.$$

Mit $N \longrightarrow \infty$ erhält man die Behauptung. □

2.39 Satz. *Seien X ein Hilbertraum und $S = \{e_n : n \in \mathbb{N}\}$ ein abzählbares Orthonormalsystem.*

a) Für alle $x \in X$ konvergiert die Reihe $\sum_{n\in\mathbb{N}} \langle x, e_n\rangle\, e_n$ in X.

b) Sei $c_n \in \mathbb{C}$ für $n \in \mathbb{N}$ mit $\sum_{n\in\mathbb{N}} |c_n|^2 < \infty$. Dann konvergiert die Reihe $\sum_{n\in\mathbb{N}} c_n\, e_n$ in X.

c) $x - \sum_{n\in\mathbb{N}} \langle x, e_n\rangle\, e_n \in S^\perp$.

Beweis. a) Nach der Besselschen Ungleichung (Satz 2.26 c)) gilt für alle $N \in \mathbb{N}$

$$\sum_{n=1}^{N} |\langle x, e_n\rangle|^2 \leq \|x\|^2\,.$$

Also ist $\sum_{n\in\mathbb{N}} |\langle x, e_n\rangle|^2 < \infty$. Nach dem Satz von Pythagoras gilt

$$\Big\|\sum_{n=N}^{M} \langle x, e_n\rangle\, e_n\Big\|^2 = \sum_{n=N}^{M} |\langle x, e_n\rangle|^2 \longrightarrow 0 \quad (N, M \longrightarrow \infty).$$

Also ist $\left(\sum_{n=1}^{N} \langle x, e_n\rangle e_n\right)_{N\in\mathbb{N}}$ eine Cauchyfolge, und $y := \sum_{n=1}^{\infty} \langle x, e_n\rangle e_n \in X$ existiert.

b) wurde im Beweis von a) mitbewiesen.

c) Für $m \in \mathbb{N}$ gilt

$$\left\langle x - \sum_{n=1}^{\infty} \langle x, e_n\rangle e_n, e_m \right\rangle = \langle x, e_m\rangle - \sum_{n=1}^{\infty} \langle x, e_n\rangle \langle e_n, e_m\rangle = 0$$

wegen $\langle e_n, e_m\rangle = \delta_{nm}$. Man beachte hier, dass nach a) und wegen der Stetigkeit der Abbildung $X \to \mathbb{K}$, $y \mapsto \langle y, e_m\rangle$ das Skalarprodukt in die Summe gezogen werden darf. □

2.40 Satz. *Seien X ein Hilbertraum und $S = \{e_n : n \in \mathbb{N}\} \subset X$ ein abzählbares Orthonormalsystem. Dann sind äquivalent:*

(i) *S ist Orthonormalbasis.*

(ii) *$S^\perp = \{0\}$.*

(iii) *Für alle $y \in X$ gilt $y = \sum_{n\in\mathbb{N}} \langle y, e_n\rangle e_n$.*

(iv) *Für alle $x, y \in X$ gilt $\langle x, y\rangle = \sum_{n\in\mathbb{N}} \langle x, e_n\rangle \overline{\langle y, e_n\rangle}$.*

(v) *(Parsevalsche Gleichung) Es gilt*

$$\|x\|^2 = \sum_{n\in\mathbb{N}} |\langle x, e_n\rangle|^2 \quad (x \in X).$$

Beweis. (i)$\Longrightarrow$(ii). Falls ein $x \in S^\perp \setminus \{0\}$ existiert, so ist $S \cup \{\frac{x}{\|x\|}\}$ ein Orthonormalsystem.

(ii)$\Longrightarrow$(iii). Satz 2.39 c).

(iii)$\Longrightarrow$(iv). Die Reihen konvergieren absolut nach Lemma 2.38 und Satz 2.39 a), man darf also einsetzen.

(iv)$\Longrightarrow$(v). Setze $x = y$.

(v)$\Longrightarrow$(i). Falls S nicht maximal ist, wählen wir ein $x \in S^\perp$ mit $\|x\| = 1$. Es ergibt sich $\sum_{n\in\mathbb{N}} |\langle x, e_n\rangle|^2 = 0$, Widerspruch zu (v). □

2.41 Bemerkung. Ein topologischer Raum heißt nach Definition 1.4 separabel, wenn er eine abzählbare dichte Teilmenge besitzt. Da jeder Hilbertraum ein metrischer Raum ist, ergibt sich der topologische Abschluss einer Menge durch Hinzufügen aller Häufungspunkte (vgl. Bemerkung 1.13), was in der Anwendung häufig praktischer ist.

Ein normierter Raum X ist genau dann separabel, wenn es ein abzählbares linear unabhängiges $S \subset X$ gibt mit $\overline{\operatorname{span} S} = X$. Insbesondere ist ein Hilbertraum genau dann separabel, wenn er eine höchstens abzählbare Orthonormalbasis besitzt.

Um das zu sehen, betrachtet man alle Linearkombinationen der Form $\sum_{k=1}^n a_k s_k$, $a_k \in \mathbb{Q} + i\mathbb{Q}$, $s_k \in S$.

2.42 Satz. *Sei X ein unendlich-dimensionaler separabler Hilbertraum mit Hilbertraumbasis $\{e_i\}_{i\in\mathbb{N}}$. Dann ist die Abbildung $X \to \ell^2$, $x \mapsto (\langle x, e_i\rangle)_{i\in\mathbb{N}}$ ein isometrischer Isomorphismus von Hilberträumen (d.h. linear, bijektiv und isometrisch).*

Beweis. Die Linearität ist klar, Isometrie und damit Injektivität nach Satz 2.40 (v), die Surjektivität nach Satz 2.39 b). □

2.43 Beispiele (Hilbertraumdimension). a) Der Raum $L^2(\Omega)$ mit $\Omega \subset \mathbb{R}^n$ offen ist ein separabler Hilbertraum. Speziell ist durch $e_n(x) := \frac{1}{\sqrt{2\pi}} e^{inx}$ $(n \in \mathbb{Z})$ eine Hilbertraumbasis des Hilbertraums $L^2((-\pi,\pi),\mathbb{C})$ gegeben (Fourierreihen). In diesem Fall erhält man

$$\langle f, e_n\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}dx =: \hat{f}(n) \quad (n \in \mathbb{Z}),$$

die Fourier-Koeffizienten von f. Die Parsevalsche Gleichung besagt dann

$$\|f\|_{L^2(\mathbb{R})} = \|(\hat{f}(n))_{n\in\mathbb{Z}}\|_{\ell^2(\mathbb{Z})},$$

und die Abbildung $f \mapsto (\hat{f}(n))_{n\in\mathbb{Z}}$ ist ein Hilbertraum-Isomorhpismus.

Eine Orthonormalbasis von $L^2(\mathbb{R})$ ist durch die Hermite-Funktionen gegeben, die die Form $\psi_n(x) = c_n(x - \frac{d}{dx})^n e^{-x^2/2}$ mit geeigneter Konstante c_n besitzen.

b) Es gibt auch Prähilberträume mit überabzählbaren Orthonormalsystemen.

[[Dies zeigt das folgende Beispiel: Sei Y die Menge aller Funktionen $f\colon \mathbb{R} \to \mathbb{R}$ mit $f|_{(-T,T)} \in L^2(-T,T)$ für alle $T > 0$, für welche

$$\|f\|_1 := \left(\lim_{T\to\infty} \frac{1}{T}\int_{-T}^{T} |f(x)|^2\, \mathrm{d}x\right)^{1/2}$$

existiert. Dann ist $\|\cdot\|_1$ eine Halbnorm auf Y, aber keine Norm (so gilt etwa für die charakteristische Funktion $f := \chi_{(-1,1)}$ zwar $\|f\|_1 = 0$, aber $f \neq 0$).

Sei $N := \{f \in Y\colon \|f\|_1 = 0\}$. Dann ist N ein abgeschlossener Unterraum. Sei $X := Y/N$ der Quotientenraum mit induzierter Norm $\|[f]\|_2 := \inf_{g\in N} \|f - g\|_1$ $(f \in Y)$. Durch

$$\langle [f], [g]\rangle_2 := \lim_{T\to\infty} \frac{1}{T}\int_{-T}^{T} f_1(x)g_1(x)\, \mathrm{d}x \quad (f_1 \in [f], g_1 \in [g])$$

wird X zu einem Prähilbertraum, und es gilt $\langle [f], [f]\rangle_2 = \|[f]\|_2$.

Zu $\alpha > 0$ definiere $v_\alpha(x) := \sin(\alpha x)(x \in \mathbb{R})$. Wegen

$$\frac{1}{T}\int_{-T}^{T} v_\alpha(x)^2\,\mathrm{d}x = \frac{1}{T}\left(T - \frac{\sin(s\alpha T)}{2\alpha}\right) \overset{T\to\infty}{\longrightarrow} 1$$

gilt $v_\alpha \in Y$ und $\|v_\alpha\|_1 = \|[v_\alpha]\|_2 = 1$. Andererseits erhält man für $\alpha \neq \beta$

$$\begin{aligned}\frac{1}{T}\int_{-T}^{T} \sin(\alpha r)\sin(\beta r)\,\mathrm{d}r &= \frac{1}{2T}\int_{-T}^{T} \cos((\alpha-\beta)r) - \cos((\alpha+\beta)r)\,\mathrm{d}r\\ &= \frac{1}{2T}\left[\frac{\sin((\alpha-\beta)r)}{\alpha-\beta} - \frac{\sin((\alpha+\beta)r)}{\alpha+\beta}\right]_{r=-T}^{r=T}\\ &= \frac{1}{T}\left[\frac{\sin((\alpha-\beta)T)}{\alpha-\beta} - \frac{\sin((\alpha+\beta)T)}{\alpha+\beta}\right] \overset{T\to\infty}{\longrightarrow} 0.\end{aligned}$$

Daher gilt $\langle [v_\alpha], [v_\beta]\rangle_2 = 0$ für $\alpha \neq \beta$, d.h. $\{v_\alpha\}_\alpha$ ist ein überabzählbares Orthonormalsystem in X.]]

2.44 Beispiele (Separabilität)**.** a) Der Raum $C([a,b])$ mit $\|\cdot\|_\infty$-Norm ist ein separabler Banachraum, denn die Polynome mit rationalen Koeffizienten liegen dicht. Die Räume $\mathbb{R}^n$ und ℓ^2 sind separabel nach Bemerkung 2.41. Ebenso separabel ist der Raum $L^2(\Omega)$ für eine offene Teilmenge $\Omega \subset \mathbb{R}^n$.

b) Der Banachraum ℓ^∞ ist ein Beispiel für einen nicht separablen Raum: Angenommen es existiert eine dichte Teilmenge $\{x^{(k)} \colon k \in \mathbb{N}\} \subset \ell^\infty$. Schreibe $x^{(k)} = (x_n^{(k)})_{n\in\mathbb{N}}$ und betrachte

$$y_k := \begin{cases} x_k^{(k)} + 1, & : |x_k^{(k)}| < 1,\\ 0 & : \text{ sonst.}\end{cases}$$

Dann gilt $|y_k| \leq 2$ und damit $y := (y_k)_{k\in\mathbb{N}} \in \ell^\infty$. Aber für alle $k \in \mathbb{N}$ haben wir $|y_k - x_k^{(k)}| \geq 1$ und damit $\|y - x^{(k)}\|_\infty \geq 1$, Widerspruch zur Dichtheit von $\{x^{(k)} \colon k \in \mathbb{N}\}$.

c) Auch der Raum $L^\infty(\Omega)$ mit $\Omega \subset \mathbb{R}^n$ offen ist ein Beispiel für einen nicht separablen Banachraum.

3. Hahn-Banach-Sätze und Reflexivität

3.1 Worum geht's? Die Sätze von Hahn-Banach können auf zwei verschiedene Arten formuliert und interpretiert werden: Zum einen als Fortsetzungssätze, bei welchen Fortsetzungen von Funktionalen gesucht werden, zum anderen als Trennungssätze, bei welchen Teilmengen durch den Wert eines Funktionals getrennt werden. Beide Versionen sind für Anwendungen wichtig. Beiden Versionen gemeinsam ist die Aussage, dass der topologische Dualraum eines normierten Raums hinreichend viele Funktionale enthält, um ausreichend Information über den ursprünglichen Raum zu erhalten. Als ein Beispiel einer Anwendung des Satzes von Hahn-Banach wird hier auch die Injektivität der sogenannten kanonischen Einbettung auftauchen. Diese bettet einen normierten Raum in seinen Bidualraum ein. Falls diese Einbettung auch surjektiv ist, heißt der Raum reflexiv, ein Begriff, der unter anderem im Bereich der schwachen Topologien und entsprechender Konvergenz- bzw. Kompaktheitsaussagen wichtig ist.

a) Hahn–Banach-Sätze

3.2 Satz (Fortsetzungssatz von Hahn–Banach, reelle Version)**.** *Sei X ein $\mathbb{R}-$Vektorraum und $p\colon X \to \mathbb{R}$ konvex, d.h. es gelte*

$$p(\alpha x + (1-\alpha)y) \le \alpha p(x) + (1-\alpha)p(y) \quad (\alpha \in [0,1], x, y \in X).$$

Sei ferner $L \subset X$ ein linearer Teilraum und $\lambda\colon L \to \mathbb{R}$ linear mit

$$\lambda(x) \le p(x) \quad (x \in L).$$

Dann existiert ein lineares $\Lambda\colon X \to \mathbb{R}$ mit $\Lambda|_L = \lambda$ und $\Lambda(x) \le p(x) \quad (x \in X)$.

Beweis. **(i)** Fortsetzung auf $L \oplus \mathbb{R}z$:

Sei $z \in X \setminus L$ fest gewählt, und definiere $\widetilde{L} := \operatorname{span}\{L, z\} = L \oplus \mathbb{R}z$. Wir setzen jetzt λ auf $\widetilde{L}$ fort.

Für $y_1, y_2 \in L$ und $\alpha, \beta > 0$ beliebig gilt:

$$\begin{aligned}\beta\lambda(y_1) + \alpha\lambda(y_2) &= \lambda(\beta y_1 + \alpha y_2) = (\alpha+\beta)\cdot\lambda\Big(\frac{\beta}{\alpha+\beta}\,y_1 + \frac{\alpha}{\alpha+\beta}\,y_2\Big)\\ &\le (\alpha+\beta)p\Big(\frac{\beta}{\alpha+\beta}(y_1-\alpha z) + \frac{\alpha}{\alpha+\beta}(y_2+\beta z)\Big)\\ &\le \beta p(y_1 - \alpha z) + \alpha p(y_2 + \beta z).\end{aligned}$$

Damit erhalten wir

$$\frac{1}{\alpha}\left[\lambda(y_1) - p(y_1 - \alpha z)\right] \le \frac{1}{\beta}\left[p(y_2 + \beta z) - \lambda(y_2)\right] \tag{3-1}$$

Wähle

$$\widetilde{\lambda}(z) := \alpha_0 \in \left[\sup_{y_1 \in L,\, \alpha > 0} \frac{\lambda(y_1) - p(y_1 - \alpha z)}{\alpha}, \inf_{y_2 \in L,\, \beta > 0} \frac{p(y_2 + \beta z) - \lambda(y_2)}{\beta}\right]$$

und definiere $\widetilde{\lambda}(\mu z + y) := \mu \widetilde{\lambda}(z) + \lambda(y)$ auf $\widetilde{L} = L \oplus \mathbb{R} \cdot z$.

$\widetilde{\lambda}$ ist linear auf $\widetilde{L}$ nach Definition, und es gilt

$$\widetilde{\lambda}(\mu z + y) = \mu \alpha_0 + \lambda(y) \le p(\mu z + y).$$

Denn für $\mu > 0$ gilt nach Wahl von α_0 die Abschätzung

$$p\,(y_2 + \beta z) \ge \lambda(y_2) + \beta \alpha_0.$$

Setze nun $y_2 := y$ und $\beta := \mu$. Den Fall $\mu < 0$ sieht man analog.

(ii) Fortsetzung auf X:

Sei $\mathscr{M}$ die Menge aller Abbildungen $m\colon M \to \mathbb{R}$ auf einem linearen Unterraum $M \supset L$, welche linear sind und für welche gilt $m|_L = \lambda$ und $m \le p|_M$.

Durch

$$m_1 \prec m_2 :\Longleftrightarrow M_1 \subset M_2,\ m_2|_{M_1} = m_1$$

wird $\mathscr{M}$ partiell geordnet. Sei $\{m_k\}$ eine Kette in $\mathscr{M}$. Deshalb ist dann $M := \bigcup_k M_k$ ein linearer Unterraum, und durch

$$m(x) := m_k(x) \quad (x \in M_k)$$

wird eine obere Schranke $m \in \mathscr{M}$ der Kette definiert. Nach dem Zornschen Lemma existiert ein maximales Element $\Lambda \in \mathscr{M}$.

Da Λ maximal ist, ist Λ auf ganz X definiert. Sonst existierte nach Schritt 1 eine Fortsetzung auf $D(\Lambda) \oplus \mathbb{R} \cdot z$ mit $z \in X \setminus D(\Lambda)$. □

3.3 Satz (Hahn–Banach, komplexe Version)**.** *Sei X ein $\mathbb{C}-$Vektorraum und $p\colon X \to \mathbb{R}$ eine Abbildung mit*

$$p(\alpha x + \beta y) \le |\alpha|\, p(x) + |\beta|\, p(y) \quad (\alpha, \beta \in \mathbb{C} \text{ mit } |\alpha| + |\beta| = 1).$$

Sei $L \subset X$ linearer Teilraum und $\lambda\colon L \to \mathbb{C}$ sei $\mathbb{C}$–linear mit $|\lambda(x)| \le p(x)$ $(x \in L)$. Dann existiert ein $\mathbb{C}$–lineares $\Lambda\colon X \to \mathbb{C}$ mit $\Lambda|_L = \lambda$ und $|\Lambda(x)| \le p(x) \quad (x \in X)$.

Beweis. Setze $\ell(x) := \operatorname{Re}\lambda(x)$. Dann ist $\ell\colon L \to \mathbb{R}$ eine $\mathbb{R}$–lineare Abbildung mit

$$\ell(x) \le |\lambda(x)| \le p(x) \quad (x \in L).$$

Weil λ $\mathbb{C}$–linear ist, haben wir $\ell(\mathrm{i}x) = -\operatorname{Im}\lambda(x)$, und somit ist $\lambda(x) = \ell(x) - \mathrm{i}\ell(\mathrm{i}x)$. Setze ℓ nach Satz 3.2 fort zu einem $\mathbb{R}-$linearen $\mathcal{L}\colon X \to \mathbb{R}$ mit $\mathcal{L}(x) \le p(x) \quad (x \in X)$. Dann ist

$$\Lambda(x) := \mathcal{L}(x) - \mathrm{i}\mathcal{L}(\mathrm{i}x)$$

$\mathbb{R}$–linear. Wegen

$$\Lambda(\mathrm{i}x) = \mathcal{L}(\mathrm{i}x) - \mathrm{i}\mathcal{L}(-x) = \mathcal{L}(\mathrm{i}x) + \mathrm{i}\mathcal{L}(x) = \mathrm{i}\Lambda(x)$$

ist Λ tatsächlich $\mathbb{C}$–linear.

Für $\theta := \arg\Lambda(x)$ gilt:

$$|\Lambda(x)| = e^{-\mathrm{i}\theta}\Lambda(x) = \Lambda(e^{-\mathrm{i}\theta}x) = \mathcal{L}(e^{-\mathrm{i}\theta}x) \le p(e^{-\mathrm{i}\theta}x) = p(x).$$

Hier wurde $\operatorname{Re}\Lambda = \mathcal{L}$ und $\Lambda(e^{-\mathrm{i}\theta}x) = |\Lambda(x)| \in \mathbb{R}$ verwendet. □

3.4 Korollar. *Sei X normiert, $L \subset X$ linearer Teilraum und $\lambda \in L'$. Dann existiert ein $\Lambda \in X'$ mit $\|\Lambda\|_{X'} = \|\lambda\|_{L'}$, $\Lambda|_L = \lambda$.*

Beweis. Sei $p(x) := \|\lambda\|_{L'} \cdot \|x\|$. Dann ist $|\lambda(x)| \le p(x) \quad (x \in L)$.

Nach Satz 3.3 existiert eine Fortsetzung Λ mit

$$|\Lambda(x)| \le \|\lambda\|_{L'} \cdot \|x\| \quad (x \in X),$$

d.h. $\Lambda \in X'$ und $\|\Lambda\|_{X'} \le \|\lambda\|_{L'}$. Wegen $\Lambda|_L = \lambda$ ist $\|\Lambda\|_{X'} = \|\lambda\|_{L'}$. □

3.5 Korollar. *Sei X normiert, $x_0 \in X \setminus \{0\}$ fest. Dann existiert ein $\Lambda \in X'$ mit $\Lambda(x_0) = \|x_0\|$ und $\|\Lambda\|_{X'} = 1$.*

Beweis. Definiere $L := \operatorname{span}(x_0)$, $\lambda\colon L \to \mathbb{K}$, $\lambda(\alpha x_0) := |\alpha|\,\|x_0\|$, und setze nach Korollar 3.4 fort. □

3.6 Korollar. *Sei X normiert, $M \subset X$ linearer Teilraum und $x_0 \in X$. Sei*

$$d := \inf_{y \in M} \|x_0 - y\| > 0.$$

Dann existiert ein $\Lambda \in X'$ mit $\|\Lambda\| = 1$, $\Lambda(x_0) = d$ und $\Lambda|_M = 0$.

Beweis. Definiere λ auf $L := M \oplus \operatorname{span}(x_0)$ durch $\lambda(y + \alpha x_0) := \alpha d$. Dann gilt

$$\begin{aligned}\|\lambda\|_{L'} &= \sup_{\substack{\alpha\in\mathbb{K}\\ y\in M}} \frac{|\alpha d|}{\|y+\alpha x_0\|} = \sup_{\substack{0\neq\alpha\in\mathbb{K}\\ y\in M}} \frac{|\alpha d|}{\|y+\alpha x_0\|}\\ &= \sup_{\substack{0\neq\alpha\in\mathbb{K}\\ z\in M}} \frac{|\alpha d|}{\|-\alpha z+\alpha x_0\|} = \sup_{\substack{0\neq\alpha\in\mathbb{K}\\ z\in M}} \frac{d}{\|x_0 - z\|} = \frac{d}{\inf_{z\in M}\|x_0-z\|} = \frac{d}{d} = 1.\end{aligned}$$

Die Behauptung folgt nun aus Korollar 3.4. □

Die letzte Aussage kann bereits als Trennungseigenschaft gesehen werden: Der Vektor x_0 wird durch das Funktional Λ vom Unterraum M getrennt. Etwas allgemeiner gilt z.B. folgender Trennungssatz, der hier nicht bewiesen werden soll. (Falls X Hilbertraum ist, folgt der Beweis sehr schnell aus Übungsaufgabe 4.2, im allgemeinen Fall verwendet man das sogenannte Minkowski-Funktional.)

3.7 Satz (Hahn–Banach, Trennungssatz-Version). *Sei X ein normierter Raum, $M \subset X$ eine abgeschlossene und konvexe Teilmenge, und sei $x_0 \in X \setminus M$. Dann existieren ein $\Lambda \in X'$ und ein $\alpha \in \mathbb{R}$ mit*

$$\operatorname{Re}\Lambda(y) < \alpha < \operatorname{Re}\Lambda(x_0) \quad (y \in M).$$

b) Dualräume und Reflexivität

3.8 Satz. *Seien X normierter Raum und Y Banachraum. Dann ist $L(X,Y)$ Banachraum. Insbesondere ist X' Banachraum.*

Beweis. Nur die Vollständigkeit ist nichttrivial. Sei $(A_n)_{n\in\mathbb{N}}$ Cauchyfolge in $L(X,Y)$. Dann ist $(A_n x)_{n\in\mathbb{N}}$ Cauchyfolge in Y für jedes $x \in X$, wie man sich überlegt. Aus den Analysisvorlesungen ist bekannt, dass Cauchyfolgen in normierten Räumen beschränkte Folgen sind, also existiert ein $C \in \mathbb{R}$ mit $\|A_n\| \le C$ für jedes n.

Setze $Ax := \lim_n A_n x \in Y$. Dann ist A offensichtlich linear. Wegen

$$\|Ax\|_Y = \lim_n \|A_n x\|_Y \le \lim_n \|A_n\| \cdot \|x\|_X \le C\,\|x\|_X$$

ist $A \in L(X,Y)$. Da $\|(A-A_n)x\|_Y = \lim_{m\to\infty} \|(A_m - A_n)x\|_Y$, erhalten wir

$$\|A - A_n\| = \sup_{x\neq 0} \frac{\|(A-A_n)x\|_Y}{\|x\|_X} = \sup_{x\neq 0} \lim_{m\to\infty} \frac{\|(A_m-A_n)x\|_Y}{\|x\|_X} \le \lim_{m\to\infty} \|A_m - A_n\|$$
$$< \varepsilon$$

für $n \ge n_0$, d.h. es gilt $A_n \longrightarrow A$ in $L(X,Y)$. □

3.9 Lemma. *Sei X ein normierter Raum. Die Abbildung $X \to X''$, $x \mapsto \widetilde{x}$ mit*

$$\widetilde{x}(\lambda) := \lambda(x) \quad (\lambda \in X')$$

ist linear und isometrisch. Insbesondere gilt

$$\|x\|_X = \sup_{\lambda \in X',\, \|\lambda\|_{X'}=1} |\lambda(x)| \quad (x \in X).$$

Beweis. Es gilt

$$(\widetilde{\alpha x + \beta y})(\lambda) = \lambda(\alpha x + \beta y) = \alpha\lambda(x) + \beta\lambda(y) = \alpha\widetilde{x}(\lambda) + \beta\widetilde{y}(\lambda),$$

d.h. die Abbildung $x \mapsto \widetilde{x}$ ist linear. Weiter ist

$$\|\widetilde{x}\|_{X''} = \sup_{\|\lambda\| \le 1} |\widetilde{x}(\lambda)| = \sup_{\|\lambda\| \le 1} |\lambda(x)| \le \sup_{\|\lambda\| \le 1} \|\lambda\| \cdot \|x\|_X \le \|x\|_X .$$

Nach Lemma 3.5 existiert zu jedem $x \in X$ ein $\lambda_0 \in X'$ mit $\|\lambda_0\|_{X'} = 1$ und $\lambda_0(x) = \|x\|_X$. Damit gilt $\|\widetilde{x}\|_{X''} = \sup_{\|\lambda\|_{X'} \le 1} |\lambda(x)| \ge |\lambda_0(x)| = \|x\|_X$. □

3.10 Definition. Ein normierter Raum X heißt reflexiv, falls die kanonische Einbettung $X \hookrightarrow X''$ aus Lemma 3.9 surjektiv ist.

3.11 Beispiele. a) Jeder Hilbertraum ist reflexiv nach dem Satz von Riesz.

b) Sei $1 < p < \infty$ und $(X, \mathscr{A}, \mu)$ ein Maßraum. Nach einem Satz von Riesz ist die Abbildung

$$T\colon L^q(\mu) \to (L^p(\mu))', \quad (Tg)(f) := \int fg\,\mathrm{d}\mu \tag{3-2}$$

ein isometrischer Isomorphismus von Banachräumen, wobei $q \in (1, \infty)$ definiert ist durch $p^{-1} + q^{-1} = 1$. Damit ist $L^p(\mu)$ für $1 < p < \infty$ reflexiv.

[[Sei $\Lambda \in (L^p(\mu))''$. Definiere das Funktional $\Lambda_1 := \Lambda \circ T \in (L^q(\mu))'$. Nach dem Satz von Riesz, angewendet auf den Raum $(L^q(\mu))'$, existiert genau eine Funktion $h \in L^p(\mu)$, so dass für alle $g \in L^q(\mu)$ der Wert $\Lambda_1(g)$ gegeben ist durch

$$\Lambda_1(g) = \int g \cdot h\,\mathrm{d}\mu. \tag{3-3}$$

Somit gilt für jedes $\lambda \in (L^p(\mu))'$

$\Lambda(\lambda) = \Lambda_1(T^{-1}\lambda)$	$\Lambda = \Lambda_1 \circ T^{-1}$
$= \int (T^{-1}\lambda) \cdot h\,\mathrm{d}\mu$	wegen (3-3)
$= \lambda(h)$	wegen (3-2)

$$= \widetilde{h}(\lambda).$$

Wir haben gesehen, dass $\Lambda = \widetilde{h}$ gilt, d.h. dass die Abbildung $h \mapsto \widetilde{h},\ X \to X''$, surjektiv ist. Die $L^p(\mu)$-Räume sind für $1 < p < \infty$ also reflexiv.]]

c) In der Situation von b) gilt die Isomorphie $\big(L^1(\mu)\big)' \equiv L^\infty(\mu)$, aber andererseits $L^1(\mu) \subsetneqq \big(L^\infty(\mu)\big)'$, d.h. $L^1(\mu)$ ist nicht reflexiv. Diese Aussage wird nicht bewiesen.

4. Lineare Operatoren: Grundbegriffe

4.1 Worum geht's? Eine der zentralen Begriffe der Funktionalanalysis ist der eines linearen Operators. Während im Endlich-Dimensionalen alle linearen Abbildungen stetig und damit beschränkt sind (und durch Matrizen dargestellt werden können), treten gerade in vielen Anwendungen wie etwa bei Differentialgleichungen unbeschränkte Operatoren auf. Ein typisches Beispiel ist - bei geeigneter Wahl der Räume - der Ableitungsoperator. In diesem Abschnitt geht es um grundlegende Begriffe und Eigenschaften wie das Spektrum eines Operators und Eigenschaften der Resolvente.

a) Operatoren und Spektrum

Im folgenden sei stets $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

4.2 Beispiel (Shift-Operatoren). a) Definiere den Rechtsshift $S_R \in L(\ell^2)$ durch

$$S_R\big((x_n)_{n\in\mathbb{N}}\big) := (y_n)_{n\in\mathbb{N}}, \quad y_n := \begin{cases} 0 & : n = 1, \\ x_{n-1} & : n > 1. \end{cases}$$

Dann ist S stetig mit Norm 1 (sogar eine Isometrie, d.h. es gilt $\|Sx\|_2 = \|x\|_2 \quad (x \in \ell^2)$), injektiv aber nicht surjektiv. Analog ist der Linksshift

$$S_L\big((x_n)_{n\in\mathbb{N}}\big) := (x_{n+1})_{n\in\mathbb{N}}$$

stetig mit Norm 1, surjektiv aber nicht injektiv.

b) Sei $\ell^2(\mathbb{Z};\mathbb{K}) := \{x\colon \mathbb{Z} \to \mathbb{K} : \|x\|_2 := (\sum_{n\in\mathbb{Z}} |x_n|^2)^{1/2} < \infty\}$ und S_R der Rechtsshift

$$S_R\big((x_n)_{n\in\mathbb{Z}}\big) := (x_{n-1})_{n\in\mathbb{Z}}.$$

Dann ist S_R ein Norm-Isomorphismus, d.h. S_R ist bijektiv, linear, S_R und S_R^{-1} sind stetig, und S_R ist eine Isometrie.

4.3 Beispiel (Ableitungsoperator). a) Sei $X := C^1([0,1])$ mit Norm $\|f\|_X := \sup\{|f(t)| + |f'(t)|\colon t \in [0,1]\}$ und $Y := C([0,1])$, versehen mit der Norm $\|f\|_Y := \|f\|_\infty := \sup\{|f(t)|\colon t \in [0,1]\}$. Dann sind X, Y Banachräume, und der Ableitungsoperator

$$T\colon X \to Y,\ f \mapsto f'$$

ist ein linearer stetiger Operator mit Norm 1.

b) Wählt man in a) für X die Supremumsnorm $\|f\|_\infty$, so ist der Ableitungsoperator T nicht mehr stetig, denn für $f_n(t) := t^n$ gilt $\|f_n\|_\infty = 1$ und $\|Tf_n\|_\infty = n$, d.h. $\|T\| = \infty$. Jetzt ist X nur noch normierter Vektorraum, aber nicht mehr vollständig.

Das letzte Beispiel (Teil b)) ist typisch: Hier ist der Definitionsbereich $D(T)$ des Operators T ein linearer dichter Teilraum eines Banachraums X, und T bildet nach X ab. In diesem Sinn ist T ein Operator „in“ X.

4.4 Definition. Seien X, Y normierte Räume.

a) Ein linearer Operator $T\colon X \supset D(T) \to Y$ ist eine lineare Abbildung vom Definitionsbereich $D(T) \subset X$ nach Y, wobei $D(T)$ ein linearer Unterraum von X ist. Die Menge $G(T) := \{(x, Tx)\colon x \in D(T)\}$ heißt der Graph von T.

b) Der Operator T heißt abgeschlossen, wenn $G(T)$ eine abgeschlossene Teilmenge von $X \oplus Y$ ist.

c) Der Operator T heißt abschließbar, wenn es einen abgeschlossenen linearen Operator $\overline{T}$ gibt mit $G(\overline{T}) = \overline{G(T)}$. Der Operator $\overline{T}$ heißt Abschließung oder der Abschluss von T.

4.5 Bemerkung. a) Wie üblich bei Abbildungen, ist die Stetigkeit eines linearen Operators $T\colon X \supset D(T) \to Y$ unter Verwendung der Relativtopologie auf $D(T)$ erklärt. Damit ist T genau dann stetig, wenn für alle Folgen $(x_n)_{n\in\mathbb{N}} \subset D(T)$ mit $x_n \longrightarrow 0$ gilt $Tx_n \longrightarrow 0$ (hierbei haben wir benutzt, dass eine Abbildung zwischen metrischen Räumen genau dann stetig ist, wenn sie folgenstetig ist).

b) Ein linearer Operator T ist genau dann abgeschlossen, wenn für alle Folgen $(x_n)_{n\in\mathbb{N}} \subset D(T)$ mit $x_n \longrightarrow x \in X$ und $Tx_n \longrightarrow y \in Y$ gilt $x \in D(T)$ und $Tx = y$.

4.6 Lemma. *Seien X, Y normierte Räume und $T \in L(X, Y)$. Dann ist T abgeschlossen.*

Beweis. Wir verwenden Bemerkung 4.5 b). Sei $(x_n)_{n\in\mathbb{N}} \subset D(T)$ mit $x_n \longrightarrow x$ in X und $Tx_n \longrightarrow y$ in Y. Dann gilt trivialerweise $x \in D(T) = X$, und da T stetig ist, folgt $Tx_n \longrightarrow Tx$, d.h. $Tx = y$. □

4.7 Lemma. *Seien X, Y Banachräume und $T\colon X \supset D(T) \to Y$ ein linearer Operator. Dann definiert*

$$\|x\|_T := \|x\|_X + \|Tx\|_Y \quad (x \in D(T))$$

eine Norm auf $D(T)$, die sog. Graphennorm. Der normierte Raum $(D(T), \|\cdot\|_T)$ ist genau dann ein Banachraum, wenn der Operator T abgeschlossen ist.

Beweis. (i) Sei T abgeschlossen und $(x_n)_{n\in\mathbb{N}} \subset D(T)$ eine Cauchyfolge bzgl. der Norm $\|\cdot\|_T$. Dann ist nach Definition der Norm $\|\cdot\|_T$ sowohl die Folge $(x_n)_{n\in\mathbb{N}} \subset$

X bzgl. der Norm $\|\cdot\|_X$ als auch die Folge $(Tx_n)_{n\in\mathbb{N}} \subset Y$ bzgl. der Norm $\|\cdot\|_Y$ eine Cauchyfolge. Da X und Y Banachräume sind, existieren $x \in X$ und $y \in Y$ mit $\|x_n - x\|_X \longrightarrow 0$ und $\|Tx_n - y\|_Y \longrightarrow 0$. Da T abgeschlossen ist, folgt nach Bemerkung 4.5 b) $x \in D(T)$ und $Tx = y$. Insbesondere folgt $\|x_n - x\|_T \longrightarrow 0$. Also ist $(D(T), \|\cdot\|_T)$ ein Banachraum.

(ii) Die andere Richtung der Äquivalenz zeigt man analog. □

Im folgenden schreiben wir für einen Operator $T\colon X \supset D(T) \to X$ statt $T - \lambda\,\mathrm{id}_X$ einfach $T - \lambda$.

4.8 Definition. Sei X ein Banachraum und $T\colon X \supset D(T) \to X$ ein linearer Operator.

a) $\rho(T) := \{\lambda \in \mathbb{C} : \overline{R(T-\lambda)} = X, (T-\lambda)$ injektiv, $(T-\lambda)^{-1}\colon R(T-\lambda) \to X$ stetig$\}$ heißt die Resolventenmenge von T.

b) $\sigma(T) := \mathbb{C} \setminus \rho(T)$ heißt das Spektrum von T.

c) $\sigma_p(T) := \{\lambda \in \mathbb{C}\colon T - \lambda$ nicht injektiv$\}$ heißt das Punktspektrum von T (die Menge aller Eigenwerte von T).

d) $\sigma_c(T) := \{\lambda \in \mathbb{C}\colon T-\lambda$ injektiv, $\overline{R(T-\lambda)} = X$, $(T-\lambda)^{-1}\colon R(T-\lambda) \to X$ nicht stetig$\}$ heißt das kontinuierliche Spektrum von T.

e) $\sigma_r(T) := \{\lambda \in \mathbb{C}\colon T-\lambda$ injektiv, $\overline{R(T-\lambda)} \neq X\}$ heißt das residuelle Spektrum (oder Restspektrum) von T.

4.9 Bemerkung. a) Nach Definition gilt

$$\mathbb{C} = \rho(T) \,\dot{\cup}\, \sigma(T) = \rho(T) \,\dot{\cup}\, \sigma_p(T) \,\dot{\cup}\, \sigma_c(T) \,\dot{\cup}\, \sigma_r(T),$$

wobei $\dot{\cup}$ die disjunkte Vereinigung bezeichnet.

b) Nach dem Satz vom stetigen Inversen (wird später bewiesen) folgt für abgeschlossene Operatoren T aus der Bijektivität von $(T-\lambda)\colon D(T) \to X$ bereits die Stetigkeit von $(T-\lambda)^{-1}\colon X \to X$. Somit gilt für T abgeschlossen

$$\begin{aligned} \rho(T) &= \{\lambda \in \mathbb{C} \,|\, T-\lambda\colon D(T) \to X \text{ bijektiv}\}, \\ \sigma_c(T) &= \{\lambda \in \mathbb{C} \,|\, T-\lambda\colon D(T) \to X \text{ injektiv}, \overline{R(T-\lambda)} = X,\ R(T-\lambda) \neq X\}. \end{aligned}$$

Andererseits gilt: Falls $T - \lambda\colon D(T) \to X$ bijektiv ist und $(T-\lambda)^{-1}\colon X \to X$ stetig ist, so folgt bereits die Abgeschlossenheit von T. Denn nach Lemma 4.6 ist $(T-\lambda)^{-1}$ abgeschlossen, und damit ist auch $T-\lambda$ abgeschlossen (wie man z.B. mit Bemerkung 4.5 b) sieht). Somit ist auch T abgeschlossen (wieder mit Bemerkung 4.5 b)).

c) Falls $\dim X < \infty$, so ist $\sigma_c(T) = \sigma_r(T) = \emptyset$.

4.10 Definition. Sei $T\colon X \supset D(T) \to X$ ein abgeschlossener linearer Operator.

a) Für $\lambda \in \rho(T)$ heißt $R_\lambda(T) := (T-\lambda)^{-1}$ die Resolvente von T. Die Abbildung $\rho(T) \to L(X), \lambda \mapsto R_\lambda(T)$, heißt Resolventenabbildung.

b) Für $\lambda \in \sigma_p(T)$ heißt $\ker(T-\lambda)$ der geometrische Eigenraum von T zu λ und

$$\{x \in D(T) : \exists\, n \in \mathbb{N}\colon x \in D(T^n) \text{ und } (T-\lambda)^n x = 0\}$$

der algebraische Eigenraum von T zu λ.

Hierbei haben wir den Definitionsbereich für die Verknüpfung von Operatoren (insbesondere den Definitionsbereich von $(T-\lambda)^n$) auf naheliegende Weise definiert:

4.11 Definition. Seien X, Y, Z normierte Räume. Seien ferner $S, \widetilde{S}$ lineare Operatoren von X nach Y und T ein linearer Operator von Y nach Z.

a) Der Operator $S + \widetilde{S}$ ist definiert durch

$$D(S+\widetilde{S}) := D(S) \cap D(\widetilde{S}) \text{ und } (S+\widetilde{S})x := Sx + \widetilde{S}x \quad (x \in D(S+\widetilde{S})).$$

b) Der Operator $TS\colon X \to Z$ ist definiert durch

$$D(TS) := \{x \in D(S)\colon Sx \in D(T)\} \text{ und } (TS)x := T(Sx) \quad (x \in D(TS)).$$

c) Wir schreiben $S \subset \widetilde{S}$, falls $D(S) \subset D(\widetilde{S})$ und $\widetilde{S}|_{D(S)} = S$ gilt.

4.12 Beispiel. Sei $X = \ell^2$ und $S = S_R \in L(X)$ der Rechts-Shift aus Beispiel 4.2, d.h. $S(x_1, x_2, \dots) := (0, x_1, x_2, \dots)$. Dann ist $D(S) = \ell^2$, $\ker S = \{0\}$ und $\|e_1 - y\| \geq 1$ für alle $y \in R(S)$, wobei $e_1 := (1, 0, 0, \dots)$. Daher ist $\overline{R(S)} \neq X$ und damit $0 \in \sigma_r(S)$.

b) Eigenschaften der Resolventenabbildung

4.13 Lemma (Neumannsche Reihe)**.** *Sei X ein Banachraum und $T \in L(X)$ mit $\|T\| < 1$. Dann existiert $(1-T)^{-1} \in L(X)$, und es gilt $(1-T)^{-1} = \sum_{n=0}^\infty T^n$ und $\|(1-T)^{-1}\| \leq \frac{1}{1-\|T\|}$.*

Beweis. Es gilt

$$\sum_{n=0}^N \|T^n\| \leq \sum_{n=0}^N \|T\|^n \leq \sum_{n=0}^\infty \|T\|^n = \frac{1}{1-\|T\|},$$

d.h. die Reihe konvergiert absolut. Damit existiert $S := \sum_{n=0}^\infty T^n \in L(X)$, und es gilt $\|S\| \le \frac{1}{1-\|T\|}$.

Es gilt $ST = TS = \sum_{n=0}^\infty T^{n+1} = S - 1$, d.h. $S(1-T) = (1-T)S = 1$ und damit $S = (1-T)^{-1}$. $\square$

4.14 Satz. *Sei X ein $\mathbb{C}$-Banachraum und $T\colon X \supset D(T) \to X$ ein abgeschlossener linearer Operator. Dann ist $\rho(T)$ offen und somit $\sigma(T)$ abgeschlossen.*

Beweis. Falls $\rho(T) = \emptyset$, so ist nichts zu zeigen. Sei also $\lambda_0 \in \rho(T)$. Es gilt

$$T - \lambda = T - \lambda_0 - (\lambda - \lambda_0) = (T - \lambda_0)\big[1 - (\lambda - \lambda_0)(T - \lambda_0)^{-1}\big].$$

Nach Definition der Resolventenmenge ist $(T - \lambda_0)^{-1}$ beschränkt. Für $\lambda \in \mathbb{C}$ mit $|\lambda - \lambda_0| \cdot \|(T - \lambda_0)^{-1}\| < 1$ existiert

$$\big[1 - (\lambda - \lambda_0)(T - \lambda_0)^{-1}\big]^{-1} \in L(X)$$

nach Lemma 4.13. Damit existiert

$$(T - \lambda)^{-1} = \big[1 - (\lambda - \lambda_0)(T - \lambda_0)^{-1}\big]^{-1}(T - \lambda_0)^{-1} \in L(X).$$

Somit gilt

$$\Big\{\lambda \in \mathbb{C}\colon |\lambda - \lambda_0| < \big\|(T - \lambda_0)^{-1}\big\|^{-1}\Big\} \subset \rho(T),$$

also ist $\rho(T)$ offen. $\square$

4.15 Korollar. a) *Für $\lambda_0 \in \rho(T)$ gilt*

$$\|R_{\lambda_0}(T)\| \ge \big[\operatorname{dist}(\lambda_0, \sigma(T))\big]^{-1}.$$

b) *Für $\lambda_0 \in \rho(T)$ und $\lambda \in \mathbb{C}$ mit $|\lambda - \lambda_0| < \|R_{\lambda_0}(T)\|^{-1}$ gilt*

$$R_\lambda(T) = \sum_{n=0}^\infty (\lambda - \lambda_0)^n R_{\lambda_0}(T)^{n+1}.$$

Beweis. a) folgt aus der letzten Zeile im Beweis von Satz 4.14, b) aus der Darstellung von $R_\lambda(T)$ im Beweis von Satz 4.14 und der Neumann–Reihe. $\square$

4.16 Satz. *Sei X ein $\mathbb{C}$-Banachraum und $T \in L(X)$. Dann ist das Spektrum $\sigma(T) \subset \mathbb{C}$ kompakt und nichtleer.*

Beweis. Hier wird nur die Kompaktheit von $\sigma(T)$ bewiesen. Für $\lambda \in \mathbb{C}$ mit $|\lambda| > \|T\|$ ist $T - \lambda = (-\lambda)(1 - \lambda^{-1}T)$ nach Lemma 4.13 in $L(X)$ invertierbar, d.h. $\lambda \in \rho(T)$. Also ist $\sigma(T)$ eine abgeschlossene Teilmenge von $\{\lambda \in \mathbb{C} \colon |\lambda| \leq \|T\|\}$ und damit kompakt. □

4.17 Bemerkung. Der Beweis von $\rho(T) \neq \emptyset$ für $T \in L(X)$ verwendet die Holomorphie der Resolventenabbildung und den Satz von Liouville aus der Funktionentheorie. Dabei heißt eine Banachraum-wertige Funktion $f \colon U \to X$, $U \subset \mathbb{C}$ offen, holomorph, falls für alle $z_0 \in U$ der Grenzwert $\lim_{z \to z_0} \frac{f(z)-f(z_0)}{z-z_0}$ existiert, wobei die Konvergenz in der Norm von X zu verstehen ist. Mit Hilfe der Reihendarstellung von $R_\lambda(T)$ aus Korollar 4.15 b) kann man zeigen, dass die Resolventenabbildung eine holomorphe Abbildung von $\rho(T)$ in den Banachraum $L(X)$ ist. Wäre $\rho(T) = \mathbb{C}$, so würde aus $\|R_\lambda(T)\| \to 0$ für $|\lambda| \to \infty$ und dem Satz von Liouville folgen, dass diese Abbildung unabhängig von λ ist.

4.18 Beispiel. Die folgenden Beispiele zeigen, dass für unbeschränkte Operatoren sehr wohl die Fälle $\sigma(T) = \mathbb{C}$ und $\sigma(T) = \emptyset$ auftreten können.

a) Sei $X = C([0,1])$ und $Tf := f'$ für $f \in D(T) := C^1([0,1])$. Dann ist T unbeschränkt, da $\|Tf_n\| = n$ und $\|f_n\| = 1$ gilt für $f_n(t) := t^n$.

Wir zeigen, dass T abgeschlossen ist: Sei $(f_n)_{n\in\mathbb{N}} \subset D(T)$ mit Konvergenzen $f_n \longrightarrow f$ in X und $Tf_n = f_n' \longrightarrow g$ in X. Da $(f_n)_{n\in\mathbb{N}}$ und $(f_n')_{n\in\mathbb{N}}$ gleichmäßig konvergieren, gilt $f \in C^1([0,1])$ und $f_n' \longrightarrow f'$. Somit ist $f \in D(T)$ und $g = f' = Tf$.

Weil für jedes $\lambda \in \mathbb{C}$ die Funktion $f(t) := e^{\lambda t}$ in $\ker(T - \lambda)$ liegt, gilt gilt $\sigma(T) = \sigma_p(T) = \mathbb{C}$.

b) Sei $X := C_0([0,1]) := \{f \in C([0,1]) \colon f(0) = 0\}$ und $Tf := f'$ für $f \in D(T) := \{f \in X \colon f' \in X\}$. Sei $\lambda \in \mathbb{C}$, und seien $f, g \in X$. Betrachte die Gleichung $(T - \lambda)f = g$, d.h. $f' - \lambda f = g$. Versehen mit der Anfangsbedingung $f(0) = 0$ hat diese gewöhnliche Differentialgleichung die eindeutige Lösung

$$f(t) = e^{\lambda t} \int_0^t e^{-\lambda s} g(s)\,\mathrm{d}s.$$

Es gilt $f'(0) = g(0) + \lambda f(0) = 0$, d.h. $f' \in X$ und damit $f \in D(T)$. Somit ist $T - \lambda \colon D(T) \to X$ bijektiv für alle $\lambda \in \mathbb{C}$, d.h. $\rho(T) = \mathbb{C}$.

Der Vergleich von a) und b) zeigt, dass eine kleine Änderung des Grundraums X das Spektrum eines unbeschränkten Operators erheblich beeinflussen kann.

4.19 Definition und Satz (Adjungierte beschränkter Operatoren)**.** *Seien X, Y Banachräume und $T \in L(X,Y)$. Für jedes $f \in Y'$ wird durch $f_1 := f \circ T$, d.h. $f_1(x) := f(Tx)$ $(x \in X)$ ein beschränktes lineares Funktional $f_1 \in X'$ definiert. Die*

Abbildung $T'\colon Y' \to X'$, $f \mapsto f \circ T$ heißt (Banachraum-)adjungierter Operator zu T. Es gilt $T' \in L(Y', X')$. Die Abbildung $T \mapsto T'$, $L(X,Y) \to L(Y',X')$ ist eine Isometrie.

Manchmal schreibt man $\langle x, g\rangle_{X\times X'} := g(x)$ für $x \in X$ und $g \in X'$. In dieser Schreibweise wird die Gleichung $f(Tx) = (T'f)(x)$ zu

$$\langle Tx, f\rangle_{Y\times Y'} = \langle x, T'f\rangle_{X\times X'}.$$

Beweis. Wegen $f_1 = f \circ T$ ist die Linearität und die Stetigkeit von f_1 klar. Wegen

$$|f_1(x)| = |f(Tx)| \le \|f\|_{Y'} \cdot \|T\| \cdot \|x\|\,, \qquad x \in X,$$

ist $\|f_1\|_{X'} \le \|T\| \cdot \|f\|_{Y'}$, d.h. $\|T'\| \le \|T\|$. Der Operator T' ist linear wegen

$$T'(\alpha f + \beta g)(x) = (\alpha f + \beta g)(Tx) = \alpha f(Tx) + \beta g(Tx).$$

Ebenso ist die Abbildung $T \mapsto T'$ linear.

Zu zeigen ist noch, dass $\|T\| \le \|T'\|$ gilt. Nach Lemma 3.5 a) existiert zu $x \in X$ ein $f_x \in Y'$ mit $\|f_x\|_{Y'} = 1$ und $f_x(Tx) = \|Tx\|_Y$. Damit ist

$$\|Tx\|_Y = |f_x(Tx)| = |(T'f_x)(x)| \le \|T'\| \cdot \|x\|_X\,. \qquad \square$$

4.20 Bemerkung. Es gelten die üblichen Rechenregeln für den adjungierten Operator: Falls $T \in L(X,Y)$ und $S \in L(Y,Z)$, so ist $(ST)' = T'S'$. Denn es gilt $((ST)'f)(x) = f(STx) = (S'f)(Tx) = [T'(S'f)](x)$ $(x \in X)$, also $(ST)'f = T'S'f$ für jedes $f \in Z'$.

Falls $T \in L(X,Y)$ invertierbar ist, so gilt $(T^{-1})' = (T')^{-1}$. Dies folgt aus $(T^{-1})'T' = (TT^{-1})' = \mathrm{id}_Y' = \mathrm{id}_{Y'}$ und $T'(T^{-1})' = (T^{-1}T)' = \mathrm{id}_X' = \mathrm{id}_{X'}$.

4.21 Definition (Adjungierte unbeschränkter Operatoren)**.** Seien X, Y Banachräume und $T\colon X \to Y$ ein linearer dicht definierter Operator (d.h. $\overline{D(T)} = X$). Definiere

$$D(T') := \{f \in Y'\colon\ \exists\, f_1 \in X' \text{ mit } f(Tx) = f_1(x) \quad (x \in D(T))\}$$

und

$$T'f := f_1 \quad (f \in D(T')).$$

Kurz kann man auch schreiben: $D(T') = \{f \in Y'\colon\ x \mapsto f(Tx) \in X'\}$. Die Abbildung $T'\colon Y' \to X'$ mit Definitionsbereich $D(T')$ heißt (Banachraum-)Adjungierte von T.

4.22 Bemerkung. a) $T'f$ ist eindeutig definiert. Denn seien $f_1, f_2 \in X'$ mit $f_1(x) = f(Tx) = f_2(x) \quad (x \in D(T))$. Da $\overline{D(T)} = X$ und f_1, f_2 stetig sind, folgt $f_1 = f_2$ auf ganz X. Aus der Eindeutigkeit von $T'f$ folgt dann auch die Linearität von T'.

b) Es gilt $(f, g) \in G(T')$ genau dann, wenn $g(x) = f(Tx)$ $(x \in D(T))$.

c) Der Operator T' ist abgeschlossen. Denn sei $(f_n)_{n\in\mathbb{N}} \subset D(T')$ mit Konvergenzen $f_n \longrightarrow f$ in Y' und $T'f_n \longrightarrow g$ in X'. Dann gilt für $x \in D(T)$:

$$f(Tx) = \lim_{n\to\infty} f_n(Tx) = \lim_{n\to\infty} (T'f_n)(x) = g(x).$$

Damit haben wir $f \in D(T')$ und $(f, g) \in G(T')$ nach b).

Hilberträume sind Spezialfälle von Banachräumen, daher übertragen sich die obigen Definitionen auch auf Hilberträume. Man verwendet aber meistens den Begriff der Hilbertraum-Adjungierten. Grundlage dafür ist die Beschreibung des Dualraums durch den Satz von Riesz (Satz 2.32).

4.23 Definition und Satz. *Seien X, Y Hilberträume, und $T \in L(X, Y)$. Dann existiert zu jedem $y \in Y$ genau ein $x^* \in X$ mit*

$$\langle Tx, y\rangle_Y = \langle x, x^*\rangle_X \quad (x \in X).$$

Setze $T^*y := x^*$. Die Abbildung $T^* \in L(Y, X)$ heißt (Hilbertraum-)Adjungierte zu T.

Beweis. Das folgt sofort aus Satz 4.19 und dem Satz von Riesz. Man beachte, dass die Abbildung aus dem Satz von Riesz

$$i_X : X \to X', \; x \mapsto \langle \cdot, x\rangle$$

isometrisch aber konjugiert linear ist. Damit hängen Hilbertraum- und Banachraum-adjungierte über $T^* = i_X^{-1} \circ T' \circ i_Y$ zusammen. □

4.24 Definition. Seien X, Y Hilberträume und $T : X \to Y$ ein linearer dicht definierter Operator. Definiere

$$D(T^*) := \{y \in Y : \; x \mapsto \langle Tx, y\rangle \text{ ist stetiges lineares Funktional auf } D(T)\}.$$

Für $y \in D(T^*)$ existiert ein eindeutiges $x^* \in X$ mit

$$\langle Tx, y\rangle_Y = \langle x, x^*\rangle_X \quad (x \in D(T)).$$

Definiere $T^* : Y \to X$ durch $T^*y := x^* \quad (y \in D(T^*))$. Der Operator T^* heißt (Hilbertraum-)Adjungierte von T.

4.25 Definition. a) Seien X, Y Hilberträume und $T \in L(X, Y)$. Dann heißt T unitär, falls $TT^* = \mathrm{id}_Y$ und $T^*T = \mathrm{id}_X$ gilt.

b) Sei X ein Hilbertraum und T ein linearer dicht definierter Operator in X. Dann heißt T

(i) selbstadjungiert, falls $T = T^*$,

(ii) normal, falls $TT^* = T^*T$,

(iii) wesentlich selbstadjungiert, falls T abschließbar ist und $\overline{T}$ selbstadjungiert ist,

(iv) symmetrisch, falls $T \subset T^*$ gilt.

Bei dieser Definition ist zu beachten, dass zwei Operatoren genau dann gleich sind, falls sie gleiche Definitionsbereiche haben und darauf gleiche Werte annehmen.

5. Distributionen und Sobolevräume

5.1 Worum geht's? In der Theorie der partiellen Differentialgleichungen ist es oftmals sinnvoll, den Begriff der bekannten klassischen Differenzierbarkeit aufzulockern. Dies ermöglicht es etwa, Lösungsbegriffe zu definieren, die zwar schwächer sind als bereits bekannte, trotzdem aber für viele Anwendungen ausreichend sind. Die Theorie der Distributionen stellt nun Begriffsbildungen zur Verfügung, die es uns im Folgenden erlauben werden gewisse Sachverhalte elegant zu beschreiben. Insbesondere werden wir in der Lage sein, der Diracschen δ-„Funktion" einen präzisen Sinn zu geben. Dies wiederum wird von großer Wichtigkeit für die darauf folgende Potentialtheorie sein. Zu erwähnen ist, dass der Name „Distributionen" von Laurent Schwartz eingeführt wurde.

a) Distributionen

In diesem Kapitel werden wir die praktische Multiindex-Schreibweise verwenden:

5.2 Definition (Multiindex-Schreibweise)**.** Seien $x, \xi \in \mathbb{R}^n$ und $\alpha \in \mathbb{N}_0^n$. Dann definieren wir

$$\begin{aligned} x\xi &:= \sum_{j=1}^n x_j \xi_j, \\ |x| &:= \Big(\sum_{j=1}^n x_j^2 \Big)^{1/2}, \\ x^\alpha &:= x_1^{\alpha_1} \cdot \ldots \cdot x_n^{\alpha_n}, \\ |\alpha| &:= \alpha_1 + \cdots + \alpha_n. \end{aligned}$$

Für $f \in C^{|\alpha|}(\mathbb{R}^n)$ sei

$$\partial^\alpha f(x) := \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} f(x).$$

5.3 Definition. Sei $G \subset \mathbb{R}^n$ offen.

a) Für eine Funktion $\varphi\colon G \to \mathbb{C}$ heißt $\operatorname{supp}\varphi := \overline{\{x \in G\colon \varphi(x) \neq 0\}}$ der Träger von φ.

Die Menge $C_0^\infty(G) := \{\varphi \in C^\infty(G)\colon \operatorname{supp}\varphi \subset G, \operatorname{supp}\varphi \text{ kompakt}\}$ heißt die Menge der Testfunktionen auf G (oder die Menge der C^∞-Funktionen mit kompaktem Träger). Manchmal sieht man auch die Bezeichnung $C_c^\infty(G) := C_0^\infty(G)$.

c) Wir schreiben $K \Subset G$, falls K eine kompakte Teilmenge von G ist. Sei nun

$(\varphi_\ell)_{\ell\in\mathbb{N}} \subset C_0^\infty(G)$, dann definieren wir die Konvergenz wie folgt:

$$\varphi_\ell \to 0\ (\ell\to\infty) \text{ in } \mathscr{D}(G) :\Longleftrightarrow \begin{cases} (i)\ \ \exists K \Subset G\colon \forall \ell\in\mathbb{N}\colon \ \operatorname{supp}\varphi_\ell \subset K \\ (ii)\ \forall\alpha\in\mathbb{N}_0^n\colon \ \sup\limits_{x\in K} |\partial^\alpha\varphi_\ell(x)| \to 0\ (\ell\to\infty). \end{cases}$$

c) Die Menge $\mathscr{D}'(G)$ wird definiert als die Menge aller Abbildungen $u\colon \mathscr{D}(G)\to\mathbb{C}$ mit den Eigenschaften

(i) $u\colon \mathscr{D}(G)\to\mathbb{C}$ ist linear,

(ii) für alle Folgen $(\varphi_\ell)_{\ell\in\mathbb{N}}$ mit $\varphi_\ell\to 0$ $(\ell\to\infty)$ in $\mathscr{D}(G)$ gilt $u(\varphi_\ell)\to 0$.

$\mathscr{D}'(G)$ heißt Menge der Distributionen von G.

5.4 Bemerkung. a) Es existiert eine Topologie $\mathcal{T}$ auf $\mathscr{D}(G)$ (in Form einer lokalkonvexen Topologie), so dass die oben definierte Konvergenz in $\mathscr{D}(G)$ gerade der Konvergenz bzgl. $\mathcal{T}$ entspricht und eine lineare Abbildung $u\colon \mathscr{D}(G)\to\mathbb{C}$ genau dann stetig ist, wenn c) (ii) gilt. Versieht man $\mathscr{D}(G)$ mit dieser Topologie, so ist $\mathscr{D}'(G) = \{u\colon \mathscr{D}(G)\to\mathbb{C} \mid u \text{ linear und stetig}\}$.

b) Es existieren ausreichend viele Testfunktionen; ein Beispiel einer Testfunktion lässt sich mit Hilfe der exp-Funktion konstruieren. Tatsächlich liegen die Testfunktionen sogar dicht in $L^p(G)$ für $1\le p<\infty$ (aber nicht in $L^\infty(G)$). Der Beweis dieser Aussage verwendet etwa den sogenannten Friedrichsschen Glättungsoperator.

5.5 Beispiel (reguläre Distributionen)**.** Sei $f\in L^1_{\mathrm{loc}}(G) := \{f\colon G\to\mathbb{C}\colon \forall K\subset G,$ K kompakt: $f|_K\in L^1(K)\}$. Dann definiert

$$[f]\colon \mathscr{D}(G)\to\mathbb{R}, \qquad \varphi\mapsto [f](\varphi) := \int_G f(x)\varphi(x)\,\mathrm{d}x$$

eine Distribution, denn

(i) $[f]$ ist offensichtlich linear.

(ii) Sei $(\varphi_\ell)_{\ell\in\mathbb{N}}$ eine Folge in $\mathscr{D}(G)$ mit $\varphi_\ell\to 0$ in $\mathscr{D}(G)$ für $\ell\longrightarrow\infty$, dann folgt:

$$|[f]\varphi_\ell| \le \int_G |f\varphi_\ell|\,\mathrm{d}x \le \|f\|_{L^1(K)}\cdot\sup_{x\in K}|\varphi_\ell(x)| \longrightarrow 0$$

für $\ell\longrightarrow\infty$ und ein passendes $K\Subset G$. In diesem Zusammenhang spricht man auch davon, dass $[f]$ eine von f erzeugte Distribution ist. Eine von einer L^1_{loc}-Funktion erzeugte Distribution heißt *reguläre* Distribution.

5.6 Beispiel (Dirac-Distribution)**.** Ein Beispiel für eine nicht-reguläre Distribution ist die sog. Dirac-Distribution (Diracsche „Deltafunktion“), die für festes $x_0 \in G$ definiert ist durch

$$\delta_{x_0} \colon \mathscr{D}(G) \to \mathbb{C}, \quad \varphi \mapsto \delta_{x_0}(\varphi) := \varphi(x_0).$$

Offensichtlich ist $\delta_{x_0} \colon \mathscr{D}(G) \to \mathbb{C}$ linear. Sei $(\varphi_\ell)_{\ell \in \mathbb{N}} \subset \mathscr{D}(G)$ eine Folge mit $\varphi_\ell \to 0$ in $\mathscr{D}(G)$. Dann gilt

$$|\delta_{x_0}(\varphi_\ell)| = |\varphi_\ell(x_0)| \leq \sup_{x \in K} |\varphi_\ell(x)| \longrightarrow 0$$

für $\ell \to \infty$ und jedes $K \Subset G$ mit $x_0 \in K$.

Die Dirac-Distribution ist jedoch nicht regulär. Denn sonst existiert ein $f \in L^1_{\text{loc}}(G)$ existiert mit

$$\delta_{x_0}(\varphi) = \varphi(x_0) = \int_G f(x)\varphi(x)\,\mathrm{d}x \quad (\varphi \in \mathscr{D}(G)).$$

Wir wählen $\varepsilon > 0$ so, dass $\overline{B(x_0, \varepsilon)} \subset G$ und $\int_{B(x_0,\varepsilon)} |f(x)|\,\mathrm{d}x < 1$ gilt. Weiter wählen wir eine Testfunktion φ, für die einerseits $\operatorname{supp} \varphi \subset B(x_0, \varepsilon)$ und andererseits $\forall x \in G \colon \varphi(x_0) \geq \varphi(x) \geq 0$ sowie $\varphi(x_0) > 0$ gilt. Damit ergibt sich dann aber

$$\delta_{x_0}(\varphi) = \varphi(x_0) = |\varphi(x_0)| = \left| \int_G f(x)\varphi(x)\,\mathrm{d}x \right| \leq \varphi(x_0) \int_{B(x_0,\varepsilon)} |f(x)|\,\mathrm{d}x < \varphi(x_0),$$

Widerspruch.

Wie zu Anfang dieses Kapitels bereits erwähnt wurde, haben wir das Ziel, den klassischen Ableitungsbegriff aufzulockern oder zu verallgemeinern. Das bedeutet aber, dass sich dieser Ableitungsbegriff bei klassisch differenzierbaren Funktionen nicht von dem klassischen Begriff unterscheiden sollte. Es sei $f \in C^k(\mathbb{R}^n)$ und $[f]$ die von f erzeugte reguläre Distribution. Offenbar gilt dann für alle $\varphi \in \mathscr{D}(\mathbb{R}^n)$ und alle $\alpha \in \mathbb{N}_0^n$ mit $|\alpha| \leq k$

$$[\partial^\alpha f](\varphi) = \int_{\mathbb{R}^n} (\partial^\alpha f)\varphi\,\mathrm{d}x = (-1)^{|\alpha|} \int_{\mathbb{R}^n} f \partial^\alpha \varphi\,\mathrm{d}x = (-1)^{|\alpha|}[f](\partial^\alpha \varphi).$$

Dies motiviert die folgende

5.7 Definition. Für beliebiges $u \in \mathscr{D}'(G)$ sei für $\alpha \in \mathbb{N}_0^n$

$$\partial^\alpha u \colon \mathscr{D}(G) \to \mathbb{C}, \qquad \varphi \mapsto (-1)^{|\alpha|} u(\partial^\alpha \varphi).$$

$\partial^\alpha u$ heißt Ableitung der Distribution u vom Grad $|\alpha|$.

Es ist klar, dass $\partial^\alpha u \in \mathscr{D}'(G)$ ist, da mit $\varphi_\ell \to 0$ in $\mathscr{D}(G)$ auch $\partial^\alpha \varphi_\ell \to 0$ in $\mathscr{D}(G)$ gilt. Also ist jede Distribution beliebig oft differenzierbar.

5.8 Beispiel (Heaviside-Funktion)**.** Sei

$$h(x) := \begin{cases} 1 & : x \geq 0, \\ 0 & : x < 0. \end{cases}$$

Dann ist $h \in L^1_{\text{loc}}(\mathbb{R})$, und für $\varphi \in \mathscr{D}(\mathbb{R})$, $\operatorname{supp} \varphi \subset [a,b]$ mit $a \leq 0 \leq b$ gilt

$$[h]'(\varphi) = -\int_a^b h(x)\varphi'(x)\,\mathrm{d}x = -\int_0^b \varphi'(x)\,\mathrm{d}x = \varphi(0) = \delta_0(\varphi).$$

Also erhalten wir $[h]' = \delta_0$ in $\mathscr{D}'(G)$.

b) Sobolevräume: Definition und erste Eigenschaften

Im folgenden sei $G \subset \mathbb{R}^n$ ein Gebiet und $1 \leq p < \infty$.

Für eine Distribution $u \in \mathscr{D}'(G)$ und einen Multiindex $\alpha \in \mathbb{N}_0^n$ schreiben wir

$$\partial^\alpha u \in L^p(G),$$

falls eine Funktion $f \in L^p(G)$ existiert mit $\partial^\alpha u = [f]$ in $\mathscr{D}'(G)$. Hier ist $[f]$ wieder die zu f gehörige reguläre Distribution.

5.9 Definition (Sobolevräume)**.** a) Zu $s \in \mathbb{N}_0$ definiere

$$W^{s,p}(G) := \{u \in \mathscr{D}'(G) \colon\ \partial^\alpha u \in L^p(G) \quad (0 \leq |\alpha| \leq s)\}.$$

Als Norm in $W^{s,p}(G)$ definiert man

$$\|u\|_{W^{s,p}(G)} := \|u\|_{s,p,G} := \Bigg(\sum_{0 \leq |\alpha| \leq s} \|\partial^\alpha u\|^p_{L^p(G)} \Bigg)^{1/p}.$$

b) Zu $s \in \mathbb{N}_0$ definiere $H^{s,p}(G)$ als die Vervollständigung von $\{u \in C^s(G) \colon\ \|u\|_{s,p,G} < \infty\}$. Im Falle $p = 2$ schreiben wir auch $H^s(G)$ statt $H^{s,2}(G)$.

c) Zu $s \in \mathbb{N}_0$ definiere $H_0^{s,p}(G)$ als den Abschluss von $C_0^\infty(G)$ im Raum $H^{s,p}(G)$.

5.10 Bemerkung. a) In der Definition von $W^{s,p}(G)$ wird insbesondere $u \in L^p(G)$ gefordert. Daher kann man auch schreiben

$$W^{s,p}(G) = \{u \in L^p(G) \colon\ \partial^\alpha u \in L^p(G) \quad (0 \leq |\alpha| \leq s)\}.$$

b) Die Vervollständigung eines metrischen Raums kann abstrakt definiert werden. Im Fall von $H^{s,p}(G)$ ist aber wegen $\|\cdot\|_{L^p(G)} \leq \|\cdot\|_{s,p,G}$ offensichtlich $H^{s,p}(G) \subset L^p(G)$, d.h. ein Element der abstrakten Vervollständigung kann mit einer Funktion in $L^p(G)$ identifiziert werden.

c) Im Fall $p = 2$ erhält man das Skalarprodukt

$$\langle u, v \rangle_{H^s(G)} := \sum_{|\alpha| \le s} \langle \partial^\alpha u, \partial^\alpha v \rangle_{L^2(G)} .$$

5.11 Lemma. *Die Räume $H^{s,p}(G)$ und $W^{s,p}(G)$ sind Banachräume.*

Beweis. Der Raum $H^{s,p}(G)$ ist als Vervollständigung eines normierten Raumes konstruktionsgemäß ein Banachraum. Sei also $(u_\ell)_{\ell\in\mathbb{N}} \subset W^{s,p}(G)$ eine Cauchyfolge. Nach Definition der Norm ist für $0 \le |\alpha| \le s$ auch $(\partial^\alpha u_\ell)_{\ell\in\mathbb{N}} \subset L^p(G)$ eine Cauchyfolge, daher existiert ein $u_\alpha \in L^p(G)$ mit $\partial^\alpha u_\ell \longrightarrow u_\alpha$ in $L^p(G)$. Setze $u := u_{(0,\dots,0)}$.

Für die zugehörigen regulären Distributionen gilt mit der Hölder-Ungleichung für alle $\varphi \in \mathscr{D}(G)$

$$\begin{aligned}
|[\partial^\alpha u_\ell](\varphi) - [u_\alpha](\varphi)| &= \left| \int_G (\partial^\alpha u_\ell - u_\alpha)(x)\varphi(x)\,\mathrm{d}x \right| \\
&\le \|\partial^\alpha u_\ell - u_\alpha\|_{L^p(G)} \cdot \|\varphi\|_{L^q(G)} \longrightarrow 0 \quad (\ell \longrightarrow \infty),
\end{aligned}$$

wobei $\frac{1}{p} + \frac{1}{q} = 1$ ist. Also gilt

$$(\partial^\alpha [u])(\varphi) = (-1)^{|\alpha|}[u](\partial^\alpha \varphi) = \lim_{\ell\to\infty} (-1)^{|\alpha|}[u_\ell](\partial^\alpha\varphi) = \lim_{\ell\to\infty} [\partial^\alpha u_\ell](\varphi) = [u_\alpha](\varphi)$$

für alle $\varphi \in \mathscr{D}(G)$. Somit ist $\partial^\alpha u = u_\alpha$ in $\mathscr{D}'(G)$, d.h. $u \in W^{s,p}(G)$.

Es folgt

$$\|u_\ell - u\|_{s,p,G}^p = \sum_{0\le|\alpha|\le s} \|\partial^\alpha u_\ell - \partial^\alpha u\|_{L^p(G)}^p \longrightarrow 0 \quad (\ell \longrightarrow \infty),$$

also haben wir $u_\ell \longrightarrow u$ in $W^{s,p}(G)$, und $W^{s,p}(G)$ ist ein Banachraum. □

Die beiden Definitionen von Sobolevräumen sind für allgemeine Gebiete äquivalent. Der folgende Satz wurde erst 1964 bewiesen (während die ersten Definitionen bereits 1938 formuliert wurden).

5.12 Satz („$H = W$“)**.** *Für $1 \le p < \infty$ und $s \in \mathbb{N}_0$ gilt*

$$W^{s,p}(G) = H^{s,p}(G).$$

Beweisskizze. **(i)** Es gilt $H^{s,p}(G) \subset W^{s,p}(G)$: Sei $u \in H^{s,p}(G)$ und $(u_\ell)_{\ell\in\mathbb{N}} \subset C^s(G)$ eine Cauchyfolge bzgl. $\|\cdot\|_{s,p,G}$, welche gegen u konvergiert. Nach Definition der $\|\cdot\|_{s,p,G}$-Norm gilt wieder $\partial^\alpha u_\ell \longrightarrow u_\alpha$ mit $u_\alpha \in L^p(G)$. Wie im letzten Beweis sieht man $\partial^\alpha u = u_\alpha$ in $\mathscr{D}'(G)$ und damit $u \in W^{s,p}(G)$.

(ii) Für die Inklusion $W^{s,p}(G) \subset H^{s,p}(G)$ ist zu zeigen, dass $C^s(G) \cap H^{s,p}(G)$ dicht in $W^{s,p}(G)$ liegt. Unter Verwendung des Friedrichsschen Glättungsoperators kann man sogar zeigen, dass

$$\{\varphi \in C^\infty(G)\colon\ \|\varphi\|_{s,p,G} < \infty\}$$

dicht in $W^{s,p}(G)$ liegt. Dies geschieht über eine kompakte Ausschöpfung von G und eine zugehörige Partition der Eins. Die Details sind z.B. im Buch von Adams [1] beschrieben. □

Falls $u \in C^1(\overline{G})$ und G ein Gebiet mit C^1-Rand ist, so gilt nach dem Satz von Gauß für $f \in C^1(\overline{G}; \mathbb{C}^n)$ die Gleichheit

$$\int_G u(x)\operatorname{div} f(x)dx = -\int_G (\nabla u(x))^\top f(x)dx + \int_{\partial G} u(x)n(x)^\top f(x)dS(x). \tag{5-1}$$

Dabei ist $n(x) \in \mathbb{R}^n$ für $x \in \partial G$ der äußere Normalenvektor und dS steht für das Oberflächenmaß auf ∂G. Die Divergenz eines Vektorfeldes $f \in C^1(\overline{G}; \mathbb{C}^n)$ war definiert als $\operatorname{div} f := \sum_{j=1}^n \partial_{x_j} f_j$. Es gilt $u|_{\partial G} = 0$ genau dann, wenn für alle $f \in C^1(\overline{G}; \mathbb{C}^n)$ das letzte Integral in (5-1) verschwindet, d.h. falls gilt

$$\langle u, \operatorname{div} f\rangle_{L^2(G)} = -\langle \nabla u, f\rangle_{L^2(G;\mathbb{C}^n)} \quad (f \in C^1(\overline{G}; \mathbb{C}^n)). \tag{5-2}$$

Dabei definiert man

$$\langle f, g\rangle_{L^2(G;\mathbb{C}^n)} := \int_G \langle f(x), g(x)\rangle_{\mathbb{C}^n} dx \quad (f, g \in L^2(G; \mathbb{C}^n))$$

mit dem Standard-Skalarprodukt $\langle\cdot,\cdot\rangle_{\mathbb{C}^n}$ auf $\mathbb{C}^n$.

Der Raum $H_0^1(G)$ war definiert als Abschluss von $\mathscr{D}(G)$ im Raum $H^1(G)$. Da alle Testfunktionen auf dem Rand ∂G den Wert 0 annehmen, ist für $u \in H^1(G)$ die Bedingung $u \in H_0^1(G)$ als eine Verallgemeinerung der Randbedingung $u|_{\partial G} = 0$ (Dirichlet-Randbedingung) zu verstehen. Der folgende Satz zeigt, dass auch im Sobolevraum eine zu (5-2) analoge Beschreibung gilt.

5.13 Satz. *Sei $u \in H^1(G)$. Dann gilt $u \in H_0^1(G)$ genau dann, wenn für alle $f \in L^2(G; \mathbb{C}^n)$ mit $\operatorname{div} f \in L^2(G)$ die Gleichheit $\langle u, \operatorname{div} f\rangle_{L^2(G)} = -\langle \nabla u, f\rangle_{L^2(G;\mathbb{C}^n)}$ gilt.*

Der Satz wird hier nicht bewiesen. Man beachte, dass alle Ableitungen hier distributionell zu verstehen sind.

c) Wichtige Sätze aus der Theorie der Sobolevräume

Die folgenden Ergebnisse aus der Theorie der Sobolevräume werden nicht bewiesen.

5.14 Satz (Sobolevscher Einbettungssatz). *Seien $s \in \mathbb{N}$, $k \in \mathbb{N}_0$ und $1 \le p < \infty$, und sei $G \subset \mathbb{R}^n$ ein C^1-Gebiet.*

Falls $s > \frac{n}{p} + k$, dann gilt

$$W^{s,p}(G) \hookrightarrow C_b^k(G).$$

Insbesondere existiert ein $C > 0$, so dass für alle $u \in W^{s,p}(G)$ die Abschätzung

$$\|u\|_{C_b^k G)} \le C \|u\|_{W^{s,p}(G)}$$

gilt.

5.15 Definition. Seien X und Y normierte Räume und $K\colon X \longrightarrow Y$ eine (nicht unbedingt lineare) Abbildung. Dann heißt K kompakt, falls für jede beschränkte Folge $(x_n)_{n\in\mathbb{N}} \subset X$ die Folge $(Kx_n)_{n\in\mathbb{N}} \subset Y$ eine konvergente Teilfolge besitzt.

5.16 Satz (Rellich–Kondrachov). *Sei $G \subset \mathbb{R}^n$ ein beschränktes C^1-Gebiet.*

a) Für $1 \le p < \infty$ und $s \in \mathbb{N}$ ist die Einbettung

$$W^{s,p}(G) \hookrightarrow L^p(G)$$

kompakt.

b) Für $1 \le p < \infty$ und $s \in \mathbb{N}$ mit $sp > n$ ist die Einbettung

$$W^{s,p}(G) \hookrightarrow C_b(\overline{G})$$

kompakt.

Die folgende Ungleichung ist sehr wichtig, um Abschätzungen beweisen zu können.

5.17 Satz (Erste Poincarésche Ungleichung). *Es sei $G \subset \mathbb{R}^n$ ein Gebiet, welches in eine Richtung beschränkt ist. Dann existiert eine Konstante $b > 0$ so, dass*

$$\|u\|_{L^2(G)} \le b \|\nabla u\|_{L^2(G;\mathbb{C}^n)} \quad (u \in H_0^1(G)).$$

Damit ist durch

$$|u|_{H^1(G)} := \Big(\sum_{|\alpha|=1} \|\partial^\alpha u\|_{L^2(G)}^2 \Big)^{1/2}$$

auf $H_0^1(G)$ eine Norm gegeben, welche zur Norm $\|\cdot\|_{H^1(G)}$ äquivalent ist.

Beweis. Es sei G ohne Einschränkung in x_n-Richtung beschränkt, d.h. es gelte

$$G \subset \{x \in \mathbb{R}^n : 0 \le x_n \le b\}$$

für ein $b > 0$. Wir beweisen die Aussage zunächst für $u \in \mathscr{D}(G)$, wobei wir $u(x) := 0$ für $x \in \mathbb{R}^n \setminus G$ setzen. Offenbar gilt nach dem Hauptsatz der Differential– und Integralrechnung

$$u(x) = \int_0^{x_n} 1 \cdot \partial_n u(x_1, \dots, x_{n-1}, t) dt.$$

Damit können wir unter Verwendung der Cauchy–Schwarz–Ungleichung wie folgt abschätzen:

$$|u(x)|^2 \le b \int_0^{x_n} |\partial_n u(x_1, \dots, x_{n-1}, t)|^2 dt \le b \int_0^b |\nabla u(x_1, \dots, x_{n-1}, t)|^2 dt.$$

Somit folgt

$$\begin{aligned} \|u\|^2_{L^2(G)} &\le b \int_0^b \left\{ \int_{\mathbb{R}^{n-1}} \int_0^b |\nabla u(x_1, \dots, x_{n-1}, t)|^2 dt dx_1 \cdots dx_{n-1} \right\} dx_n \\ &\le b^2 \|\nabla u\|^2_{L^2(G)}, \end{aligned}$$

und das war zu zeigen. Da $\mathscr{D}(G) \subset H^1_0(G)$ dicht liegt, folgt die Abschätzung für beliebige $u \in H^1_0(G)$ durch Approximation in der $\|\cdot\|_{H^1(G)}$-Norm. □

5.18 Satz (Zweite Poincarésche Ungleichung). *Sei $G \subset \mathbb{R}^n$ ein beschränktes C^1-Gebiet. Dann existiert ein $c > 0$ mit*

$$\|u\|_{L^2(G)} \le c \left(\|\nabla u\|_{L^2(G)^n} + \left| \int_G u(x)\,\mathrm{d}x \right| \right) \quad (u \in H^1(G)).$$

Damit ist durch

$$\|u\|_{H^1_*(G)} := \left(\|\nabla u\|^2_{L^2(G)^n} + \left| \langle u, 1 \rangle_{L^2(G)} \right|^2 \right)^{1/2}$$

eine Norm auf $H^1(G)$ gegeben, welche zur Norm $\|\cdot\|_{H^1(G)}$ äquivalent ist.

Beweis. Wir nehmen indirekt an, dass eine Folge $(u_\ell)_{\ell\in\mathbb{N}} \subset H^1(G)$ existiert mit $\|u_\ell\|_{L^2(G)} = 1$ und

$$\|\nabla u_\ell\|_{L^2} + \left| \langle u_\ell, 1 \rangle_{L^2(G)} \right| < \frac{1}{\ell} \quad (\ell \in \mathbb{N}).$$

Nach dem Satz von Rellich–Kondrachov existiert eine Teilfolge $(\widetilde{u}_\ell)_{\ell\in\mathbb{N}}$, die etwa gegen $u_0 \in L^2(G)$ konvergiert. Wegen $\|\nabla \widetilde{u}_\ell\|_{L^2} \to 0$ folgt, dass $(\widetilde{u}_\ell)_{\ell\in\mathbb{N}} \subset H^1(G)$ eine

Cauchyfolge ist. Da $H^1(G)$ vollständig ist, existiert ein $\overline{u}_0 \in H^1(G)$ mit $\widetilde{u}_\ell \to \overline{u}_0$ in $H^1(G)$. Wegen der Eindeutigkeit des Grenzwertes ist $u_0 = \overline{u}_0$. Wegen $\langle \widetilde{u}_\ell, 1\rangle \longrightarrow 0$ folgt $\langle u_0, 1\rangle = 0$. Andererseits gilt $\nabla u_0 = \lim_{n\to\infty} \nabla \widetilde{u}_\ell = 0$, also $u_0 = \text{const}$. Insgesamt folgt also $u_0 = 0$, was im Widerspruch zu

$$\|u_0\|_{L^2} = \lim_{\ell\to\infty} \|\widetilde{u}_\ell\|_{L^2} = 1$$

steht. □

6. Klassische Sätze der Funktionalanalysis

6.1 Worum geht's? Zu den klassischen Sätzen der Funktionalanalysis gehört (neben dem Satz von Hahn-Banach) der Satz von Baire mit seinen Folgerungen. Der Satz von Baire ist wesentliches Beweismittel für einige Sätze aus der Operatortheorie, welche wichtige Aussagen über abgeschlossene Operatoren liefern, z.B. der Satz vom stetigen Inversen und der Satz vom abgeschlossenen Graphen.

Weitere klassische Sätze (ebenfalls unter Verwendung des Satzes von Baire beweisbar) sind das Prinzip der gleichmäßigen Beschränktheit bzw. der Satz von Banach-Steinhaus. Dieses Prinzip besagt, dass man aus einer punktweisen Abschätzung bereits eine gleichmäßige Abschätzung einer Familie von Operatoren erhalten kann.

6.2 Satz (Bairescher Kategoriensatz). *Sei (X, d) ein vollständiger metrischer Raum, $A_n \subset X$ abgeschlossen. Falls $A := \bigcup_{n\in\mathbb{N}} A_n$ eine offene Kugel enthält, so existiert ein $n_0 \in \mathbb{N}$ so, dass A_{n_0} (schon) eine offene Kugel enthält.*

Beweis. Sei $B(x_0, r) \subset A$ eine offene Kugel. Angenommen, kein A_n enthält eine offene Kugel (also: jede offene Kugel $B(x, \varepsilon)$ „ragt über jedes A_n hinaus"), d.h.

$$\forall\, n \in \mathbb{N}\;\; \forall\, \varepsilon > 0\, \forall\, x \in X \colon (X \setminus A_n) \cap B(x, \varepsilon) \neq \emptyset.$$

Dann ist einerseits $(X \setminus A_1) \cap B(x_0, r)$ nichtleer (enthält also ein x_1), andererseits offen (als Durchschnitt zweier offener Mengen), also gibt es ein $\varepsilon_1 \in (0, 1/2)$ mit

$$B(x_1, \varepsilon_1) \subset (X \setminus A_1) \cap B(x_0, r).$$

Wähle nun im nächsten Schritt erst ein x_2 und danach ein ε_2, sodass $B(x_2, \varepsilon_2) \subset (X \setminus A_2) \cap B(x_1, \varepsilon_1)$ (offen, nicht leer) und $0 < \varepsilon_2 < \frac{1}{4}$.

Allgemein wähle $x_{n+1}, \varepsilon_{n+1}$ mit $B(x_{n+1}, \varepsilon_{n+1}) \subset (X \setminus A_{n+1}) \cap B(x_n, \varepsilon_n)$ und $0 < \varepsilon_{n+1} < 2^{-n-1}$. Zusätzlich können wir ε_{n+1} so klein wählen, daß $B(x_{n+1}, \varepsilon_{n+1}) \subset B(x_n, \gamma_n \varepsilon_n)$ gilt, mit einer geeigneten Konstanten $\gamma_n < 1$.

Wegen $\varepsilon_n \longrightarrow 0$ und $x_{n+1} \in B(x_n, \varepsilon_n)$ ist $(x_n)_{n\in\mathbb{N}}$ Cauchyfolge, d.h. $x_n \longrightarrow x \in X$. Hier verwenden wir, dass X vollständig ist. Wegen

$$d(x, x_n) = \lim_{m\to\infty} \underbrace{d(x_m, x_n)}_{<\gamma_n\varepsilon_n \text{ falls } m>n} \leq \gamma_n \varepsilon_n < \varepsilon_n$$

haben wir dann auch $x \in \bigcap\limits_{n\in\mathbb{N}} B(x_n, \varepsilon_n)$.

Aber es gilt sowohl

$$\bigcap_{n\in\mathbb{N}} B(x_n, \varepsilon_n) \subset \bigcap_{n\in\mathbb{N}} (X \setminus A_n) = X \setminus A$$

als auch
$$\bigcap_{n\in\mathbb{N}} B(x_n,\varepsilon_n) \subset B(x_1,\varepsilon_1) \subset B(x_0,r) \subset A.$$
Somit ist $\bigcap_{n\in\mathbb{N}} B(x_n,\varepsilon_n)=\emptyset$ im Widerspruch zu $x\in\bigcap_{n\in\mathbb{N}} B(x_n,\varepsilon_n)$. □

6.3 Bemerkung. Eine Menge $A\subset X$ heißt nirgends dicht, falls $\overline{A}$ keine inneren Punkte enthält, d.h. $\mathring{\overline{A}}=\emptyset$. Dies ist äquivalent dazu, dass $\overline{A}$ keine offene Kugel enthält.

Eine Menge $A\subset X$ heißt von erster Kategorie (mager) in X, falls $A\subset\bigcup_{n=1}^\infty A_n$ mit nirgends dichten Mengen A_n gilt. Gibt es keine solche Darstellung, heißt A von zweiter Kategorie.

Damit erhalten wir folgende Formulierung des Satzes von Baire: Sei (X,d) vollständiger metrischer Raum. Falls $A\subset X$ eine offene Kugel enthält, dann ist A von zweiter Kategorie in X. Insbesondere ist X von zweiter Kategorie in sich.

6.4 Satz (Prinzip der gleichmäßigen Beschränktheit)**.** *Sei (X,d) vollständiger metrischer Raum, $\mathcal{T}$ eine Familie stetiger Abbildungen $f\colon X\to\mathbb{K}$. Die Familie $\mathcal{T}$ sei punktweise gleichmäßig beschränkt, d.h. es gilt*
$$\forall\, x\in X\ \exists\, c_x>0\ \forall\, f\in\mathcal{T}\colon\ |f(x)|\le c_x.$$
Dann existiert eine offene Kugel K und ein $c>0$ mit
$$\forall\, x\in K\ \forall\, f\in\mathcal{T}\colon\ |f(x)|\le c.$$

Beweis. Die Menge
$$A_n := \{x\in X\colon\ \forall\, f\in\mathcal{T}\colon |f(x)|\le n\} = \bigcap_{f\in\mathcal{T}}\{x\in X\colon |f(x)|\le n\}$$
ist (als Durchschnitt abgeschlossener Mengen) abgeschlossen. Für $x\in X$ existiert nach Voraussetzung ein $c_x>0$ mit $|f(x)|\le c_x\quad(f\in\mathcal{T})$, d.h. es existiert eine natürliche Zahl n mit $x\in A_n$. Somit ist $X=\bigcup_{n\in\mathbb{N}} A_n$.

Nach dem Satz von Baire existiert ein $n_0\in\mathbb{N}$ und eine offene Kugel $K\subset A_{n_0}$. Damit ist $|f(x)|\le n_0\quad(x\in K, f\in\mathcal{T})$. □

6.5 Satz (von Banach–Steinhaus)**.** *Sei X Banachraum und Y normierter Raum. Sei $\mathcal{T}\subset L(X,Y)$ eine punktweise gleichmäßig beschränkte Familie, d.h. es gelte*
$$\forall\, x\in X\ \exists\, c_x>0\ \forall\, T\in\mathcal{T}\colon\ \|Tx\|_Y\le c_x.$$
Dann existiert ein $c>0$ mit $\|T\|\le c\quad(T\in\mathcal{T})$.

Beweis. Definiere $\mathcal{T}' := \{f_T : T \in \mathcal{T}\}$ als Familie stetiger Abbildungen $f_T : X \to \mathbb{K}$ durch $f_T(x) := \|Tx\|_Y$. Nach Voraussetzung ist diese Familie $\mathcal{T}'$ punktweise gleichmäßig beschränkt.

Nach Satz 6.4 existieren $B(x_0, r_0)$ und $c' > 0$ mit

$$\forall\, x \in B(x_0, r_0)\ \forall\, T \in \mathcal{T}\colon\ \|Tx\|_Y \le c'.$$

Sei nun $x \in X$, $\|x\| = 1$ und $T \in \mathcal{T}$. Dann gilt

$$\begin{aligned}\|Tx\|_Y &= \frac{2}{r_0}\ \Big\|T\Big(\frac{r_0}{2}x - x_0 + x_0\Big)\Big\|_Y \le \\ &\le \frac{2}{r_0}\ \Big(\Big\|T\big(\underbrace{\frac{r_0}{2}x - x_0}_{\in B(x_0,r_0)}\big)\Big\|_Y + \Big\|\underbrace{Tx_0}_{\in B(x_0,r_0)}\Big\|_Y\Big) \le \frac{4}{r_0}c' =: c.\end{aligned}$$

Somit gilt $\|T\| \le c \quad (T \in \mathcal{T})$. □

Man beachte für den folgenden Satz, dass eine Abbildung offen heißt, falls offene Mengen auf offene Mengen abgebildet werden (siehe Definition 1.2).

6.6 Satz (Prinzip der offenen Abbildung)**.** *Seien X, Y Banachräume und $T \in L(X,Y)$ surjektiv. Dann ist T offen.*

Beweis. Wir setzen $B_n := B(0_X, n) \subset X$ für $n \in \mathbb{N}$.

(i) Es existiert ein $\varepsilon > 0$ mit $B(0_Y, \varepsilon) \subset \overline{T(B_1)}$: Da T surjektiv ist, gilt $Y = \bigcup_{n\in\mathbb{N}} \overline{T(B_n)}$. Nach dem Satz von Baire ist das Innere von $\overline{T(B_n)}$ für mindestens ein $n \in \mathbb{N}$ nichtleer, d.h. es existieren $n \in \mathbb{N}$, $\varepsilon_0 > 0$ und $y_0 \in Y$ mit $B(y_0, \varepsilon_0) \subset \overline{T(B_n)}$.

Da T surjektiv ist, existiert ein $x_0 \in X$ mit $Tx_0 = y_0$. Es ist $B(y_0, \varepsilon_0) = Tx_0 + B(0_Y, \varepsilon_0)$ und damit

$$B(0_Y, \varepsilon_0) \subset \overline{T(B_n)} - Tx_0 = \overline{T(nB_1 - x_0)} \subset \overline{T(mB_1)} = m\,\overline{T(B_1)}$$

für ein $m \in \mathbb{N}$. Beachte dabei, dass $nB_1 - x_0 \subset mB_1$ für großes m gilt. Wir erhalten $B(0_Y, \varepsilon) \subset \overline{T(B_1)}$ für $\varepsilon := \varepsilon_0/m$.

(ii) Es gilt $\overline{T(B_1)} \subset T(B_2)$: Dazu sei $y \in \overline{T(B_1)}$ und ε wie im Schritt 1. Nach Definition des Abschlusses gibt es ein $x_1 \in B_1$ mit $y - Tx_1 \in B(0_Y, \frac{\varepsilon}{2})$.

Nach Schritt (i) ist $B(0_Y, \frac{\varepsilon}{2}) \subset \overline{T(B(0_X, \frac{1}{2}))}$. Wähle nun $x_2 \in B(0_X, \frac{1}{2})$ mit

$$y - Tx_1 - Tx_2 \in T(B(0_X, \tfrac{1}{4})) \subset \overline{T(B(0_X, \tfrac{1}{4}))}.$$

Wir erhalten iterativ eine Folge $(x_n)_{n\in\mathbb{N}} \subset X$ mit $x_n \in B(0_X, 2^{-n+1})$ und $y - \sum_{i=1}^n Tx_i \in B(0_Y, 2^{-n}\varepsilon)$. Nach Wahl der x_n ist $x := \sum_{n=1}^\infty x_n$ absolut konvergent. Es gilt $x \in B_2$ wegen $\|x\|_X \le \sum_n \|x_n\|_X < 2$.

Aus der Stetigkeit von T erhalten wir $y = \sum_{i=1}^\infty Tx_i = T(\sum_{i=1}^\infty x_i) = Tx \in T(B_2)$.

(iii) T ist eine offene Abbildung: Sei $U \subset X$ eine offene Menge, und sei $x \in U$. Dann existiert ein $\delta > 0$ mit $B(x, \delta) \subset U$. Nach (ii) gilt $B(0_Y, \varepsilon) \subset T(B_2)$ und damit $B(0_Y, \frac{\delta\varepsilon}{2}) \subset T(B(0_X, \delta))$. Damit gilt

$$B(Tx, \tfrac{\delta\varepsilon}{2}) = Tx + B(0_Y, \tfrac{\delta\varepsilon}{2}) \subset Tx + T(B(0_X, \delta)) = T(B(x, \delta)) \subset T(U).$$

Also ist $T(U) \subset Y$ offen. □

6.7 Korollar. *Seien X, Y Banachräume und $T\colon X \supset D(T) \to Y$ ein abgeschlossener linearer Operator. Falls $R(T)$ abgeschlossen ist, dann ist*

$$T\colon (D(T), \|\cdot\|_X) \to (R(T), \|\cdot\|_Y)$$

offen.

Beweis. Versehen mit der Graphennorm $\|x\|_T := \|x\|_X + \|Tx\|_Y$, ist $D(T)$ ein Banachraum nach Lemma 4.7. Der Operator $\widetilde{T}\colon (D(T), \|\cdot\|_T) \to R(T),\ x \mapsto Tx$, ist surjektiv per Definition, und er ist beschränkt mit Operatornorm ≤ 1, also ist er stetig. Der lineare Raum $R(T)$ ist (als abgeschlossener Unterraum des Banachraums Y) ein Banachraum. Also sind alle Voraussetzungen von Satz 6.6 erfüllt, und

$$\widetilde{T}\colon (D(T), \|\cdot\|_T) \to (R(T), \|\cdot\|_Y)$$

ist eine offene Abbildung.

Sei U eine offene Teilmenge von $(D(T), \|\cdot\|_X)$, und sei $u \in U$. Dann existiert ein $\varepsilon > 0$ so, dass $\{v \in D(T)\colon \|u - v\|_X < \varepsilon\} \subset U$. Dann gilt aber erst recht $\{v \in D(T)\colon \|u - v\|_T < \varepsilon\} \subset U$, d.h. die Menge U ist auch eine offene Teilmenge von $(D(T), \|\cdot\|_T)$. Da $\widetilde{T}$ offen ist, ist $T(U) = \widetilde{T}(U)$ offen in $(R(T), \|\cdot\|_Y)$. □

6.8 Satz (Stetigkeit des Inversen)**.** *Seien X, Y Banachräume und $T\colon X \subset D(T) \to Y$ ein abgeschlossener linearer Operator mit $\ker T = \{0\}$ und $R(T)$ abgeschlossen. Dann ist $T^{-1}\colon (R(T), \|\cdot\|_Y) \to (X, \|\cdot\|_X)$ stetig.*

Beweis. Weil T injektiv ist, existiert $T^{-1}\colon R(T) \to D(T)$ als lineare Bijektion. Die Stetigkeit von T^{-1} ist per Definition äquivalent zur Offenheit von T. Nun wende man Korollar 6.7 an. □

6.9 Satz (Satz vom abgeschlossenen Graphen)**.** *Seien X, Y Banachräume, $T\colon X \supset D(T) \to Y$ abgeschlossener linearer Operator. Falls $D(T)$ abgeschlossen ist, so ist T stetig.*

Beweis. Da T abgeschlossen ist, ist $G(T)$ mit $\|(x,Tx)\|_G := \|x\|_X + \|Tx\|_Y$ als abgeschlossener Unterraum von $X \oplus Y$ ein Banachraum. Die Projektion $\pi_1 \colon G(T) \to X$, $(x, Tx) \mapsto x$, ist injektiv und hat Operatornorm ≤ 1, ist also stetig, und damit ein abgeschlossener linearer Operator. Der Wertebereich $R(\pi_1) = D(T)$ ist abgeschlossen in X. Nach Satz 6.8 ist π_1^{-1} stetig als Abbildung von $(D(T), \|\cdot\|_X)$ nach $(G(T), \|\cdot\|_G)$. Ebenso ist $\pi_2 \colon G(T) \to Y$, $(x,Tx) \mapsto Tx$, stetig als Abbildung von $(G(T), \|\cdot\|_G)$ nach $(R(T), \|\cdot\|_Y)$. Damit ist $T = \pi_2 \circ \pi_1^{-1}$ stetig. □

6.10 Korollar. *Seien $X_1 = (X, \|\cdot\|_1)$ und $X_2 = (X, \|\cdot\|_2)$ Banachräume mit*

$$\|x\|_1 \leq c\,\|x\|_2 \quad (x \in X)$$

für eine Konstante $c > 0$.

Dann sind die Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ äquivalent, d.h. es gibt eine Konstante $c' > 0$ mit $c'\,\|x\|_2 \leq \|x\|_1 \leq c\,\|x\|_2 \quad (x \in X)$.

Beweis. Satz 6.8, angewandt auf $\mathrm{id}\colon (X, \|\cdot\|_2) \to (X, \|\cdot\|_1)$, $x \mapsto x$. □

6.11 Bemerkung. Sei X Banachraum, Y normierter Raum und $T\colon X \to Y$ ein Isomorphismus normierter Räume, d.h. T linear, bijektiv und T, T^{-1} stetig. Dann ist auch Y ein Banachraum. Denn falls $(y_n)_{n\in\mathbb{N}} \subset Y$ eine Cauchyfolge ist, so auch $(x_n)_{n\in\mathbb{N}}$ mit $x_n := T^{-1}y_n$. Damit existiert $x := \lim_n x_n \in X$, und für $y := Tx \in Y$ gilt $y_n \longrightarrow y$.

Man beachte, dass hier die Linearität entscheidend ist, vergleiche $\arctan\colon \mathbb{R} \to (-\pi/2, \pi/2)$.

6.12 Lemma. *Seien X, Y Banachräume und $T\colon X \supset D(T) \to Y$ abgeschlossener linearer Operator. Dann sind äquivalent:*

(i) *Es existiert ein $C > 0$ mit $\|Tx\|_Y \geq C\,\|x\|_X \quad (x \in D(T))$.*

(ii) *T ist injektiv und $R(T)$ ist abgeschlossen.*

Beweis. (i)$\Longrightarrow$(ii). Der Operator $T\colon (D(T), \|\cdot\|_T) \to (R(T), \|\cdot\|_Y)$ ist surjektiv per Definition und offensichtlich injektiv. Weiterhin ist er beschränkt mit Operatornorm ≤ 1, also stetig. Sein Inverses ist ebenfalls ein beschränkter Operator, wegen (i).

Also ist $T\colon (D(T), \|\cdot\|_T) \to (R(T), \|\cdot\|_Y)$ ein Isomorphismus von normierten Räumen. Da $(D(T), \|\cdot\|_T)$ Banachraum ist (denn T ist abgeschlossener Operator), ist $R(T)$ abgeschlossen nach Bemerkung 6.11.

(ii)$\Longrightarrow$(i) folgt direkt aus dem Satz vom stetigen Inversen. □

7. Nützliches über das Spektrum

In vielen Anwendungen ist es wichtig, das Spektrum eines Operators gut zu kennen. So kann man etwa am Spektrum ablesen, ob die Lösung einer Differentialgleichung für große Zeiten gegen Null konvergiert (Stabilität). Es gibt eine ganze Reihe kleiner Aussagen über das Spektrum, welche bei der Bestimmung des Spektrum helfen können. Am einfachsten (aus der Sicht des Spektrums gesehen) sind selbstadjungierte Operatoren. Zum Glück sind das auch die Operatoren, welche am häufigsten vorkommen. Für Anwendungen sehr nützlich ist auch der Begriff des approximativen Spektrums.

Im folgenden seien X und Y $\mathbb{C}$-Hilberträume.

7.1 Satz. *Sei $T\colon X \supset D(T) \to Y$ ein dicht definierter Operator. Dann gilt*

a) $R(T)^\perp = \overline{R(T)}^\perp = \ker T^*$.

b) $\overline{R(T)} = (\ker T^*)^\perp$.

Beweis. a) Es gilt $y \in R(T)^\perp$ genau dann, wenn für alle $x \in D(T)$ gilt $\langle Tx, y\rangle = 0$. Dies ist äquivalent zu $y \in D(T^*)$ und $T^*y = 0$, also zu $y \in \ker T^*$.

b) Nach a) gilt $\overline{R(T)} = (R(T))^{\perp\perp} = (\ker T^*)^\perp$.

□

7.2 Bemerkung. Falls $T\colon X \supset D(T) \to X$ ein dicht definierter und abgeschlossener Operator ist, so ist T^* ebenfalls abgeschlossen und dicht definiert, und es gilt $T^{**} = T$. Wendet man Satz 7.1 auf T^* an, so erhält man

$$R(T^*)^\perp = \ker T, \quad \overline{R(T^*)} = (\ker T)^\perp.$$

7.3 Lemma. *a) Sei $T \in L(X)$. Dann ist T genau dann selbstadjungiert, falls $\langle Tx, x\rangle \in \mathbb{R}$ gilt für alle $x \in X$.*

b) Sei $T\colon X \supset D(T) \to X$ dicht definiert. Dann ist T genau dann symmetrisch, falls $\langle Tx, x\rangle \in \mathbb{R}$ $(x \in D(T))$.

Beweis. a) (i) Sei $T = T^*$. Dann ist $\langle Tx, x\rangle = \langle x, Tx\rangle = \overline{\langle Tx, x\rangle} \in \mathbb{R}$.

(ii) Seien $x, y \in X$ und $\alpha \in \mathbb{C}$. Nach Voraussetzung ist

$$\langle T(x+\alpha y), x+\alpha y\rangle = \overline{\langle T(x+\alpha y), x+\alpha y\rangle}$$

und damit
$$\alpha \langle Ty, x\rangle + \overline{\alpha}\langle Tx, y\rangle = \alpha \langle y, Tx\rangle + \overline{\alpha}\langle x, Ty\rangle .$$
Für $\alpha = 1$ erhält man
$$\langle Ty, x\rangle + \langle Tx, y\rangle = \langle y, Tx\rangle + \langle x, Ty\rangle .$$
Für $\alpha = \mathrm{i}$ erhält man
$$\langle Ty, x\rangle - \langle Tx, y\rangle = \langle y, Tx\rangle - \langle x, Ty\rangle .$$
Somit folgt
$$\langle Ty, x\rangle = \langle y, Tx\rangle \quad (x, y \in X).$$
Nach Definition von T^* gilt also $T = T^*$.

b) Die Rechnung unter a) zeigt, dass $\langle Tx, y\rangle = \langle x, Ty\rangle$ $(x, y \in D(T))$ gilt. Dies ist äquivalent zu $T \subset T^*$, d.h. zur Symmetrie von T. $\square$

7.4 Satz (Spektrum selbstadjungierter Operatoren)**.** *Sei T ein selbstadjungierter linearer (nicht notwendig beschränkter) Operator mit dichtem Definitionsbereich $D(T)$. Dann gilt*

(i) $\sigma(T) \subset \mathbb{R}$.

(ii) $\|(T-\lambda)^{-1}\| \le |\operatorname{Im}\lambda|^{-1} \quad (\lambda \in \mathbb{C}\setminus\mathbb{R})$.

(iii) *Für $\lambda \in \sigma_p(T)$ sind geometrischer und algebraischer Eigenraum identisch.*

(iv) *Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal.*

(v) $\sigma_r(T) = \emptyset$.

Beweis. (i) Sei $\lambda \in \mathbb{C}\setminus\mathbb{R}$ und $x \in D(T)$. Nach Lemma 7.3 gilt $\operatorname{Im}\langle Tx, x\rangle = 0$. Dann ist mit Cauchy–Schwarz
$$\|(T-\lambda)x\| \cdot \|x\| \ge |\langle (T-\lambda)x, x\rangle| \ge |\operatorname{Im}\langle (T-\lambda)x, x\rangle| = |\operatorname{Im}\lambda| \cdot \|x\|^2 . \qquad (7\text{-}1)$$
Nach Lemma 6.12 ist $T-\lambda$ injektiv und $R(T-\lambda)$ abgeschlossen. Weiterhin ist auch $T-\overline{\lambda}$ injektiv wegen $\overline{\lambda} \in \mathbb{C}\setminus\mathbb{R}$.

$T-\lambda$ ist surjektiv, denn mit Satz 7.1 b) haben wir
$$R(T-\lambda) = \overline{R(T-\lambda)} = (\ker((T-\lambda)^*))^\perp = \big(\ker(T-\overline{\lambda})\big)^\perp = \{0\}^\perp = X.$$

Insgesamt haben wir gezeigt, dass für $\lambda \in \mathbb{C}\setminus\mathbb{R}$ der Operator $T-\lambda$ bijektiv ist.

(ii) Setze $y := (T-\lambda)x$ in (7-1) und erhalte
$$\|y\| \ge |\operatorname{Im}\lambda| \cdot \big\|(T-\lambda)^{-1}y\big\| .$$

(iii) Sei $N_\lambda^{(a)}(T)$ der algebraische Eigenraum zum Eigenwert λ von T. Die Inklusion $\ker(T-\lambda) \subset N_\lambda^{(a)}(T)$ gilt immer. Sei also $x \in N_\lambda^{(a)}(T) \setminus \ker(T-\lambda)$ mit $\lambda \in \mathbb{R}$. Dann gilt $x \in D((T-\lambda)^n)$ und $(T-\lambda)^n x = 0$ für ein $n \geq 2$, aber $(T-\lambda)x \neq 0$. Wegen $T = T^*$ und $\lambda \in \mathbb{R}$ ist dann

$$\left\|(T-\lambda)^{n-1}x\right\|^2 = \left\langle (T-\lambda)^n x, (T-\lambda)^{n-2}x \right\rangle = 0.$$

Induktiv folgt $(T-\lambda)^{n-2}x = 0, \ldots, (T-\lambda)x = 0$, Widerspruch.

(iv) Das folgt wie in der linearen Algebra. Seien x_1, x_2 Eigenvektoren zu $\lambda_1 \neq \lambda_2$ mit $\lambda_{1,2} \in \mathbb{R}$. Dann gilt

$$\lambda_1 \langle x_1, x_2 \rangle = \langle Tx_1, x_2 \rangle = \langle x_1, Tx_2 \rangle = \lambda_2 \langle x_1, x_2 \rangle .$$

Wegen $\lambda_1 \neq \lambda_2$ folgt $\langle x_1, x_2 \rangle = 0$.

(v) Angenommen, es existiert ein $\lambda \in \sigma_r(T)$. Dann ist $\lambda \in \mathbb{R}$ wegen (i), sowie $T-\lambda$ injektiv und $\overline{R(T-\lambda)} \neq X$ wegen der Definition von σ_r. Schließlich ist dann

$$\overline{R(T-\lambda)} = \left((R(T-\lambda))^\perp\right)^\perp = (\ker((T-\lambda)^*))^\perp = (\ker(T-\lambda))^\perp = \{0\}^\perp = X.$$

Widerspruch. $\square$

Der folgende Satz wird genauso wie Satz 7.4 bewiesen.

7.5 Satz (Spektrum unitärer Operatoren)**.** *Sei X ein Hilbertraum und $T \in L(X)$ unitär. Dann gilt*

(i) $\sigma(T) \subset \{\lambda \in \mathbb{C} \colon |\lambda| = 1\}$.

(ii) $\|(T-\lambda)^{-1}\| \leq \frac{1}{\left||\lambda|-1\right|}$ *für* $|\lambda| \neq 1$.

Die Aussagen (iii)–(v) *von Satz 7.4 gelten analog.*

7.6 Beispiel. Sei T ein selbstadjungierter (nicht unbedingt beschränkter) Operator in X. Dann heißt $U := (T-\mathrm{i})(T+\mathrm{i})^{-1}$ die Cayley–Transformierte des Operators T. Dieser Operator U ist unitär. Die Cayley–Transformation ist umkehrbar: sei $U \in L(X)$ ein unitärer Operator, für den $1-U$ injektiv ist. Dann ist der Operator $T := \mathrm{i}(1+U)(1-U)^{-1}$ selbstadjungiert mit Definitionsbereich $R(1-U)$, und seine Cayley–Transfomierte ist wieder gleich U. Der Nutzen dieser Transformation besteht darin, dass sich einige Aussagen über unbeschränkte selbstadjungierte Operatoren gewinnen lassen, wenn man ihre Cayley–Transformierten studiert, welche den Vorteil haben, beschränkt zu sein.

7.7 Definition (numerischer Wertebereich)**.** Für $T \in L(X)$ ist der numerische Wertebereich definiert durch

$$W(T) := \{\langle Tx, x\rangle : \|x\| = 1\}.$$

7.8 Lemma. *Sei X ein Hilbertraum und $T \in L(X)$. Dann gilt $\sigma(T) \subset \overline{W(T)}$.*

Beweis. Sei $\lambda \notin \overline{W(T)}$. Wir wollen zeigen, dass $\lambda \in \rho(T)$. Für $\|x\| = 1$ gilt nun

$$\begin{aligned} 0 < d := \operatorname{dist}(\lambda, \overline{W(T)}) &\le |\lambda - \langle Tx, x\rangle\,| = |\,\langle(\lambda - T)x, x\rangle\,| \\ &\le \|(T-\lambda)x\| \cdot \|x\| = \|(T-\lambda)x\|\,. \end{aligned}$$

Damit haben wir $\|(T-\lambda)x\| \ge d\,\|x\| \quad (x \in X)$. Also ist nach Lemma 6.12 der Operator $T-\lambda$ injektiv und $R(T-\lambda)$ abgeschlossen. Falls $T-\lambda$ nicht surjektiv ist, dann existiert ein $x_0 \in R(T-\lambda)^\perp$ mit $\|x_0\| = 1$, und es ist

$$0 = \langle (T-\lambda)x_0, x_0\rangle = \langle Tx_0, x_0\rangle - \lambda,$$

was im Widerspruch steht zu $\lambda \notin \overline{W(T)}$. □

7.9 Lemma (approximative Eigenwerte). *Sei $T\colon X \supset D(T) \to X$ ein abgeschlossener linearer Operator. Die Menge der approximativen Eigenwerte ist definiert als*

$$\sigma_{\text{app}}(T) := \{\lambda \in \mathbb{C}\colon\ \exists\, (x_n)_{n\in\mathbb{N}} \subset D(T),\ \|x_n\| = 1 :\ \|(T-\lambda)x_n\| \longrightarrow 0\}.$$

Dann gilt

$$\sigma_p(T) \cup \sigma_c(T) \subset \sigma_{\text{app}}(T) \subset \sigma(T).$$

Insbesondere gilt für selbstadjungierte Operatoren T

$$\sigma(T) = \sigma_p(T) \cup \sigma_c(T) = \sigma_{\text{app}}(T).$$

Beweis. (i) Sei $\lambda \in \sigma_{\text{app}}(T)$. Falls $\lambda \in \rho(T)$, so wäre $(T-\lambda)^{-1}$ stetig, d.h. für alle $x_n \in D(T) \setminus \{0\}$ ist

$$\frac{\|x_n\|}{\|(T-\lambda)x_n\|} \le \big\|(T-\lambda)^{-1}\big\| < \infty.$$

Dies ist aber ein Widerspruch zur Definition der approximativen Eigenwerte.

(ii) Für $\lambda \in \sigma_p(T)$ setze $x_n := x$ mit einem normierten Eigenvektor x zum Eigenwert λ.

Sei $\lambda \in \sigma_c(T)$. Dann ist $T-\lambda$ injektiv und $R(T-\lambda)$ nicht abgeschlossen. Nach Lemma 6.12 existiert kein $C > 0$ mit $\|(T-\lambda)x\| \ge C\,\|x\|$. Somit existiert eine Folge $(x_n)_{n\in\mathbb{N}}$ mit $\|x_n\| = 1$ und $\|(T-\lambda)x_n\| \longrightarrow 0$. □

7.10 Lemma. *Sei $T \in L(X)$ selbstadjungiert. Für $m := \inf_{\|x\|=1} \langle Tx, x\rangle$ und $M := \sup_{\|x\|=1} \langle Tx, x\rangle$ gilt $\sigma(T) \subset [m, M]$ und $m, M \in \sigma(T)$.*

Beweis. Die Inklusion $\sigma(T) \subset \overline{W(T)} \subset [m, M]$ gilt nach Lemma 7.8.

Sei $(x_n)_{n\in\mathbb{N}} \subset X$ mit $\|x_n\| = 1$ und $\langle Tx_n, x_n\rangle \longrightarrow m$. Nach Definition von m ist die Bilinearform $[x, y] := \langle (T - m)x, y\rangle$ positiv semidefinit, und nach Cauchy–Schwarz (angewandt auf $[\cdot, \cdot]$) gilt

$$\begin{aligned}\|(T-m)x_n\|^2 &= [x_n, (T-m)x_n] \le [x_n, x_n]^{1/2}[(T-m)x_n, (T-m)x_n]^{1/2}\\ &= \langle (T-m)x_n, x_n\rangle^{1/2} \cdot \big\langle (T-m)^2 x_n, (T-m)x_n\big\rangle^{1/2} \longrightarrow 0 \quad (n \longrightarrow \infty),\end{aligned}$$

da $\langle (T-m)x_n, x_n\rangle \longrightarrow 0$ und $\langle (T-m)^2 x_n, (T-m)x_n\rangle$ beschränkt ist. Also ist $m \in \sigma_{\text{app}}(T) \subset \sigma(T)$. Genauso zeigt man $M \in \sigma(T)$. □

7.11 Definition und Satz. *Sei $T \in L(X)$. Für den Spektralradius*

$$r(T) := \inf_{n\in\mathbb{N}} \|T^n\|^{1/n}$$

gilt

$$r(T) = \lim_{n\to\infty} \|T^n\|^{1/n}.$$

Beweis. **(i)** Wir zeigen folgende Aussage: Sei $(a_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ mit $0 \le a_{n+m} \le a_n a_m$ für alle $n, m \in \mathbb{N}$. Dann gilt $(a_n)^{1/n} \longrightarrow a := \inf_n (a_n)^{1/n} \quad (n \longrightarrow \infty)$.

Um dies zu beweisen, sei $\varepsilon > 0$. Wähle $N \in \mathbb{N}$ mit $(a_N)^{1/N} < a + \varepsilon$ und setze $b(\varepsilon) := \max\{a_1, \ldots, a_N\}$. Schreibe nun $n \in \mathbb{N}$ in der Form $n = kN + r$ mit $0 \le r < N$. Dann gilt

$$\begin{aligned}(a_n)^{1/n} &= (a_{kN+r})^{1/n} \le (a_N^k a_r)^{1/n}\\ &\le (a+\varepsilon)^{kN/n} b^{1/n} = (a+\varepsilon)(a+\varepsilon)^{-r/n} b^{1/n}\\ &= (a+\varepsilon)\Big(\frac{b}{(a+\varepsilon)^r}\Big)^{1/n}\\ &< a + 2\varepsilon,\end{aligned}$$

falls n hinreichend groß ist.

(ii) Setzt man $a_n := \|T^n\|$, so gilt $0 \le a_{n+m} \le a_n a_m$ $(n, m \in \mathbb{N})$ wegen der Submultiplikativität der Operatornorm, und mit (i) folgt $r(T) = \lim_{n\to\infty} \|T^n\|^{1/n}$. □

7.12 Satz. *Sei $T \in L(X)$ selbstadjungiert. Dann gilt*

$$r(T) = \|T\| = \max\{|\lambda| \colon \lambda \in \sigma(T)\} = \sup\{|\langle Tx, x\rangle| \colon \|x\| = 1\}.$$

Beweis. **(i)** Es gilt

$$\begin{aligned}\|T\| &= \sup_{x\in X,\,\|x\|=1}\|Tx\| = \sup_{x\in X,\,\|x\|=1}(\langle Tx,Tx\rangle)^{1/2}\\ &= \sup_{x\in X,\,\|x\|=1}\big(\langle T^2x,x\rangle\big)^{1/2} \le \big\|T^2\big\|^{1/2},\end{aligned}$$

d.h. $\|T\|^2 \le \|T^2\|$. Aus der Submultiplikativität der Operatornorm folgt die andere Ungleichung, d.h. es gilt $\|T^2\| = \|T\|^2$. Iterativ folgt damit $\|T^{2n}\| = \|T\|^{2n}$ für $n\in\mathbb{N}$. Damit ist $r(T) = \lim_{n\to\infty}\|T^n\|^{1/n} = \|T\|$.

(ii) Sei $M(T) := \max\{|\lambda|\colon \lambda\in\sigma(T)\}$. Nach Lemma 7.10 gilt $M(T) = \sup\{|\langle Tx,x\rangle|\colon \|x\|=1\}$.

(iii) Für $|\lambda| > \|T\|$ ist $\lambda\in\rho(T)$ (Neumann-Reihe, vgl. Satz 4.16). Somit gilt $M(T)\le\|T\|$.

(iv) Zeige $\|T\|\le\sup\{|\langle Tx,x\rangle|\colon \|x\|=1\}$: Für $x,y\in X$ gilt

$$\langle T(x\pm y),x\pm y\rangle = \langle Tx,x\rangle \pm 2\,\mathrm{Re}\langle Tx,y\rangle + \langle Ty,y\rangle$$

und damit

$$\begin{aligned}\mathrm{Re}\langle Tx,y\rangle &= \frac14\big(\langle T(x+y),x+y\rangle - \langle T(x-y),x-y\rangle\big)\\ &\le \frac{M(T)}{4}\big(\|x+y\|^2+\|x-y\|^2\big)\\ &= \frac{M(T)}{2}\big(\|x\|^2+\|y\|^2\big),\end{aligned}$$

wobei die Parallelogrammgleichung verwendet wurde. Wählt man $x\in X$ mit $\|x\|=1$ und setzt $y := \frac{Tx}{\|Tx\|}$, so erhält man

$$\|Tx\| = \mathrm{Re}\langle Tx,y\rangle \le \frac{M(T)}{2}\big(\|x\|^2+\|y\|^2\big) = M(T)$$

und damit $\|T\|\le M(T)$. □

7.13 Bemerkung. a) Für nicht selbstadjungierte Operatoren gilt die Aussage von Satz 7.12 i. allg. nicht, wie man an folgendem Beispiel sieht. Sei $X = L^2([0,1])$ und

$$(Ax)(t) := \int_0^t x(\tau)\,\mathrm{d}\tau.$$

Der Operator A ist ein Beispiel eines Volterra–Operators. Es gilt $\sigma(A)=\{0\}$.

b) Es gibt selbstadjungierte Operatoren, welche keinen Eigenwert besitzen. Betrachte dazu wieder $X = L^2([0,1])$ und $(Ax)(t) := tx(t)$ (Multiplikationsoperator). Dann ist $\sigma_p(A)=\emptyset$.

7.14 Bemerkung. Die meisten obigen Aussagen und Beweise gelten nicht nur für beschränkte Operatoren in Hilbert- bzw. Banachräumen, sondern allgemein für Elemente von Banachalgebren bzw. C^*-Algebren. Die sog. Gelfand–Theorie von C^*-Algebren ermöglicht es, einen abstrakten Zugang zu Spektrum und Spektralsatz zu finden.

7.15 Beispiel (Multiplikationsoperator)**.** Sei $X = L^2((0,3))$ und für $f \in X$ sei Tf definiert durch $(Tf)(x) := \psi(x)f(x)$ $(x \in (0,3))$. Dabei sei die Funktion ψ definiert durch

$$\psi(x) := \begin{cases} x, & \text{falls } x \in (0,1], \\ 1, & \text{falls } x \in [1,2], \\ x-1, & \text{falls } x \in [2,3). \end{cases}$$

Wir untersuchen das Spektrum des Operators T.

(i) Offensichtlich ist T linear, und wegen

$$\|Tf\|^2 = \int_0^3 |\psi(x)f(x)|^2 dx \le \|\psi\|_\infty^2 \int_0^3 |f(x)|^2 dx = 4\|f\|^2$$

ist $T \in L(X)$ mit $\|T\| \le 2$.

(ii) Es gilt für alle $f \in X$

$$\langle Tf, f\rangle = \int_0^3 \psi(x)|f(x)|^2 dx \in \mathbb{R}.$$

Nach Lemma 7.3 ist T selbstadjungiert.

(iii) Wegen $0 \le \psi(x) \le 2$ $(x \in (0,3))$ gilt für alle $f \in X$ mit $\|f\| = 1$

$$\langle Tf, f\rangle = \int_0^3 \psi(x)|f(x)|^2 dx \in [0,2].$$

Also gilt für den numerischen Wertbereich $W(T) \subset [0,2]$ und damit $\sigma(T) \subset [0,2]$.

(iv) Sei $\lambda \in (0,2) \setminus \{1\}$. Dann ist die Menge $\psi^{-1}(\{\lambda\})$ einelementig und damit eine Nullmenge. Somit gilt: Falls $f \in X$ mit $Tf = \lambda f$, so gilt $(\psi(x) - \lambda)f(x) = 0$ fast überall und daher $f(x) = 0$ fast überall, d.h. $f = 0$ in X. Also ist λ kein Eigenwert.

Sei andererseits $\lambda = 1$. Dann ist z.B. $f(x) := \chi_{(1,2)}(x)$ ein Eigenvektor von T. Somit ist $1 \in \sigma_p(T)$. (Man sieht sofort, dass die geometrische Vielfachheit ∞ ist.)

(v) Sei $\lambda \in (0,2)$, und sei $x_0 \in (0,3)$ mit $\psi(x_0) = \lambda$. Da ψ Lipschitz-stetig mit Konstante 1 ist, gilt $|\psi(x) - \psi(x_0)| \le \frac{1}{n}$ falls $|x - x_0| \le \frac{1}{n}$. Zu $n \in \mathbb{N}$ sei $I_n \subset (0,3)$ ein offenes Intervall der Länge $\frac{1}{n}$ mit $x_0 \in I_n$. Für $f_n := \sqrt{n}\chi_{I_n}$ gilt dann $f_n \in X$ mit $\|f_n\| = 1$ sowie

$$\|(T-\lambda)f_n\|^2 = \int_{I_n} |\psi(x) - \psi(x_0)|^2 n dx \le \int_{I_n} \tfrac{1}{n^2} n dx = \tfrac{1}{n^2} \to 0 \ (n \to \infty).$$

Also gilt $\lambda \in \sigma_{\text{app}}(T) = \sigma(T)$.

(vi) Da $\sigma(T)$ abgeschlossen ist, folgt $0 \in \sigma(T)$ und $2 \in \sigma(T)$. Wie in (iv) folgt, dass weder 0 noch 2 ein Eigenwert von T ist.

Insgesamt haben wir also $\sigma_p(T) = \{1\}$, $\sigma_c(T) = [0,2]\setminus\{1\}$ und (natürlich) $\sigma_r(T) = \emptyset$. Mit Satz 7.12 folgt nun auch $\|T\| = \max_{\lambda\in\sigma(T)} |\lambda| = 2$.

7.16 Beispiel (Laplace-Operator)**.** Man definiert den Laplace-Operator in $\mathbb{R}^n$ durch

$$D(\Delta) := H^2(\mathbb{R}^n) \subset L^2(\mathbb{R}^n)$$

und $\Delta\colon L^2(\mathbb{R}^n) \supset D(\Delta) \to L^2(\mathbb{R}^n)$, $u \mapsto \Delta u := \sum_{j=1}^n \frac{\partial^2}{\partial {x_j}^2} u$.

Dann ist Δ ein selbstadjungierter Operator, und es gilt

$$\sigma(\Delta) = \sigma_c(\Delta) = (-\infty, 0].$$

Der Beweis erfordert einige Hilfsmittel, die nicht in dieser Vorlesung besprochen werden. Es ist aber leicht zu sehen, dass Δ symmetrisch ist.

Falls $G \subset \mathbb{R}^n$ ein beschränktes Gebiet ist, so wird der Dirichlet-Laplace-Operator definiert durch

$$D(\Delta_D) := H^2(G) \cap H_0^1(G) \subset L^2(G), \quad \Delta_D u := \Delta u.$$

Wieder ist Δ_D selbstadjungiert, und es gilt $\sigma(\Delta_D) = \sigma_p(\Delta_D) = \{\lambda_j\colon j \in \mathbb{N}\}$ mit $0 > \lambda_1 \geq \lambda_2 \geq \ldots$ und $\lambda_j \to -\infty$ $(j \to \infty)$.

8. Der Spektralsatz für selbstadjungierte Operatoren

8.1 Worum geht's? Der Spektralsatz ist einer der wichtigsten Sätze der Operatortheorie. Es handelt sich dabei um eine Verallgemeinerung des bekannten Satzes aus der Linearen Algebra, nach welchem selbstadjungierte Matrizen orthogonal diagonalisierbar sind, also zunächst einmal um eine Strukturaussage. Diese Darstellung linearer selbstadjungierter Operatoren kann nun verwendet werden, um etwa Funktionen von Operatoren zu definieren, was wichtige Anwendungen z.B. für partielle Differentialgleichungen besitzt. Man erhält einen Funktionalkalkül für Operatoren.

a) Motivation und orthogonale Projektionen

Im Folgenden sei H ein $\mathbb{C}$-Hilbertraum. Die orthogonale Diagonalisierbarkeit einer hermiteschen Matrix kann mit Hilfe orthogonaler Projektionen geschrieben werden. Dazu zunächst ein kleines Lemma:

8.2 Definition. Sei $M \subset H$ ein abgeschlossener Unterraum. Dann heißt die Abbildung $P\colon H \to H$, $x \mapsto x_1$ mit $x = x_1 + x_2$, $x_1 \in M$, $x_2 \in M^\perp$, die orthogonale Projektion von H auf M.

Man beachte, dass die Abbildung nach dem Projektionssatz wohldefiniert ist.

8.3 Lemma. *a) Sei $M \subset H$ abgeschlossener Unterraum und P die orthogonale Projektion von H auf M. Dann ist $P \in L(H)$ mit*

$$\|P\|_{L(H)} = \begin{cases} 1, & \text{falls } M \neq \{0\}, \\ 0, & \text{falls } M = \{0\}. \end{cases}$$

Es gilt $\ker P = M^\perp$ und $R(P) = M$.

b) Ein Operator $P \in L(H)$ ist genau dann eine orthogonale Projektion, wenn $P^2 = P = P^*$ gilt.

Beweis. a) Offensichtlich ist P linear. Nach dem Satz von Pythagoras gilt $\|Px\|^2 = \|x_1\|^2 \leq \|x_1\|^2 + \|x_2\|^2 = \|x\|^2$ $(x \in H)$ und damit $P \in L(H)$ mit $\|P\|_{L(H)} \leq 1$. Falls $M = \{0\}$, so ist $P = 0$. Ansonsten gilt für $x \in M \setminus \{0\}$ die Gleichheit $Px = x$ und damit $\|P\|_{L(H)} = 1$.

Direkt nach Definition von P folgt $R(P) = M$ und $\ker P = M^\perp$.

b) (i) Sei P die orthogonale Projektion auf M. Die Gleichheit $P^2 = P$ ist klar nach Definition von P. Seien $x, y \in H$ mit $x = x_1 + x_2$, $y = y_1 + y_2$, wobei $x_1, y_1 \in M$

und $x_2, y_2 \in M^\perp$. Dann gilt

$$\langle Px, y\rangle = \langle x_1, y_1 + y_2\rangle = \langle x_1, y_1\rangle + \underbrace{\langle x_1, y_2\rangle}_{=0} = \langle x, y_1\rangle = \langle x, Py\rangle\,,$$

d.h. es ist tatsächlich $P = P^*$.

(ii) Sei nun $P = P^* = P^2 \in L(H)$. Setze $M := R(P)$. Für eine Folge $(y_n)_{n\in\mathbb{N}} \subset M$ mit $y_n \longrightarrow y$ existiert eine Folge $(x_n)_{n\in\mathbb{N}} \subset H$ mit $y_n = Px_n$, und es ist

$$Py_n = P^2x_n = Px_n = y_n, \tag{8-1}$$

und damit

$$\|y_n - Py\| = \|P(y_n - y)\| \le \|P\|_{L(H)} \cdot \|y_n - y\| \longrightarrow 0 \quad (n \longrightarrow \infty),$$

d.h. $Py = y$ wegen der Eindeutigkeit des Grenzwertes. Somit ist M abgeschlossen. Aus Satz 7.1 und $P = P^*$ folgt dann $M^\perp = \ker P$.

Weil M abgeschlossen ist, existiert die orthogonale Projektion $\widetilde{P}$ auf den Unterraum M, und diese erfüllt $\widetilde{P} = \widetilde{P}^2 = \widetilde{P}^*$ wegen Schritt b) (i).

Für $x, y \in H$ haben wir die Zerlegungen

$$x = x^{(1)} + x^{(2)}, \qquad y = y^{(1)} + y^{(2)}, \qquad x^{(1)}, y^{(1)} \in M, \quad x^{(2)}, y^{(2)} \in M^\perp,$$

und es folgt mit (8-1)

$$\langle \widetilde{P}x, y\rangle = \langle x, \widetilde{P}y\rangle = \langle x, y^{(1)}\rangle = \langle x, P(y^{(1)} + y^{(2)})\rangle = \langle Px, y\rangle.$$

Also gilt $\widetilde{P} = P$, und P ist eine orthogonale Projektion. □

8.4 Beispiel. Sei $v \in H$ mit $\|v\| = 1$. Dann ist $P\colon X \to X$, $x \mapsto \langle x, v\rangle v$ die orthogonale Projektion auf $\operatorname{span}\{v\}$. Denn es gilt $R(P) = \operatorname{span}\{v\}$ und $P^2 = P = P^*$ wegen

$$\begin{aligned} P^2x &= \Big\langle \langle x, v\rangle v, v\Big\rangle v = \langle x, v\rangle\, \|v\|^2\, v = Px, \\ \langle Px, y\rangle &= \Big\langle \langle x, v\rangle v, y\Big\rangle = \langle x, v\rangle\langle v, y\rangle = \Big\langle x, \langle y, v\rangle v\Big\rangle = \langle x, Py\rangle. \end{aligned}$$

Sei nun $\{v_1, \dots, v_n\} \subset H$ ein Orthonormalsystem, und sei $P_j := \langle \cdot, v_j\rangle v_j$ die orthogonale Projektion auf $\operatorname{span}\{v_j\}$ für $j = 1, \dots, n$. Dann gilt $P_jP_k = 0$ falls $j \ne k$, und

$$P\colon X \to X,\ x \mapsto \sum_{j=1}^n \langle x, v_j\rangle v_j$$

ist die Orthogonalprojektion auf $M := \operatorname{span}\{v_1, \dots, v_n\}$ wegen $P^2 = P = P^*$ und $R(P) = M$.

8.5 Bemerkung. Wir beginnen mit einer Darstellung bekannter Ergebnisse aus der linearen Algebra. Sei $T = T^* \in L(\mathbb{C}^n) = \mathbb{C}^{n\times n}$ eine hermitesche Matrix. Dann existiert eine Orthonormalbasis $v_1, \dots, v_n$ von Eigenvektoren von T zu Eigenwerten λ_j. Sei $P_j = \langle\cdot, v_j\rangle v_j$ die Projektion auf $\operatorname{span}\{v_j\}$. Dann gelten für alle $x \in \mathbb{C}^n$ wegen der Basiseigenschaft und wegen $Tv_j = \lambda_j v_j$ die Darstellungen

$$\begin{aligned} x &= \sum_{j=1}^n \langle x, v_j\rangle v_j = \sum_{j=1}^n P_j x, \\ Tx &= \sum_{j=1}^n \lambda_j \langle x, v_j\rangle v_j = \sum_{i=1}^n \lambda_j P_j x. \end{aligned}$$

Man kann diese Darstellung noch nach gleichen Eigenwerten zusammenfassen: Sei $\lambda \in \sigma(T)$ ein (eventuell mehrfacher) Eigenwert von T, etwa $\lambda = \lambda_1 = \dots = \lambda_d$ mit $d \le n$. Dann ist $E(\{\lambda\}) := P_1 + \dots + P_d$ nach Beispiel 8.4 die Orthogonalprojektion auf $\operatorname{span}\{v_1, \dots, v_d\} = \ker(T - \lambda)$, d.h. auf den Eigenraum zum Eigenwert λ. Insgesamt erhält man für alle $x \in H$

$$\begin{aligned} x &= \sum_{\lambda\in\sigma(T)} E(\{\lambda\})x, \\ Tx &= \sum_{\lambda\in\sigma(T)} \lambda E(\{\lambda\})x. \end{aligned}$$

Ein Vorteil dieser Darstellung liegt darin, dass Funktionen der Matrix T einfach dargestellt werden können. Z.B. gilt $T^N x = \sum_{\lambda\in\sigma(T)} \lambda^N E(\{\lambda\})x$ und $\exp(T)x = \sum_{\lambda\in\sigma(T)} e^\lambda E(\{\lambda\})x$. Allgemein kann für eine beliebige Funktion $f\colon \sigma_p(T) \to \mathbb{C}$ durch

$$f(T)x := \sum_{\lambda\in\sigma(T)} f(\lambda)E(\{\lambda\})x \quad (x \in \mathbb{C}^n)$$

eine Matrix $f(T)$ definiert werden. Die obigen Darstellungen für x, Tx und die Definition für $f(T)x$ sollen nun auf Operatoren $T = T^* \in L(H)$ übertragen werden.

Die Menge der orthogonalen Projektionen besitzen eine partielle Ordnung, welche durch folgendes Lemma charakterisiert wird.

8.6 Lemma. *Seien P_1, P_2 orthogonale Projektionen auf abgeschlossene Unterräume M_1 bzw. $M_2 \subset H$.*

a) P_1P_2 ist genau dann orthogonale Projektion, falls $P_1P_2 = P_2P_1$ gilt. In diesem Fall ist P_1P_2 orthogonale Projektion auf den Unterraum $M_1 \cap M_2$.

b) Es sind äquivalent:

(i) $M_1 \subset M_2$.

(ii) *Es gilt* $\|P_1x\| \le \|P_2x\| \quad (x \in H)$.

(iii) *Es gilt* $P_1 \leqslant P_2$ *im Sinne von* $\langle P_1x, x\rangle \le \langle P_2x, x\rangle \quad (x \in H)$.

(iv) *Es gilt* $P_1P_2 = P_2P_1 = P_1$.

Beweis. a) (i). Es gelte $P_1P_2 = P_2P_1$. Dann erhalten wir

$$(P_1P_2)^2 = P_1P_2P_1P_2 = P_1^2P_2^2 = P_1P_2$$

und

$$(P_1P_2)^* = (P_2P_1)^* = P_1^*P_2^* = P_1P_2.$$

Also ist P_1P_2 eine orthogonale Projektion.

(ii). Sei P_1P_2 orthogonale Projektion. Dann gilt für $x, y \in H$

$$\langle x, P_2P_1y\rangle = \langle x, P_2^*P_1^*y\rangle = \langle P_1P_2x, y\rangle = \langle x, P_1P_2y\rangle .$$

Daher ist $P_1P_2 = P_2P_1$.

In diesem Fall gilt $R(P_2P_1) \subset R(P_2) = M_2$ und $R(P_2P_1) = R(P_1P_2) \subset M_1$. Zu $x \in M_1 \cap M_2$ ist $x = P_1x = P_2x$, d.h. $(P_2P_1)x = x$. Insgesamt erhalten wir $R(P_2P_1) = M_1 \cap M_2$.

Der Beweis von Teil b) sei als Übungsaufgabe überlassen. □

8.7 Lemma. *Sei* $(P_n)_{n\in\mathbb{N}} \subset L(H)$ *eine Folge orthogonaler Projektionen in einem Hilbertraum* H *mit* $P_m \leqslant P_n$ *für* $m \le n$*. Dann konvergiert* P_n *stark gegen eine orthogonale Projektion* $P \in L(H)$.

Beweis. Für $x \in H$ ist $(\|P_nx\|)_{n\in\mathbb{N}}$ beschränkt, monoton steigend (Lemma 8.6 b)), also konvergent. Für $m \le n$ ist

$$\begin{aligned}\|P_nx - P_mx\|^2 &= \underbrace{\langle P_nx, P_nx\rangle}_{=\|P_nx\|^2} - \underbrace{\langle P_nx, P_mx\rangle}_{=\langle P_mP_nx, x\rangle = \langle P_mx, x\rangle = \|P_mx\|^2} - \langle P_mx, P_nx\rangle + \underbrace{\langle P_mx, P_mx\rangle}_{=\|P_mx\|^2}\\ &= \|P_nx\|^2 + \|P_mx\|^2 - 2\|P_mx\|^2 \longrightarrow 0 \quad (m, n \longrightarrow \infty).\end{aligned}$$

Also existiert $\lim_{n\to\infty} P_nx \quad (x \in H)$, und der Grenzwert hängt linear von x ab; wir können ihn also Px nennen, mit $P \in L(H)$.

Es gilt $\langle Px, y\rangle = \lim_{n\to\infty} \langle P_nx, y\rangle = \lim_{n\to\infty} \langle x, P_ny\rangle = \langle x, Py\rangle$ und

$$\langle P^2x, y\rangle = \langle Px, Py\rangle = \lim_{n\to\infty} \langle P_nx, P_ny\rangle = \lim_{n\to\infty} \langle P_nx, y\rangle = \langle Px, y\rangle .$$

Somit gilt $P^2 = P = P^*$ und $P_n \xrightarrow{s} P$. □

b) PV-Maße und der Spektralsatz

Wir kommen noch einmal auf den Matrizenfall zurück. In der Situation von Beispiel 8.4 war

$$Tx = \sum_{\lambda \in \sigma(T)} \lambda E(\{\lambda\})x \quad (x \in \mathbb{C}^n).$$

Liest man hier $E\colon \sigma(T) \to L(\mathbb{C}^n)$ als eine matrizenwertige σ-additive Abbildung, also als ein matrizenwertiges Maß, kann man die rechte Seite als ein Integral interpretieren, und man erhält

$$T = \int_{\sigma(T)} \lambda dE(\lambda).$$

Man beachte dabei, dass $\sigma(T) = \{\lambda_1, \dots, \lambda_m\}$ die disjunkte Vereinigung der einelementigen Mengen $\{\lambda_1\}, \dots, \{\lambda_m\}$ ist und dass $\mathrm{id}\colon \sigma(T) \to \mathbb{C}$ eine Funktion mit endlichem Wertebereich, also eine Stufenfunktion ist.

Dies soll im Folgenden formalisiert und auf den Fall unendlicher Spektren $\sigma(T)$ verallgemeinert werden. Dabei sei stets $(X, \mathscr{A})$ ein Messraum.

8.8 Definition. Eine Abbildung $E\colon \mathscr{A} \to L(H)$ heißt ein projektorwertiges Maß (PV-Maß), falls gilt:

(i) $E(A)$ ist orthogonale Projektion $(A \in \mathscr{A})$.

(ii) Sei $(A_n)_{n\in\mathbb{N}} \subset \mathscr{A}$ eine Familie paarweise disjunkter Mengen. Dann gilt

$$\Big[E\Big(\bigcup_{n\in\mathbb{N}} A_n\Big)\Big]x = \sum_{n\in\mathbb{N}} E(A_n)x \quad (x \in H).$$

(iii) Es gilt $E(X) = \mathrm{id}_H$.

Eine Menge $A \in \mathscr{A}$ heißt eine E-Nullmenge, falls $E(A) = 0$ (dabei ist die 0 auf der rechten Seite der Nulloperator in H).

Falls X topologischer Raum ist und $\mathscr{A} = \mathscr{B}(X)$ die Borel–σ–Algebra, so besitzt das PV-Maß kompakten Träger, falls eine kompakte Menge $K \in \mathscr{B}(X)$ existiert mit $E(K) = \mathrm{id}_H$.

8.9 Bemerkung. Sei E ein PV-Maß. Durch direktes Nachrechnen kann man die folgenden Aussagen beweisen.

a) $E(\emptyset) = 0_{L(H)}$.

b) $E(B \setminus A) = E(B) - E(A)$ für $A, B \in \mathscr{A}$ mit $A \subset B$.

c) $E(A \cup B) + E(A \cap B) = E(A) + E(B) \quad (A, B \in \mathscr{A})$.

d) Seien $A_n \in \mathscr{A}$ mit $A_n \subset A_{n+1}$ $(n \in \mathbb{N})$. Dann ist $E(\bigcup_n A_n) = \text{s-lim}_n E(A_n)$, also mit Konvergenz in der starken Operatortopologie auf der rechten Seite.

Analog gilt für $A_n \in \mathscr{A}$ mit $A_n \supset A_{n+1}$ $(n \in \mathbb{N})$ die Gleichheit $E(\bigcap_{n\in\mathbb{N}} A_n) = \text{s-lim}_{n\to\infty} E(A_n)$.

e) $E(A \cap B) = E(A)E(B) = E(B)E(A) \quad (A, B \in \mathscr{A})$.

f) Seien $A, B \in \mathscr{A}$ mit $A \cap B = \emptyset$. Dann ist $R(E(A)) \perp R(E(B))$.

g) Sei $x \in H$. Dann ist $E_x \colon \mathscr{A} \to [0, \infty)$ mit

$$E_x(A) := \langle E(A)x, x\rangle = \|E(A)x\|^2$$

ein endliches Maß mit $\|E_x\|_M = E_x(X) = \|x\|^2$.

h) Seien $x, y \in H$. Dann ist $E_{x,y} \colon \mathscr{A} \to \mathbb{C}$ mit

$$E_{x,y}(A) := \langle E(A)x, y\rangle$$

ein komplexes Maß mit $\|E_{x,y}\|_M \le \|x\| \cdot \|y\|$. Dabei bezeichnet $\|\cdot\|_M$ die Totalvariation.

8.10 Definition. Sei E ein PV-Maß. Sei $f \colon X \to \mathbb{C}$ eine Stufenfunktion, d.h. es existiert eine Darstellung der Form $f = \sum_{i=1}^n f_i \chi_{A_i}$ mit $f_i \in \mathbb{C}$ und $A_i \in \mathscr{A}$ disjunkt. Dann heißt

$$\int f \, \mathrm{d}E := \sum_{i=1}^n f_i E(A_i) \in L(H)$$

das Integral von f bzgl. E.

Im Folgenden schreiben wir etwas präziser für eine Funktion $f \colon X \to \mathbb{C}$ statt $\|f\|_\infty$ auch

$$\|f\|_{\mathscr{L}^\infty(X)} := \sup_{\lambda \in X} |f(\lambda)|.$$

8.11 Lemma. *Sei E ein PV-Maß und seien f, g Stufenfunktionen.*

a) Die Abbildung $f \mapsto \int f \, \mathrm{d}E$ (vom Vektorraum der Stufenfunktionen nach $L(H)$) ist linear.

b) Für $x \in H$ gilt $\big\|(\int f \, \mathrm{d}E)x\big\|^2 = \int |f|^2 \, \mathrm{d}E_x \le \|f\|^2_{\mathscr{L}^\infty(X)} \|x\|^2$.

c) Es gilt $(\int f \, \mathrm{d}E) \circ (\int g \, \mathrm{d}E) = \int fg \, \mathrm{d}E$.

d) Es gilt $(\int f \, \mathrm{d}E)^* = \int \overline{f} \, \mathrm{d}E$.

Beweis. a) ist klar.

b) Unter Verwendung des Satzes von Pythagoras erhalten wir

$$\begin{aligned}\Big\|\Big(\int f\,\mathrm{d}E\Big)x\Big\|^2 &= \Big\|\sum_{i=1}^n f_i E(A_i)x\Big\|^2\\ &= \sum_{i=1}^n |f_i|^2\,\|E(A_i)x\|^2 = \int |f|^2\,\mathrm{d}E_x\\ &\le \|f\|^2_{\mathscr{L}^\infty(X)}\,\|E_x\|_M = \|f\|^2_{\mathscr{L}^\infty(X)}\,\|x\|^2 .\end{aligned}$$

c) Mit Bemerkung 8.9 gilt

$$\begin{aligned}\Big(\int f\,\mathrm{d}E\Big)\Big(\int g\,\mathrm{d}E\Big) &= \Big(\sum_{i=1}^n f_i E(A_i)\Big)\Big(\sum_{j=1}^m g_j E(B_j)\Big)\\ &= \sum_{i,j} f_i g_j E(A_i)E(B_j) = \sum_{i,j} f_i g_j E(A_i\cap B_j)\\ &= \int fg\,\mathrm{d}E.\end{aligned}$$

d) folgt direkt aus der Definition des Integrals. □

8.12 Definition. Sei E ein PV-Maß auf $(X,\mathscr{A})$ mit Werten in $L(H)$. Dann wird die Menge aller beschränkten messbaren Funktionen $f\colon X\to\mathbb{C}$ mit $B(X)$ bezeichnet. Für $f\in B(X)$ sei $(f_n)_{n\in\mathbb{N}}\subset B(X)$ eine Folge von Stufenfunktionen, welche gleichmäßig gegen f konvergiert. Definiere das Integral

$$\int f\,\mathrm{d}E := \int f(\lambda)\,\mathrm{d}E(\lambda) := \lim_{n\to\infty}\int f_n\,\mathrm{d}E \in L(H),$$

mit Konvergenz in der starken Operatortopologie. Für $A\in\mathscr{A}$ setzt man $\int_A f\,\mathrm{d}E := \int \chi_A f\,\mathrm{d}E$.

8.13 Bemerkung. a) Man beachte, dass das Integral wegen Lemma 8.11 b) wohldefiniert ist.

b) Die Eigenschaften von Lemma 8.11 übertragen sich in üblicher Weise auf messbare beschränkte Funktionen.

c) Falls $K\in\mathscr{A}$ ist mit $E(K)=\mathrm{id}_H$, so ist $\int f\,\mathrm{d}E = \int_K f\,\mathrm{d}E$ für alle $f\in B(X)$ (denn $E(X\setminus K)=0$).

8.14 Satz. *Sei $X\subset\mathbb{R}$ kompakt, und sei $E\colon\mathscr{B}(X)\to L(H)$ ein PV-Maß.*

a) Durch $T:=\int\lambda\,\mathrm{d}E(\lambda)$ wird ein selbstadjungierter Operator $T\in L(H)$ definiert.

b) Zu $f \in B(X)$ definiert man den Operator $f(T) := \int f(\lambda) dE(\lambda) \in L(H)$. Dann gilt:

(i) $f(T)$ ist ein normaler Operator für alle $f \in B(X)$ mit $\|f(T)\|_{L(H)} \leq \|f\|_{\mathscr{L}^\infty(X)}$.

(ii) Für $f, g \in B(X)$ ist $f(T)g(T) = (fg)(T) = g(T)f(T)$ und $\overline{f}(T) = (f(T))^*$.

Beweis. a) Da $X \subset \mathbb{R}$ kompakt ist, gilt $\mathrm{id}_X \in B(X)$, und damit ist $T := \int \lambda dE(\lambda) \in L(H)$ wohldefiniert. Wegen $\lambda = \overline{\lambda}$ auf X gilt nach Lemma 8.11 auch $T^* = \int \overline{\lambda} dE(\lambda) = T$, d.h. T ist selbstadjungiert.

b) Dies folgt direkt aus Lemma 8.11 (für beschränkte messbare Funktionen). Man beachte, dass $f(T)f(T)^* = f(T)\overline{f}(T) = (f\overline{f})(T) = |f|^2(T) = f(T)^* f(T)$ gilt und daher $f(T)$ normal ist. □

Der folgende Satz ist einer der wichtigsten Sätze der Funktionalanalysis und besagt, dass zu jedem Operator auch ein PV-Maß gefunden werden kann, so dass die obige Darstellung gilt.

8.15 Satz (Spektralsatz für beschränkte selbstadjungierte Operatoren)**.** *Sei $T \in L(H)$ selbstadjungiert. Dann existiert ein eindeutig bestimmtes PV-Maß*

$$E \colon \mathscr{B}(\sigma(T)) \to L(H)$$

mit

$$T = \int_{\sigma(T)} \lambda \, \mathrm{d}E(\lambda).$$

Für alle $f \in B(\sigma(T))$ wird durch $f(T) := \int_{\sigma(T)} f(\lambda) dE(\lambda)$ ein normaler Operator $f(T) \in L(H)$ definiert, und es gelten folgende Eigenschaften (Funktionalkalkül):

(i) Für $f \in B(\sigma(T))$ ist $\|f(T)\|_{L(H)} \leq \|f\|_{\mathscr{L}^\infty(\sigma(T))}$.

(ii) Für $f, g \in B(\sigma(T))$ ist $f(T)g(T) = (fg)(T) = g(T)f(T)$ und $\overline{f}(T) = (f(T))^*$.

(iii) Sei $(f_n)_{n\in\mathbb{N}} \subset B(\sigma(T))$ eine Folge und $f \in B(\sigma(T))$ mit $f_n \to f$ punktweise und $\|f_n\|_{\mathscr{L}^\infty(\sigma(T))} \leq C < \infty$ $(n \in \mathbb{N})$. Dann gilt $f_n(T)x \to f(T)x$ $(n \to \infty)$ für alle $x \in H$.

Eine Beweisidee für diesen Satz wird etwas später diskutiert.

8.16 Bemerkung. Man kann in Satz 8.15 das PV-Maß auf ganz $\mathbb{R}$ fortsetzen, indem man

$$\widetilde{E}(A) := E(A \cap \sigma(T)) \quad (A \in \mathscr{B}(\mathbb{R}))$$

definiert. Schreibt man wieder E statt $\widetilde{E}$, so kann man den Spektralsatz auch in der Form

$$T = \int_{\mathbb{R}} \lambda dE(\lambda)$$

schreiben. Aus dieser Definition folgt sofort: Sei $U \subset \rho(T) \cap \mathbb{R}$ eine offene Teilmenge von $\mathbb{R}$. Dann gilt $E(U) = 0$.

Da $\rho(T)$ offen ist, folgt für alle $\lambda_0 \in \rho(T)$: Es existiert ein $\varepsilon > 0$ mit $E((\lambda_0 - \varepsilon, \lambda_0 + \varepsilon)) = 0$.

c) Folgerungen aus dem Spektralsatz, Funktionalkalkül

8.17 Bemerkung. Sei $T \in L(H)$ ein selbstadjungierter Operator und sei $E\colon \mathscr{B}(\sigma(T)) \to L(H)$ die zugehörige Spektralschar. Dann gilt:

(i) $E(\sigma(T)) = \mathrm{id}_H$,

(ii) für alle $A \in \mathscr{B}(\sigma(T))$ ist $E(A) = \chi_A(T)$,

(iii) für alle Polynome $p\colon \mathbb{R} \to \mathbb{C}$, $\lambda \mapsto \sum_{n=0}^N c_n \lambda^n$ stimmt $p(T) = \int p(\lambda) dE(\lambda)$ mit der üblichen Definition $\sum_{n=0}^N c_n T^n$ überein.

Diese Eigenschaften folgen direkt aus der Eigenschaft des Funktionalkalküls, z.B. gilt $\chi_A(T) = \int_{\sigma(T)} \chi_A(\lambda) dE(\lambda) = E(A \cap \sigma(T)) = E(A)$. Teil (iii) folgt aus $T = \int_{\sigma(T)} \lambda dE(\lambda)$ und der Eigenschaft (ii) aus Satz 8.14.

8.18 Bemerkung. Sei $T \in L(H)$ selbstadjungiert und $E\colon \sigma(T) \to L(H)$ das zugehörige PV-Maß. Dann gilt für alle $x, y \in H$ und alle $f \in B(\sigma(T))$

$$\Big\langle \Big[\int_{\sigma(T)} f(\lambda) dE(\lambda) \Big] x, y \Big\rangle = \int_{\sigma(T)} f(\lambda) d\langle E(\lambda)x, y\rangle,$$

wobei auf der rechten Seite bezüglich des komplexen Maßes $A \mapsto \langle E(A)x, y\rangle$ integriert wird. Speziell gilt für $x = y$

$$\Big\langle \Big[\int_{\sigma(T)} f(\lambda) dE(\lambda) \Big] x, x \Big\rangle = \int_{\sigma(T)} f(\lambda) d\|E(\lambda)x\|^2$$

mit dem Maß $\|E(\cdot)x\|^2\colon \mathscr{B}(\sigma(T)) \to [0, \infty)$, $A \mapsto \|E(A)x\|^2$. Denn diese Gleichheit gilt wegen der Linearität für Stufenfunktionen, und aufgrund der Definition des Integrals als starker Limes auch für beschränkte messbare Funktionen.

8.19 Lemma. *Sei $T \in L(H)$ selbstadjungiert, und sei $x \in H$ ein Eigenvektor zum Eigenwert λ_0. Dann gilt $f(T)x = f(\lambda_0)x$ für alle $f \in B(\sigma(T))$.*

Beweis. Dies gilt offensichtlich, falls f ein Polynom ist. Da die Menge der Polynome nach dem Satz von Weierstraß dicht in $C(\sigma(T))$ bzgl. $\|\cdot\|_{\mathscr{L}^\infty(\sigma(T))}$ liegen, existiert zu jedem $f \in C(\sigma(T))$ eine Folge $(p_n)_{n\in\mathbb{N}}$ von Polynomen mit $\|f - p_n\|_{\mathscr{L}^\infty(\sigma(T))} \to 0$ $(n \to \infty)$. Damit folgt

$$\begin{aligned}\|f(T)x - f(\lambda_0)x\| &\le \|f(T) - p_n(T)\|_{L(H)}\|x\| + |p_n(\lambda_0) - f(\lambda_0)|\,\|x\| \\ &\le 2\|f - p_n\|_{\mathscr{L}^\infty(\sigma(T))}\|x\| \to 0 \quad (n \to \infty)\end{aligned}$$

für alle $f \in C(\sigma(T))$. Damit gilt die Aussage auch für stetige Funktionen.

Sei nun $f = \chi_A$ mit $A \in \mathscr{B}(\sigma(T))$. Dann existiert eine Folge $(f_n)_{n\in\mathbb{N}} \subset C(\sigma(T))$ mit $\|f_n\|_{\mathscr{L}^\infty(\sigma(T))} \le C < \infty$ und $f_n \to f$ punktweise. Nach Eigenschaft (iii) im Spektralsatz 8.15 folgt $f(T)x = f(\lambda_0)x$ für alle Stufenfunktionen f. Da jede beschränkte messbare Funktion gleichmäßiger Grenzwert von Stufenfunktionen ist, folgt die Aussage für alle $f \in B(\sigma(T))$. □

8.20 Satz (Spektrum und Spektralmaß)**.** *Sei T ein beschränkter selbstadjungierter Operator, sei $E\colon \mathscr{B}(\mathbb{R}) \to L(H)$ das zugehörige PV-Maß, und sei $\lambda_0 \in \mathbb{R}$.*

a) Es gilt $\lambda_0 \in \rho(T)$ genau dann, falls eine offene Umgebung $U \subset \mathbb{R}$ von λ_0 existiert mit $E(U) = 0_{L(H)}$.

b) Es ist $\lambda_0 \in \sigma_p(T)$ genau dann, wenn $E(\{\lambda_0\}) \ne 0_{L(H)}$. Für alle $\lambda_0 \in \mathbb{R}$ gilt $R(E(\{\lambda_0\})) = \ker(T - \lambda_0)$.

c) Es ist $\lambda_0 \in \sigma_c(T)$ genau dann, wenn $E(\{\lambda_0\}) = 0_{L(H)}$ und $E((\lambda_0 - \varepsilon, \lambda_0 + \varepsilon)) \ne 0_{L(H)}$ für alle $\varepsilon > 0$ gilt.

Beweis. a) Nach Konstruktion (siehe Bemerkung 8.16 gilt $E(\rho(T)) = 0$. Falls $\lambda_0 \in \rho(T)$, existiert eine offene Umgebung $U \subset \rho(T)$ mit $\lambda_0 \in U$, und es folgt $E(U) = 0$.

Sei andererseits $\lambda_0 \in \mathbb{R}$ und $U = (\lambda_0 - \varepsilon, \lambda_0 + \varepsilon)$ mit $E(U) = 0$. Definiere $f, g \in B(\sigma(T))$ durch $f(\lambda) := \frac{1}{\lambda-\lambda_0} \cdot \chi_{\sigma(T)\setminus U}$ und $g(\lambda) := (\lambda - \lambda_0) \cdot \chi_{\sigma(T)}$. Dann gilt

$$\begin{aligned}f(T)(T - \lambda_0) &= f(T)g(T) = (f \cdot g)(T) = \chi_{\sigma(T)\setminus U}(T) \\ &= E(\sigma(T) \setminus U) = E(\sigma(T)) = \mathrm{id}_H\,.\end{aligned}$$

Wegen $f(T)g(T) = g(T)f(T)$ folgt $g(T)f(T) = \mathrm{id}_H$, d.h. $f(T) = g(T)^{-1} = (T - \lambda_0)^{-1}$. Somit ist $\lambda_0 \in \rho(T)$.

b) Wir zeigen $R(E(\{\lambda_0\})) = \ker(T - \lambda_0)$ für alle $\lambda_0 \in \mathbb{R}$. Falls $\lambda_0 \in \rho(T)$, so sind beide Seiten gleich $\{0\}$, also können wir $\lambda_0 \in \sigma(T)$ annehmen.

Sei $x \in R(E(\{\lambda_0\}))$. Dann gilt $x = E(\{\lambda_0\})x = \chi_{\{\lambda_0\}}(T)x$, und wir erhalten

$$Tx = T\chi_{\{\lambda_0\}}(T)x = \int_{\sigma(T)} \lambda\chi_{\lambda_0}(\lambda)dE(\lambda)x = \int_{\{\lambda_0\}} \lambda dE(\lambda)x$$

$$= \lambda_0 E(\lambda_0)x = \lambda_0 x.$$

Also ist $x \in \ker(T - \lambda_0)$.

Sei andererseits $x \in \ker(T - \lambda_0)$. Dann gilt nach Lemma 8.19 für $f = \chi_{\{\lambda_0\}}$ die Gleichheit $f(T)x = f(\lambda_0)x = x$. Wir erhalten

$$x = f(T)x = \int_{\sigma(T)} f(\lambda) dE(\lambda)x = \int_{\{\lambda_0\}} 1 dE(\lambda)x = E(\{\lambda_0\})x$$

und damit $x \in R(E(\{\lambda_0\}))$.

c) folgt aus a) und b) wegen $\mathbb{R} = (\mathbb{R} \cap \rho(T)) \,\dot\cup\, \sigma_p(T) \,\dot\cup\, \sigma_c(T)$. □

8.21 Lemma. *Sei $T \in L(H)$ selbstadjungiert mit $T \geq 0$, d.h. $\langle Tx, x\rangle \geq 0$ $(x \in H)$. Für $S := \sqrt{T} \in L(H)$ gilt dann $S \geq 0$ und $S^2 = T$. Insbesondere existiert für jedes $T \in L(H)$ der Absolutbetrag $|T| := \sqrt{T^*T}$.*

Beweis. Aus $T \geq 0$ folgt mit Lemma 7.8 die Inklusion $\sigma(T) \subset [0, \infty)$, und damit ist $\lambda \mapsto \sqrt{\lambda} \in B(\sigma(T))$, und $S := \sqrt{T} \in L(H)$ ist wohldefiniert. Für alle $x \in H$ gilt mit Bemerkung 8.18

$$\langle Sx, x\rangle = \int_{\sigma(T)} \sqrt{\lambda} d\|E(\lambda)x\|^2 \geq 0.$$

Die Existenz von $|T|$ für beliebiges $T \in L(H)$ folgt aus der Selbstadjungiertheit von T^*T. □

8.22 Satz. *Sei $T \in L(H)$ selbstadjungiert, und sei $u_0 \in H$. Für $t \in [0, \infty)$ sei $u(t)$ definiert durch*

$$u(t) := e^{tT} u_0 = \int_{\sigma(T)} e^{t\lambda} dE(\lambda) u_0.$$

Dann gilt $u \in C^1([0, \infty), H)$, und u ist eine Lösung des Cauchyproblems (abstrakten Anfangswertproblems)

$$\begin{aligned} u'(t) &= Tu(t) \quad (t \in [0, \infty)), \\ u(0) &= u_0. \end{aligned}$$

Beweis. Da $\lambda \mapsto e^{t\lambda} \in B(\sigma(T))$, ist $u(t)$ für alle $t \in [0, \infty)$ wohldefiniert. Für $t = 0$ gilt

$$u(0) = \int_{\sigma(T)} 1 dE(\lambda) u_0 = \mathrm{id}_H\, u_0 = u_0.$$

Sei $t \geq 0$, und sei $h \in \mathbb{R}$ mit $|h| \leq 1$ und $t + h \geq 0$. Für $\lambda \in \sigma(T)$ definiere $f_h(\lambda) := \frac{1}{h}(e^{(t+h)\lambda} - e^{t\lambda})$ sowie $f(\lambda) := \lambda e^{t\lambda}$. Dann gilt $\sup_{|h|\leq 1} \|f_h\|_{\mathscr{L}^\infty(\sigma(T))} < \infty$ sowie $f_h \to f$ punktweise. Nach Eigenschaft (iii) in Satz 8.15 folgt $f_h(T)x \to f(T)x$ $(h \to 0)$ für alle $x \in H$. Mit dem Funktionalkalkül erhält man

$$\begin{aligned}\tfrac{1}{h}(u(t+h) - u(t)) &= \int_{\sigma(T)} f_h(\lambda) dE(\lambda) u_0 \to \int_{\sigma(T)} f(\lambda) dE(\lambda) u_0 \\ &= Te^{tT} u_0 = Tu(t) \quad (h \to 0).\end{aligned}$$

Also ist $u\colon [0,\infty) \to H$ differenzierbar, und es gilt $u'(t) = Tu(t)$ $(t \geq 0)$. Insbesondere ist $u \in C([0,\infty), H)$, und wegen $u'(t) = Tu(t)$ ist auch u' stetig, d.h. es gilt $u \in C^1([0,\infty), H)$. □

8.23 Satz (Spektralabbildungssatz)**.** *Sei $T \in L(H)$ selbstadjungiert, und sei $f\colon \sigma(T) \to \mathbb{C}$ stetig. Dann gilt*

$$\sigma(f(T)) = f(\sigma(T)) \Big(= \{f(\lambda)\colon \lambda \in \sigma(T)\} \Big).$$

Beweis. **(i)** Wir zeigen $\sigma(f(T)) \subset f(\sigma(T))$ für Polynome f: Sei f ein Polynom, und sei $\mu \in \sigma(f(T))$. Wir faktorisieren

$$f(t) - \mu = \beta_N \cdot \prod_{i=1}^N (t - \gamma_i), \qquad \beta_N \neq 0,$$

und erhalten $f(T) - \mu = \beta_N \cdot \prod_{i=1}^N (T - \gamma_i)$. Falls $\gamma_i \in \rho(T)$ für jedes i gälte, so wäre $f(T) - \mu$ bijektiv, d.h. also $\mu \in \rho(f(T))$, was nicht sein kann. Also existiert ein i_0, für das $T - \gamma_{i_0}$ nicht bijektiv ist. Das heißt aber $\gamma_{i_0} \in \sigma(T)$. Wegen $f(\gamma_{i_0}) - \mu = 0$ folgt $\mu \in f(\sigma(T))$.

(ii) Wir zeigen $f(\sigma(T)) \subset \sigma(f(T))$ für Polynome f: Sei nun $\mu \in f(\sigma(T))$, d.h. es ist $\mu = f(\gamma)$ mit einem $\gamma \in \sigma(T)$. Dann folgt $f(\gamma) - \mu = 0$, d.h.

$$f(t) - \mu = (t - \gamma)\widetilde{f}(t)$$

mit einem Polynom $\widetilde{f}$ von Grad gleich $N - 1$. Also gilt

$$f(T) - \mu = (T - \gamma)\widetilde{f}(T) = \widetilde{f}(T)(T - \gamma).$$

Da $\gamma \in \sigma(T)$, ist $T - \gamma$ nicht surjektiv und damit auch $f(T) - \mu$ nicht surjektiv, oder es ist $T - \gamma$ nicht injektiv und damit $f(T) - \mu$ nicht injektiv. In beiden Fällen folgt $\mu \in \sigma(f(T))$.

(iii) Sei nun $f \in C(\sigma(T))$.

Falls $\mu \notin f(\sigma(T))$, dann ist $g := \frac{1}{f-\mu} \in C(\sigma(T))$, und somit

$$g(T)(f(T)-\mu) = (g\cdot(f-\mu))(T) = \chi_{\sigma(T)}(T) = \mathrm{id}_H\,.$$

Genauso folgt $(f(T)-\mu)g(T) = \mathrm{id}_H$. Also ist $\mu \in \rho(f(T))$.

Sei andererseits $\mu \in f(\sigma(T))$, d.h. es gibt ein $\lambda \in \sigma(T)$ mit $\mu = f(\lambda)$. Wähle Polynome $g_n \in P(\sigma(T))$ mit $\|f-g_n\|_{\mathscr{L}^\infty(\sigma(T))} \le 1/n$ (und damit $\|f(T)-g_n(T)\|_{L(H)} \le 1/n$). Nach (i) und (ii) ist $g_n(\lambda) \in \sigma(g_n(T))$. Es ist $g_n(T)$ ein normaler Operator, der also kein Restspektrum hat. Dann ist $\sigma(g_n(T)) = \sigma_p(g_n(T)) \cup \sigma_c(g_n(T)) = \sigma_{\mathrm{app}}(g_n(T))$, d.h. es existiert eine Folge $(x_{n,m})_{m\in\mathbb{N}} \subset H$ mit $\|x_{n,m}\| = 1$ und

$$\|(g_n(T)-g_n(\lambda))x_{n,m}\| \le \tfrac{1}{m}.$$

Somit ist

$$\begin{aligned}\|(f(T)-f(\lambda))x_{n,n}\| \le \|(f(T)-g_n(T))x_{n,n}\| + \|(g_n(T)-g_n(\lambda))x_{n,n}\|\\ + |g_n(\lambda)-f(\lambda)|\cdot\|x_{n,n}\| \le \tfrac{3}{n},\end{aligned}$$

d.h. $f(\lambda) \in \sigma_{\mathrm{app}}(f(T)) \subset \sigma(f(T))$. □

8.24 Korollar. *Seien $T \in L(H)$ selbstadjungiert und $f \in C(\sigma(T))$. Dann gilt*

$$\|f(T)\|_{L(H)} = \|f\|_{\mathscr{L}^\infty(\sigma(T))}.$$

Speziell gilt für $\lambda_0 \in \rho(T)$ die Gleichheit

$$\|(T-\lambda_0)^{-1}\|_{L(H)} = \frac{1}{\mathrm{dist}(\lambda_0,\sigma(T))}.$$

Beweis. Für $f \in C(\sigma(T))$ ist

$$\begin{aligned}\|f(T)\|_{L(H)}^2 &= \sup_{\|x\|=1}\|f(T)x\|^2 = \sup_{\|x\|=1}\langle f(T)x, f(T)x\rangle = \sup_{\|x\|=1}\langle f(T)^*f(T)x,x\rangle\\ &= \sup_{\|x\|=1}\langle |f|^2(T)x,x\rangle = \sup\{|\lambda| : \lambda\in\sigma(|f|^2(T))\}\\ &= \sup\{|f|^2(\lambda) : \lambda \in \sigma(T)\} = \|f\|^2_{\mathscr{L}^\infty(\sigma(T))}.\end{aligned}$$

Dabei wurden Satz 7.12 für den selbstadjungierten Operator $f(T)^*f(T) = |f|^2(T)$ verwendet sowie der Spektralabbildungssatz 8.23.

Der Zusatz folgt aus $\|f\|_{\mathscr{L}^\infty(\sigma(T))} = \frac{1}{\mathrm{dist}(\lambda_0,\sigma(T))}$ für die Funktion $(f\colon \lambda \mapsto \frac{1}{\lambda-\lambda_0}) \in C(\sigma(T))$. □

d) Zum Beweis des Spektralsatzes und allgemeine Formulierung

Eine Möglichkeit, den Spektralsatz zu beweisen, liegt in der Konstruktion des Funktionalkalküls: Nach Bemerkung 8.17 (ii) gilt $E(A) = \chi_A(T)$ für alle $A \in \mathscr{B}(\sigma(T))$. Dies kann als Definition von $E(A)$ verwendet werden, falls man $\chi_A(T)$ geeignet definieren kann. Dazu benötigt man messbare Funktionen des Operators T, also einen sogenannten messbaren Funktionalkalkül. Dieser wird schrittweise aufgebaut:

(i) Für alle Polynome $p\colon \mathbb{R} \to \mathbb{C}$, $p(\lambda) = \sum_{k=0}^n a_k\lambda^k$, ist $p(T) := \sum_{k=0}^n a_k T^k$ in natürlicher Weise definiert. Man rechnet unter Verwendung des Spektralabbildungssatzes (siehe Beweis (i),(ii) von Satz 8.23) wie in Korollar 8.24 nach, dass

$$\|p(T)\|_{L(H)} = \|p\|_{\mathscr{L}^\infty(\sigma(T))} \tag{8-2}$$

gilt. Offensichtlich gilt für zwei Polynome p_1, p_2 auch $(p_1p_2)(T) = p_1(T)p_2(T)$ sowie $\overline{p_1}(T) = (p_1(T))^*$, d.h. die Eigenschaften eines Funktionalkalküls gelten.

(ii) Für stetige Funktionen $f \in C(\sigma(T))$ definiert man $f(T)$ durch Approximation durch Polynome. Sei $(p_n)_{n\in\mathbb{N}}$ eine Folge von Polynomen mit $\|p_n - f\|_{\mathscr{L}^\infty(\sigma(T))} \to 0$. Dann ist $(p_n)_{n\in\mathbb{N}} \subset C(\sigma(T))$ eine Cauchyfolge, und wegen (8-2) ist $(p_n(T))_{n\in\mathbb{N}} \subset L(H)$ auch eine Cauchyfolge und damit konvergent. Man setzt $f(T) := \lim_{n\to\infty} p_n(T)$. Die Eigenschaften eines Funktionalkalküls übertragen sich sofort auf stetige Funktionen.

(iii) Für die Erweiterung von stetigen Funktionen auf messbare Funktionen braucht man einige Vorbereitungen.

8.25 Definition. a) Sei $(X, \mathscr{A})$ ein Messraum. Eine Abbildung $\mu\colon \mathscr{A} \to \mathbb{C}$ heißt ein komplexes Maß, falls für jede Folge paarweise disjunkter Mengen $(A_n)_{n\in\mathbb{N}} \subset \mathscr{A}$ gilt:

$$\mu\Big(\bigcup_{n=1}^\infty A_n\Big) = \sum_{n=1}^\infty \mu(A_n),$$

wobei wir hier verlangen, dass die (unbedingt konvergente) Reihe für jede Folge von disjunkten A_n einen endlichen Wert liefert. Die Menge aller komplexen Maße wird mit $M(X, \mathscr{A})$ bezeichnet. Falls X ein topologischer Raum mit der von der Familie der in X offenen Mengen erzeugten Borel–σ-Algebra $\mathscr{B}(X)$ ist, so schreibt man $M(X) := M(X, \mathscr{B}(X))$.

b) Zu $\mu \in M(X, \mathscr{A})$ definiert man die Variation (das Variationsmaß) $|\mu|$ durch

$$|\mu|(A) := \sup_{Z\in\mathscr{Z}} \sum_{E\in Z} |\mu(E)|,$$

wobei das Supremum über die Menge $\mathscr{Z}$ aller Zerlegungen Z von A in endlich viele paarweise disjunkte Mengen aus $\mathscr{A}$ gebildet wird. Die Totalvariation oder Variationsnorm von μ ist definiert durch $\|\mu\|_M := |\mu|(X)$.

8.26 Bemerkung. Man kann zeigen, dass $|\mu|$ ein endliches (positives) Maß auf $\mathscr{A}$ ist. Es gilt ferner: $M(X, \mathscr{A})$, versehen mit der Variationsnorm, ist ein Banachraum. Beweise finden sich z.B. in [15].

8.27 Satz (Rieszscher Darstellungssatz)**.** *Sei X ein kompakter topologischer Raum. Dann ist die Abbildung*

$$T\colon M(X) \to C(X)', \quad \mu \mapsto T\mu \;\; \textit{mit } (T\mu)(f) := \int_X f \,\mathrm{d}\mu$$

ein isometrischer Isomorphismus von Banachräumen.

Dieser Satz wird hier nicht bewiesen, siehe z.B. [15].

Mit diesen Sätzen kann man den messbaren Funktionalkalkül definieren. Seien $T \in L(H)$ ein selbstadjungierter Operator und $x, y \in H$. Dann definiert

$$\ell_{x,y}\colon C(\sigma(T)) \to \mathbb{C},\; f \mapsto \langle f(T)x, y\rangle$$

eine stetige Linearform. Dabei ist die Linearität klar, die Stetigkeit folgt aus

$$|\ell_{x,y}(f)| \le \|f(T)x\| \cdot \|y\| \le \|f\|_{L^\infty(\sigma(T))} \cdot \|x\| \cdot \|y\| \,.$$

Hier wurde der stetige Funktionalkalkül verwendet. Nach dem Rieszschen Darstellungssatz 8.27 existiert ein komplexes Maß $\mu_{x,y} \in M(\sigma(T))$ mit

$$\langle f(T)x, y\rangle = \int_{\sigma(T)} f \,\mathrm{d}\mu_{x,y} \quad (f \in C(\sigma(T))) \tag{8-3}$$

Ebenfalls nach Satz 8.27 gilt $\|\mu_{x,y}\|_M = \|\ell_{x,y}\|_{(C(\sigma(T)))'} \le \|x\| \cdot \|y\|$. Die rechte Seite von (8-3) ist aber nicht nur für stetige f, sondern auch für beschränkte messbare f definiert. Für $f \in B(\sigma(T))$ betrachte also die Abbildung

$$(x, y) \mapsto \int_{\sigma(T)} f \,\mathrm{d}\mu_{x,y}.$$

Diese Abbildung ist bilinear, und wegen

$$\Big| \int_{\sigma(T)} f \,\mathrm{d}\mu_{x,y} \Big| \le \|f\|_{L^\infty(\sigma(T))} \cdot \|\mu_{x,y}\|_M \le \|f\|_{L^\infty(\sigma(T))} \cdot \|x\| \cdot \|y\|$$

auch stetig. Damit existiert ein eindeutiger Operator $f(T) \in L(H)$, der diese stetige Bilinearform darstellt, d.h. es gilt

$$\langle f(T)x, y\rangle = \int_{\sigma(T)} f \,\mathrm{d}\mu_{x,y} \quad (x, y \in H).$$

Dies definiert $f(T)$ für beschränkte messbare Funktionen. Wieder lassen sich die Eigenschaften des Funktionalkalküls nachrechnen. Schließlich wird durch $E(A) := \chi_A(T)$ das zu T gehörige PV-Maß definiert. Die Eigenschaften des PV-Maßes aus Satz 8.15 lassen sich nun direkt zeigen, was den Beweis des Spektralsatzes beendet.

8.28 Bemerkung. Der Spektralsatz wird häufig auch mit Hilfe von Spektralscharen formuliert. Dabei heißt eine Familie $\{F_\lambda\}_{\lambda\in\mathbb{R}} \subset L(H)$ eine Spektralschar, falls gilt:

(i) F_λ ist eine orthogonale Projektion für alle $\lambda \in \mathbb{R}$.

(ii) $F_\mu F_\lambda = F_\lambda F_\mu = F_\mu$ für alle $\mu \le \lambda$.

(iii) $F_\mu x \to F_\lambda x$ für $\mu \searrow \lambda$ $(x \in H)$ (Rechtsstetigkeit).

(iv) $F_\lambda x \longrightarrow 0$ für $\lambda \longrightarrow -\infty$ $(x \in H)$.

(v) $F_\lambda x \longrightarrow x$ für $\lambda \longrightarrow +\infty$ $(x \in H)$.

Sei $T \in L(H)$ selbstadjungiert und E das zugehörige PV-Maß. Dann wird durch

$$F_\lambda := E((-\infty, \lambda]) \quad (\lambda \in \mathbb{R})$$

eine Spektralschar definiert. Definiert man das Integral über Spektralscharen geeignet (etwa im Sinne eines Riemann–Stieltjes–Integrals), so gilt

$$T = \int_{\mathbb{R}} \lambda \,\mathrm{d}E(\lambda) = \int_{-\infty}^{\infty} \lambda \,\mathrm{d}F_\lambda.$$

Die Spektralscharen entsprechen den Verteilungsfunktionen in der Wahrscheinlichkeitstheorie.

Der Spektralsatz wurde oben für beschränkte selbstadjungierte Operatoren formuliert. Tatsächlich gilt dieser Satz aber für allgemeinere Operatoren: Zum einen kann die Selbstadjungiertheit $T = T^*$ durch die Normalität $TT^* = T^*T$ des Operators T ersetzt werden, zum anderen kann T auch unbeschränkt sein. Man beachte, dass zwei unbeschränkte Operatoren genau dann gleich sind, wenn sie gleichen Definitionsbereich besitzen und auf diesem übereinstimmen. Der folgende Satz wird hier nicht bewiesen, der wesentliche Teil des Beweises wurde aber oben schon skizziert.

8.29 Definition und Satz (Integral über PV-Maß für messbare Funktionen)**.** *Sei $(X, \mathscr{A})$ ein Messraum, $E\colon \mathscr{A} \to L(H)$ ein PV-Maß. Zu $x \in H$ sei wieder $E_x\colon \mathscr{A} \to [0,\infty)$ durch $E_x(A) := \|E(A)x\|^2$ definiert. Sei $f\colon X \to \mathbb{C}$ messbar.*

Definiere

$$D\left(\int f \,\mathrm{d}E\right) := \left\{x \in H\colon \int |f|^2 \,\mathrm{d}E_x < \infty\right\}.$$

Dann existiert für alle $x \in D(\int f \,\mathrm{d}E)$ eine Folge von Stufenfunktionen $f_n\colon X \to \mathbb{C}$ mit $f_n \longrightarrow f$ punktweise und $\int |f_n - f|^2 \,\mathrm{d}E_x \longrightarrow 0 \quad (n \longrightarrow \infty)$, und der Operator

$$\int f \,\mathrm{d}E\colon H \supset D\left(\int f \,\mathrm{d}E\right) \to H,\ x \mapsto \left(\int f \,\mathrm{d}E\right) x := \lim_{n\to\infty} \left(\int f_n \,\mathrm{d}E\right) x$$

ist wohldefiniert.

8.30 Satz (Spektralsatz, allgemeine Formulierung)**.** *Sei H ein $\mathbb{C}$-Hilbertraum und $T\colon H \supset D(T) \to H$ ein normaler Operator in H. Dann existiert genau ein PV-Maß $E\colon \mathscr{B}(\sigma(T)) \to L(H)$ mit*

$$T = \int_{\sigma(T)} \mathrm{id}_{\sigma(T)}\, \mathrm{d}E.$$

Für jede messbare Funktion $f\colon \sigma(T) \to \mathbb{C}$ wird durch

$$\begin{aligned} f(T) &:= \int_{\sigma(T)} f \,\mathrm{d}E, \\ D(f(T)) &:= \left\{ x \in H\colon \int_{\sigma(T)} |f|^2 \,\mathrm{d}E_x < \infty \right\} \end{aligned}$$

ein normaler Operator definiert. Falls f ein Polynom ist, stimmt $f(T)$ mit der üblichen Definition überein. Weiter gilt für messbare Funktionen $f, g\colon \sigma(T) \to \mathbb{C}$:

$$\begin{aligned} \overline{f}(T) &= (f(T))^*, \\ (f+g)(T) &\supset f(T) + g(T), \\ (fg)(T) &\supset f(T)g(T). \end{aligned}$$

9. Schwache Konvergenz

9.1 Worum geht's? In diesem Abschnitt gehen wir noch einmal auf die schwache Konvergenz ein, welche für Anwendungen wichtig ist, da die Konvergenz in der Norm häufig nicht gegeben ist. Im Gegensatz zur Normtopologie liegt in der schwachen Topologie häufiger eine kompakte Situation vor. So ist etwa die Norm-Einheitskugel in einem reflexiven Raum immer schwach folgenkompakt. Dies kann man ausnützen, um Gleichungen zu lösen: Falls man eine Folge approximativer Lösungen hat und diese in der Norm beschränkt sind, so konvergiert eine Teilfolge zumindest schwach. Damit existiert ein Grenzwert, der in vielen Fällen die Gleichung exakt löst.

Im Folgenden sei X ein normierter Raum, und wie üblich sei $X' = L(X, \mathbb{K})$ der topologische Dualraum. Versehen mit der Operatornorm $\|x'\|_{X'} := \sup_{x\in X,\, \|x\|_X=1} |x'(x)|$ ist X' stets ein Banachraum. Wir werden in diesem Abschnitt die Schreibweise $x'x := x'(x)$ für Funktionale $x' \in X'$ und Elemente $x \in X$ verwenden. Mit $J_X \colon X \to X''$, $x \mapsto \widetilde{x}$, wobei $\widetilde{x}(x') := x'x$, bezeichnen wir die kanonische Einbettung $X \hookrightarrow X''$.

Wir wiederholen noch einmal verschiedene Konvergenzbegriffe.

9.2 Definition. a) Eine Folge $(x_n)_{n\in\mathbb{N}} \subset X$ heißt normkonvergent gegen $x \in X$, falls $\|x_n - x\|_X \to 0$ $(n \to \infty)$.

b) Eine Folge $(x_n)_{n\in\mathbb{N}} \subset X$ heißt schwach konvergent, falls für alle $x' \in X'$ gilt: $x'x_n \to x'x$ $(n \to \infty)$. In diesem Fall schreibt man $x_n \rightharpoonup x$.

c) Eine Folge $(x'_n)_{n\in\mathbb{N}} \subset X'$ heißt schwach-*-konvergent gegen $x' \in X'$, falls für alle $x \in X$ gilt: $x'_n x \to x'x$ $(n \to \infty)$. In diesem Fall schreibt man $x'_n \overset{*}{\rightharpoonup} x'$.

9.3 Beispiel. Seien X ein Hilbertraum und $(e_n)_{n\in\mathbb{N}}$ ein Orthonormalsystem in X. Dann gilt $e_n \rightharpoonup 0$. Denn für alle $x \in X$ gilt $\sum_{n\in\mathbb{N}} |\langle x, e_n\rangle|^2 \le \|x\|^2$ (Besselsche Ungleichung) und damit $\langle x, e_n\rangle \to 0$ $(n \to \infty)$. Wegen $\|e_n\| = 1$ gilt aber nicht $e_n \to 0$.

9.4 Bemerkung. a) Normkonvergenz impliziert schwache Konvergenz wegen $|x'x_n - x'x| \le \|x'\|_{X'} \|x_n - x\|_X$. Die Umkehrung gilt nicht, wie das obige Beispiel zeigt.

b) Da X' selbst Banachraum ist, ist der Begriff der schwachen Konvergenz auch in X' sinnvoll. Aus der schwachen Konvergenz in X' folgt die schwach-*-Konvergenz wegen

$$|x'_n x - x'x| = |(J_X x)x'_n - (J_X x)x'| \to 0 \quad (n \to \infty).$$

Falls X reflexiv ist, so sind schwache Konvergenz und schwach-*-Konvergenz auf X' identisch.

9.5 Lemma (Eindeutigkeit der Grenzwerte)**.** *a) Falls $x_n \rightharpoonup x$ und $x_n \rightharpoonup \widetilde{x}$, so gilt $x = \widetilde{x}$.*

b) Falls $x_n' \overset{*}{\rightharpoonup} x'$ und $x_n' \overset{*}{\rightharpoonup} \widetilde{x}'$, so gilt $x' = \widetilde{x}'$.

Beweis. a) Nach Voraussetzung gilt $x'(x-\widetilde{x}) = 0$ für alle $x' \in X'$. Nach dem Satz von Hahn-Banach folgt $x - \widetilde{x} = 0$.

b) Aus der Voraussetzung folgt $x'x = \widetilde{x}'x$ für alle $x \in X$ und damit $x' = \widetilde{x}'$. □

9.6 Satz. *Sei $M \subset X$. Dann ist M genau dann beschränkt, wenn für alle $x' \in X'$ gilt: $\sup_{x\in M} |x'x| < \infty$.*

Beweis. Falls M beschränkt ist, so gilt für alle $x' \in X'$ die Abschätzung $|x'x| \le \|x'\|_{X'}\|x\|_X \le C\|x'\|_{X'}$.

Sei $\mathcal{F} := J_X(M) = \{J_X x : x \in M\} \subset X''$. Zu $f \in \mathcal{F}$ existiert ein $x \in M$ mit $f = J_X x$. Wegen $|fx'| = |(J_X x)(x')| = |x'x|$ folgt $\sup_{f\in\mathcal{F}} |fx'| < \infty$. Nach dem Satz von Banach-Steinhaus existiert ein $k > 0$ mit $\|f\|_{X''} \le k$ $(f \in \mathcal{F})$. Wegen $\|J_X x\|_{X''} = \|x\|_X$ folgt daraus $\|x\|_X \le k$ $(x \in M)$, d.h. M ist beschränkt. □

9.7 Satz. *a) Seien $(x_n)_{n\in\mathbb{N}} \subset X$ und $x \in X$ mit $x_n \rightharpoonup x$ $(n \to \infty)$. Dann ist $(x_n)_{n\in\mathbb{N}}$ normbeschränkt, und es gilt $\|x\|_X \le \liminf_{n\to\infty} \|x_n\|_X$.*

b) Sei $(x_n')_{n\in\mathbb{N}} \subset X'$ und $x' \in X'$ mit $x_n' \overset{*}{\rightharpoonup} x'$. Dann gilt $\|x'\|_{X'} \le \liminf_{n\to\infty} \|x_n'\|_{X'}$. Falls X Banachraum ist, so ist $(x_n')_{n\in\mathbb{N}} \subset X'$ beschränkt.

Beweis. a) Für alle $x' \in X'$ gilt $\sup_{n\in\mathbb{N}} |x'x_n| < \infty$. Nach Satz 9.6 ist $M := \{x_n : n \in \mathbb{N}\} \subset X$ beschränkt. Weiter gilt

$$|x'x| = \lim_{n\to\infty} |x'x_n| \le \liminf_{n\to\infty} \|x'\|_{X'}\|x_n\|_X \quad (x' \in X').$$

Wegen $\|x\|_X = \sup_{x'\in X',\,\|x'\|=1} |x'x|$ folgt $\|x\| \le \liminf_{n\to\infty} \|x_n\|$.

b) Für alle $x \in X$ gilt

$$|x'x| = \lim_{n\to\infty} |x_n'x| \le \liminf_{n\to\infty} \|x_n'\|_{X'}\|x\|_X$$

und damit $\|x'\|_{X'} = \sup_{\|x\|=1} |x'x| \le \liminf_{n\to\infty} \|x_n'\|_{X'}$.

Sei nun X ein Banachraum. Aus der schwach-*-Konvergenz folgt $\sup_{n\in\mathbb{N}} |x_n'x| < \infty$ für alle $x \in X$, und mit dem Satz von Banach-Steinhaus erhalten wir $\|x_n'\|_{X'} \le C < \infty$ $(n \in \mathbb{N})$. □

9.8 Lemma. *a) Falls $x_n \rightharpoonup x$ und $x'_n \to x'$ gilt, so folgt $x'_n x_n \to x'x$.*

b) Falls X Banachraum ist und $x_n \to x$, $x'_n \overset{*}{\rightharpoonup} x'$, so folgt $x'_n x_n \to x'x$.

Beweis. a) Nach Satz 9.7 a) existiert ein $C > 0$ mit $\|x_n\| \leq C$. Damit erhalten wir

$$|x'_n x_n - x'x| \leq \|x'_n - x'\|_{X'} \|x_n\|_X + \|x'x_n - x'x\| \leq C\|x'_n - x'\|_{X'} + \|x'x_n - x'x\| \to 0.$$

b) Analog erhält man mit Satz 9.7 b)

$$|x'_n x_n - x'x| \leq \|x'_n\|_{X'} \|x_n - x\| + |x'_n x - x'x| \to 0 \quad (n \to \infty).$$

$\square$

In Anwendungen hat man oft Folgen von Elementen eines normierten Raums oder Folgen von Funktionalen gegeben. Für die schwache Konvergenz sind daher die folgenden Begriffe sinnvoll:

9.9 Definition. a) Eine Folge $(x_n)_{n\in\mathbb{N}} \subset X$ heißt schwache Cauchy-Folge, falls $(x'x_n)_{n\in\mathbb{N}} \subset \mathbb{K}$ für jedes $x' \in X'$ Cauchyfolge ist. Der Raum X heißt schwach folgenvollständig, wenn jede schwache Cauchyfolge schwach konvergiert.

b) Eine Menge $M \subset X$ heißt schwach folgenkompakt, wenn jede Folge $(x_n)_{n\in\mathbb{N}} \subset M$ eine schwach konvergente Teilfolge besitzt.

c) Eine Menge $M \subset X'$ heißt schwach-*-folgenkompakt, wenn jede Folge $(x'_n)_{n\in\mathbb{N}} \subset M'$ eine schwach-*-konvergente Teilfolge besitzt.

9.10 Beispiel. Das folgende Beispiel zeigt, dass vollständige Räume nicht unbedingt schwach folgenvollständig sind. Sei $X = C([0,1])$, versehen mit der $\|\cdot\|_\infty$-Norm, und sei $x_n \in X$ definiert durch $x_n(t) := t^n$ $(n \in \mathbb{N})$. Man sieht sofort, dass x_n punktweise gegen $x := \chi_{\{1\}}$ konvergiert.

Nach dem Rieszschen Darstellungssatz, Satz 8.27, ist $X' = M([0,1])$, die Menge aller komplexen Borelmaße. Für $\mu \in M([0,1])$ kann man mit einer Variante der majorisierten Konvergenz zeigen, dass $\int x_n(t)d\mu(t) \to \int x(t)d\mu(t)$ $(n \to \infty)$. Damit ist $(x'x_n)_{n\in\mathbb{N}} \subset \mathbb{K}$ konvergent und somit eine Cauchyfolge für jedes $x' \in X'$, d.h. $(x_n)_{n\in\mathbb{N}}$ ist eine schwache Cauchyfolge.

Angenommen, es existiert ein $y \in X$ mit $x_n \rightharpoonup y$. Dann gilt insbesondere für die Funktionale $\delta_{t_0} \in X'$, $f \mapsto f(t_0)$ mit $t_0 \in [0,1]$ (d.h. für die Diracmaße) die Konvergenz $x_n(t_0) = \delta_{t_0}(x_n) \to \delta_{t_0}(y) = y(t_0)$. Also konvergiert x_n punktweise gegen y. Damit folgt aber $y = x = \chi_{\{1\}}$ im Widerspruch zu $y \in X$.

9.11 Satz. *Sei X normiert und separabel. Dann ist $B_{X'} := \{x' \in X' : \|x'\|_{X'} \leq 1\}$ schwach-*-folgenkompakt.*

Beweis. Sei $\{x_n : n \in \mathbb{N}\} \subset X$ eine dichte abzählbare Teilmenge, und sei $\{x'_n : n \in \mathbb{N}\} \subset B_{X'}$ eine Folge.

Da die Folge $(x'_n x_1)_{n\in\mathbb{N}} \subset \mathbb{R}$ durch $\|x_1\|$ beschränkt ist, enthält sie eine konvergente Teilfolge $(x'_{n,1}x_1)$. Analog ist $(x'_{n,1}x_2)_{n\in\mathbb{N}} \subset \mathbb{R}$ beschränkt und enthält eine konvergente Teilfolge $(x'_{n,2}x_2)$ etc.

Für die Diagonalfolge $\widetilde{x}'_n := x'_{n,n}$ konvergiert dann $(\widetilde{x}'_n x_j)_{n\in\mathbb{N}}$ für jedes $j \in \mathbb{N}$. Somit existiert für alle $y \in Y := \operatorname{span}\{x_n : n \in \mathbb{N}\}$ der Grenzwert $\widetilde{x}'y := \lim_{n\to\infty} \widetilde{x}'_n y$. Offensichtlich ist $\widetilde{x}' \in L(Y,\mathbb{K})$ mit $\|x'\|_{L(Y,\mathbb{K})} \le 1$. Wegen $\overline{Y} = X$ existiert genau eine Fortsetzung $x' \in X'$ von $\widetilde{x}'$, und es gilt $\|x'\|_{X'} \le 1$.

Für alle $x \in X$ gilt

$$|(x' - \widetilde{x}'_n)x| \le |(x' - \widetilde{x}'_n)y| + |(x' - \widetilde{x}'_n)(x - y)| \le \tfrac{\varepsilon}{2} + 2\|x - y\| < \varepsilon,$$

wobei $y \in Y$ geeignet gewählt wird und n hinreichend groß ist. Also gilt $\widetilde{x}'_n \stackrel{*}{\rightharpoonup} x'$. □

9.12 Bemerkung. a) In Satz 9.11 ist die Separabilität notwendig. Man kann zeigen, dass etwa für $X = \ell^\infty$ die Aussage nicht gilt.

b) Falls X reflexiv ist (und damit insbesondere ein Banachraum), so folgt aus Satz 9.11 sofort, dass $B_{X'}$ schwach folgenkompakt ist. Denn für reflexive Räume stimmen die schwache und schwach-*-Konvergenz auf X' überein.

c) Verwandt mit obigem Satz ist der Satz von Banach-Alaoglu: Sei X ein (nicht notwendig separabler) Banachraum, und sei $B_{X'} := \{x' \in X' : \|x\|_{X'} \le 1\}$ im Dualraum X' die Einheitskugel in X'. Dann ist $B_{X'}$ schwach-*-kompakt. Man beachte, dass hier die Kompaktheit im Sinn der Überdeckungskompaktheit definiert ist (siehe Kapitel 1), bzgl. der schwach-*-Topologie.

d) Der Zusammenhang zwischen Satz 9.11 und dem Satz von Banach-Alaoglu ist gegeben durch folgende Tatsache: Sei X ein separabler Banachraum. Dann ist $B_{X'}$, versehen mit der schwach-*-Topologie, metrisierbar. Für metrisierbare topologische Räume ist aber die Kompaktheit identisch mit der Folgenkompaktheit (Bemerkung 1.27).

9.13 Lemma. *Sei X normierter Raum und X' separabel. Dann ist auch X separabel.*

Beweis. Sei $(x'_k)_{k\in\mathbb{N}} \subset X'$ eine dichte Teilmenge. O.E. sei $x'_k \ne 0$ für alle $k \in \mathbb{N}$. Dann existiert für alle $k \in \mathbb{N}$ ein $x_k \in X$ mit $\|x_k\|_X = 1$ und $|x'_k x_k| \ge \frac{1}{2}\|x'_k\|_{X'}$. Wir setzen $M := \overline{\operatorname{span}\{x_k : k \in \mathbb{N}\}}$.

Angenommen, es gilt $M \ne X$. Dann existiert nach Korollar 3.6 zum Satz von Hahn-Banach ein $x' \in X'$ mit $\|x'\|_{X'} = 1$ und $x'|_M = 0$. Da $(x'_k)_{k\in\mathbb{N}}$ dicht in X' liegt, existiert eine Teilfolge $(x'_{k_n})_{n\in\mathbb{N}}$ mit $\|x'_{k_n} - x'\|_{X'} \to 0$ $(n \to \infty)$. Andererseits gilt wegen $\|x_{k_n}\|_X = 1$ für alle $n \in \mathbb{N}$:

$$\|x'_{k_n} - x'\| \geq |(x'_{k_n} - x')x_{k_n}| = |x'_{k_n} x_{k_n}| \geq \tfrac{1}{2}\|x'_{k_n}\|_{X'}.$$

Für $n \to \infty$ konvergiert die linke Seite gegen 0, die rechte wegen $x'_{k_n} \to x'$ aber gegen $\frac{1}{2}$, Widerspruch.

Also gilt $M = X$. In M liegen aber endliche Linearkombinationen mit rationalen Koeffizienten dicht, d.h. X ist separabel. □

Die folgende Aussage ist sehr nützlich für viele Anwendungen.

9.14 Satz. *Sei X reflexiver Banachraum. Dann ist $B_X := \{x \in X : \|x\|_X \leq 1\}$ schwach folgenkompakt.*

Beweis. Sei $(x_n)_{n\in\mathbb{N}} \subset B_X$ eine Folge, und sei $M := \overline{\operatorname{span}\{x_n : n \in \mathbb{N}\}}$. Dann ist M ein abgeschlossener Teilraum eines reflexiven Raums und damit selbst reflexiv.

[[Zu $m'' \in M''$ sei $x''(x') := m''(x'|_M)$. Dann ist $x'' \in X''$, und da X reflexiv ist, existiert ein $x \in X$ mit $x'(x) = x''(x') = m''(x'|_M)$ $(x' \in X')$. Falls $x \notin M$, existiert nach dem Satz von Hahn-Banach ein $x' \in X'$ mit $x'(x) = 1$, $x'|_M = 0$. Dies ist aber ein Widerspruch wegen $1 = x'(x) = m''(x'|_M) = m''(0) = 0$. Somit folgt $x \in M$. Für $m' \in M'$ gilt dann

$$m''(m') = m''(x'|_M) = x'(x) = m'(x),$$

wobei $x' \in X'$ eine beliebige Fortsetzung von $m' \in M'$ ist (welche wieder nach Hahn-Banach existiert).]]

Da $J_M \colon M \to M''$ ein isometrischer Isomorphismus ist, ist auch M'' separabel. Nach Lemma 9.3 ist auch M' separabel, und nach Satz 9.11 ist $B_{M''} := \{m'' \in M'' : \|m''\|_{M''} \leq 1\}$ schwach-*-folgenkompakt.

Somit besitzt $(J_M x_n)_{n\in\mathbb{N}}$ eine Teilfolge $(J_M x_{n_k})_{k\in\mathbb{N}}$ so, dass ein $m'' \in M''$ existiert mit $\|m''\|_{M''} \leq 1$ und

$$(J_M x_{n_k})m' \to m''m' \quad (k \to \infty) \text{ für alle } m' \in M'.$$

Sei $x := J_M^{-1} m'' \in M \subset X$. Da J_M eine Isometrie ist, gilt $\|x\|_X \leq 1$, und es folgt für alle $m' \in M'$

$$m' x_{n_k} \to m' x \quad (k \to \infty).$$

Also gilt $x_{n_k} \rightharpoonup x$ $(k \to \infty)$. □

9.15 Korollar. *a) Sei X ein reflexiver Banachraum. Dann ist X schwach folgenvollständig.*

b) Der Raum $C([0,1])$ ist nicht reflexiv.

Beweis. a) Sei $(x_n)_{n\in\mathbb{N}} \subset X$ eine schwache Cauchyfolge. Dann existiert nach Satz 9.6 ein $R > 0$ mit $\|x_n\| \le R$ $(n \in \mathbb{N})$. Da nach Satz 9.14 die Menge $\overline{B(0,R)} \subset X$ schwach folgenkompakt ist, existiert eine konvergente Teilfolge.

b) folgt direkt aus a) und der Tatsache, dass $C([0,1])$ nicht schwach folgenvollständig ist (Beispiel 9.10). □

9.16 Definition. Sei X normierter Raum und $M \subset X$. Dann heißt M schwach folgenabgeschlossen, falls für jede Folge $(x_n)_{n\in\mathbb{N}} \subset M$ mit $x_n \rightharpoonup x \in X$ folgt $x \in M$.

9.17 Bemerkung. a) Abgeschlossene Mengen sind nicht immer schwach folgenabgeschlossen, wie man an folgendem Beispiel sieht: Sei X Hilbertraum, $(e_n)_{n\in\mathbb{N}}$ ein Orthonormalsystem. Dann gilt $e_n \rightharpoonup 0$, d.h. die Menge $\{x \in X : \|x\| = 1\}$ ist zwar abgeschlossen, aber nicht schwach folgenabgeschlossen.

b) Falls M schwach folgenabgeschlossen ist, so ist M auch abgeschlossen, denn jede in der Norm konvergente Folge ist auch schwach konvergent.

Der folgende Satz zeigt, dass bei konvexen Mengen die beiden Versionen der Abgeschlossenheit äquivalent sind.

9.18 Satz. *Sei X normierter Raum und $M \subset X$ konvex und abgeschlossen. Dann ist M schwach folgenabgeschlossen.*

Beweis. Sei $(x_n)_{n\in\mathbb{N}} \subset M$ mit $x_n \rightharpoonup x \in X$. Falls $x \notin M$, so existiert nach dem Trennungssatz 3.7 ein $x' \in X'$ und ein $\alpha \in \mathbb{R}$ mit

$$\operatorname{Re} x'x_n < \alpha < \operatorname{Re} x'x \quad (n \in \mathbb{N}).$$

Dies widerspricht aber der Konvergenz $x'x_n \to x'x$ $(n \to \infty)$. □

9.19 Definition. Seien X ein $\mathbb{K}$-Vektorraum und $M \subset X$. Dann ist die konvexe Hülle $\operatorname{conv} M$ von M definiert als die kleinste konvexe Menge, welche eine Obermenge von M ist, d.h. es gilt

$$\operatorname{conv}(M) := \bigcap_{\substack{Z \text{ konvex} \\ M\subset Z\subset X}} Z.$$

9.20 Bemerkung. a) Die konvexe Hülle ist gleich der Menge aller endlichen Konvexkombinationen $\sum_{j=1}^n \lambda_j x_j$ mit $\lambda_j \ge 0$, $\sum_{j=1}^n \lambda_j = 1$, $x_j \in M$, $n \in \mathbb{N}$.

b) Der Abschluss $\overline{\operatorname{conv}(M)}$ ist wieder konvex. Denn seien $x, y \in \overline{\operatorname{conv}(M)}$. Dann existieren Folgen $(x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}} \subset M$ mit $x_n \to x$, $y_n \to y$, und für alle $\lambda \in [0,1]$

gilt $\lambda x_n + (1-\lambda)y_n \to \lambda x + (1-\lambda)y$. Wegen $\lambda x_n + (1-\lambda)y_n \in \operatorname{conv}(M)$ folgt $\lambda x + (1-\lambda)y \in \overline{\operatorname{conv}(M)}$.

Der folgende Satz besagt, dass man bei Übergang zu Konvexkombinationen aus schwacher Konvergenz sogar starke Konvergenz erhalten kann.

9.21 Satz (von Mazur)**.** *Seien X normierter Raum und $(x_n)_{n\in\mathbb{N}} \subset X$ mit $x_n \rightharpoonup x \in X$. Dann existiert eine Folge $(y_n)_{n\in\mathbb{N}}$ mit $y_n \in \operatorname{conv}\{x_1,\dots,x_n\}$ so, dass $y_n \to x$.*

Somit existieren $\lambda_1^{(n)},\dots,\lambda_n^{(n)} \geq 0$ mit $\sum_{j=1}^n \lambda_j^{(n)} = 1$ und

$$\sum_{j=1}^n \lambda_j^{(n)} x_j \to x \quad (n\to\infty).$$

Beweis. Die Menge $M := \overline{\operatorname{conv}(\{x_n : n\in\mathbb{N}\})}$ ist konvex und abgeschlossen. Nach Satz 9.18 ist M schwach folgenabgeschlossen, d.h. es gilt $x \in M$. Wegen $\operatorname{conv}(M) = \bigcup_{n\in\mathbb{N}} \operatorname{conv}\{x_1,\dots,x_n\}$ existiert zu jedem $k\in\mathbb{N}$ ein $n_k\in\mathbb{N}$ und ein $y_{n_k} \in \operatorname{conv}\{x_1,\dots,x_{n_k}\}$ mit $\|x - y_{n_k}\| \leq \frac{1}{k}$. Nun kann man die Folge $(y_{n_k})_{k\in\mathbb{N}}$ noch stückweise konstant ergänzen zu einer Folge $(y_n)_{n\in\mathbb{N}}$ mit $y_n \in \operatorname{conv}\{x_1,\dots,x_n\}$, etwa durch

$$\begin{aligned} y_1 = y_2 = \ldots = y_{n_1-1} &= x_1 \\ y_{n_1+1} = y_{n_1+2} = \ldots = y_{n_2-1} &= y_{n_1}, \\ y_{n_2+1} = y_{n_2+2} = \ldots = y_{n_3-1} &= y_{n_2} \qquad \text{etc.} \end{aligned}$$

□

10. Die Fourier-Transformation

10.1 Worum geht's? Die Fourier-Transformation beschreibt mathematisch die Zerlegung einer Funktion oder eines akustischen oder elektromagnetischen Signals in Grundschwingungen. Innerhalb der Mathematik tritt die Fourier-Transformation insbesondere im Rahmen der partiellen Differentialgleichungen auf. Dies liegt daran, dass die Ableitung in der Fourier-Transformierten zur Multiplikation mit der Koordinatenfunktion wird.

Die Fourier-Transformierte ist zunächst für Funktionen in $L^1(\mathbb{R}^n)$ definiert. Schränkt man die Fourier-Transformation auf den Schwartz-Raum $\mathscr{S}(\mathbb{R}^n)$ ein, so erhalten wir eine Isometrie bzgl. $\|\cdot\|_2$-Norm. Damit können wir die Fourier-Transformation auch auf $L^2(\mathbb{R}^n)$ definieren, was für Anwendungen wichtig ist, da es sich bei $L^2(\mathbb{R}^n)$ um einen Hilbertraum handelt.

Im folgenden wird die Multiindex-Schreibweise aus Definition 5.2 verwendet, wobei zusätzlich

$$D^\alpha := (-i)^{|\alpha|}\Big(\frac{\partial}{\partial x_1}\Big)^{\alpha_1}\cdots\Big(\frac{\partial}{\partial x_n}\Big)^{\alpha_n}$$

gesetzt wird.

10.2 Definition. Für $f \in L^1(\mathbb{R}^n)$ heißt $\mathscr{F}f$, definiert durch

$$(\mathscr{F}f)(\xi) := \hat{f}(\xi) := \frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n} f(x)e^{-ix\xi}dx \quad (\xi \in \mathbb{R}^n),$$

die Fourier-Transformierte von f.

10.3 Lemma (Elementare Eigenschaften)**.** *Seien $f \in L^1(\mathbb{R}^n)$, $\alpha > 0$ und $a \in \mathbb{R}^n$.*

a) Für $g(x) := f(x)e^{iax}$ gilt $\hat{g}(\xi) = \hat{f}(\xi - a)$.

b) Für $g(x) := f(x-a)$ gilt $\hat{g}(\xi) = \hat{f}(\xi)e^{-ia\xi}$.

c) Für $g(x) := \overline{f(-x)}$ gilt $\hat{g}(\xi) = \overline{\hat{f}(\xi)}$.

d) Für $g(x) := f(\frac{x}{\alpha})$ gilt $\hat{g}(\xi) = \alpha^n\hat{f}(\alpha\xi)$.

e) Sei $g \in L^1(\mathbb{R}^n)$ und $h := f * g$. Dann gilt $\hat{h}(\xi) = (2\pi)^{n/2}\hat{f}(\xi)\cdot\hat{g}(\xi)$.

Beweis. Die Aussagen a)-c) folgen unmittelbar aus der Definition, Teil d) folgt aus dem Transformationssatz. Für e) verwende man den Satz von Fubini und erhalte

$$\begin{aligned}\hat{h}(\xi) &= (2\pi)^{-n/2}\int\Big[\int f(x-y)g(y)dy\Big]e^{-ix\xi}dx\\ &= (2\pi)^{-n/2}\int\Big[\int f(x-y)g(y)dy\Big]e^{-i(x-y)\xi}e^{-iy\xi}dx\end{aligned}$$

$$= (2\pi)^{n/2} \cdot (2\pi)^{-n/2} \int g(y) \Big[(2\pi)^{-n/2} \int f(z) e^{-iz\xi} dz \Big] e^{-iy\xi} dy$$
$$= (2\pi)^{n/2} \hat{f}(\xi) \hat{g}(\xi).$$

□

10.4 Satz. *Für $f \in L^1(\mathbb{R}^n)$ ist $\mathscr{F} f \in C_0(\mathbb{R}^n)$. Die Fourier-Transformation $\mathscr{F} : L^1(\mathbb{R}^n) \to (C_0(\mathbb{R}^n), \|\cdot\|_\infty)$ ist ein stetiger linearer Operator mit Norm $\|\mathscr{F}\| \leq (2\pi)^{-n/2}$.*

Beweis. Offensichtlich ist $\mathscr{F} f$ messbar mit $\|\mathscr{F} f\|_\infty \leq (2\pi)^{-n/2} \|f\|_1$, d.h. $\mathscr{F} : L^1(\mathbb{R}^n) \to L^\infty(\mathbb{R}^n)$ ist stetig mit Norm nicht größer als $(2\pi)^{-n/2}$.

(i) Wir zeigen, dass $\mathscr{F} f$ stetig ist. Sei dazu $(\xi^{(k)})_{k\in\mathbb{N}} \subset \mathbb{R}^n$ mit $\xi^{(k)} \to \xi$. Dann gilt $e^{-ix\xi^{(k)}} \to e^{-ix\xi} \quad (x \in \mathbb{R}^n)$ und

$$|\mathscr{F} f(\xi^{(k)}) - \mathscr{F} f(\xi)| \leq (2\pi)^{-n/2} \int_{\mathbb{R}^n} |f(x)| \cdot \underbrace{|e^{-ix\xi^{(k)}} - e^{-ix\xi}|}_{\leq 2} dx \to 0$$

mit majorisierter Konvergenz.

(ii) Es gilt $\mathscr{F} f(\xi) \to 0$ für $|\xi| \to \infty$. Dazu verwenden wir, dass $\mathscr{D}(\mathbb{R}^n)$ dicht in $L^1(\mathbb{R}^n)$ ist. Da $\mathscr{F} : L^1 \to L^\infty$ stetig ist, reicht es, $\mathscr{F} f(\xi) \to 0$ für $f \in \mathscr{D}(\mathbb{R}^n)$ zu zeigen.

Sei $f \in \mathscr{D}(\mathbb{R}^n)$, $R > 0$ und $\xi \in \mathbb{R}^n$ mit $|\xi| \geq R$. Wähle j mit $|\xi_j| \geq \frac{R}{\sqrt{n}}$. Dann gilt

$$\begin{aligned} |\mathscr{F} f(\xi)| &= \Big| \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} f(x) e^{-ix\xi} dx \Big| \\ &= \Big| - \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \partial_{x_j} f(x) \, \frac{1}{-i\xi_j} \, e^{-ix\xi} dx \Big| \\ &\leq (2\pi)^{-n/2} \|\partial_{x_j} f\|_1 \cdot \frac{\sqrt{n}}{R} \ \to 0 \quad (R \to \infty). \end{aligned}$$

□

10.5 Definition. Der Schwartzraum $\mathscr{S}(\mathbb{R}^n)$ wird definiert als die Menge aller $f \in C^\infty(\mathbb{R}^n)$, für welche für alle $\alpha, \beta \in \mathbb{N}_0^n$ gilt:

$$p_{\alpha,\beta}(f) := \sup_{x\in\mathbb{R}^n} |x^\alpha D^\beta f(x)| < \infty.$$

Versehen mit der Familie $L = \{p_{\alpha,\beta} : \alpha, \beta \in \mathbb{N}_0^n\}$ wird $\mathscr{S}(\mathbb{R}^n)$ ein lokalkonvexer topologischer Raum (siehe Definition 2.10), der sogar Fréchetraum, d.h. metrisierbar und vollständig ist.

10.6 Bemerkung. a) Definiere die Familie $\widetilde{L} := \{p_N : N \in \mathbb{N}\}$ von Normen auf $\mathscr{S}(\mathbb{R}^n)$ durch

$$p_N(f) := \sup\{(1+|x|)^N |D^\alpha f(x)| : |\alpha| \le N,\ x \in \mathbb{R}^n\} \quad (f \in \mathscr{S}(\mathbb{R}^n)).$$

Dann wird durch $\widetilde{L}$ dieselbe lokalkonvexe Topologie auf $\mathscr{S}(\mathbb{R}^n)$ erzeugt. Im folgenden wird $\mathscr{S}(\mathbb{R}^n)$ stets mit dieser lokalkonvexen Topologie versehen.

Man beachte, dass für eine Folge $(f_k)_{k\in\mathbb{N}} \subset \mathscr{S}(\mathbb{R}^n)$ und für $f \in \mathscr{S}(\mathbb{R}^n)$ gilt

$$f_k \to f \ (k \to \infty) \text{ in } \mathscr{S}(\mathbb{R}^n) \quad \Longleftrightarrow \quad \forall\, N \in \mathbb{N} :\ p_N(f - f_k) \to 0 \ (k \to \infty).$$

b) Der Schwartzraum $\mathscr{S}(\mathbb{R}^n)$ liegt dicht in $L^p(\mathbb{R}^n)$ für alle $1 \le p < \infty$. Denn jedes $f \in \mathscr{S}(\mathbb{R}^n)$ liegt in $L^p(\mathbb{R}^n)$ wegen

$$\int_{\mathbb{R}^n} |f(x)|^p dx \le \int_{\mathbb{R}^n} c(1+|x|)^{-mp} dx < \infty$$

für $mp > n$. Die Dichtheit von $\mathscr{S}(\mathbb{R}^n) \subset L^p(\mathbb{R}^n)$ ist klar wegen $\mathscr{D}(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n)$.

c) Analog zur Definition von Distributionen $T \in \mathscr{D}'(\mathbb{R}^n)$ wird auch der Dualraum von $\mathscr{S}(\mathbb{R}^n)$ betrachtet. Man definiert

$$\mathscr{S}'(\mathbb{R}^n) := \{T \colon \mathscr{S}(\mathbb{R}^n) \to \mathbb{C} : T \text{ linear, stetig }\}.$$

Man beachte, dass dabei auf $\mathscr{S}(\mathbb{R}^n)$ die obige lokalkonvexe Topologie gewählt wird. Der Raum $\mathscr{S}'(\mathbb{R}^n)$ heißt der Raum der temperierten Distributionen.

10.7 Lemma. *Für $f \in \mathscr{S}(\mathbb{R}^n)$ und $\alpha \in \mathbb{N}_0^n$ gilt:*

(i) $\mathscr{F} f \in C^\infty(\mathbb{R}^n)$ und $D^\alpha(\mathscr{F} f) = (-1)^{|\alpha|}\mathscr{F}(x^\alpha f)$.

(ii) $\mathscr{F}(D^\alpha f)(\xi) = \xi^\alpha (\mathscr{F} f)(\xi) \quad (\xi \in \mathbb{R}^n)$.

(iii) $\mathscr{F} f \in \mathscr{S}(\mathbb{R}^n)$.

Beweis. (i) Da $x \mapsto x^\alpha f \in L^1(\mathbb{R}^n)$ eine integrierbare Majorante der Ableitung des Integranden ist, folgt aus dem Satz über parameterabhängige Integrale

$$\begin{aligned} D^\alpha(\mathscr{F} f)(\xi) &= (2\pi)^{-n/2} \int_{\mathbb{R}^n} f(x) D_\xi^\alpha\, e^{-ix\xi} d\xi \\ &= \frac{(-i)^{2|\alpha|}}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} f(x) x^\alpha e^{-ix\xi} d\xi = (-1)^{|\alpha|}\big(\mathscr{F}(x^\alpha f)\big)(\xi). \end{aligned}$$

(ii) Mit partieller Integration erhalten wir

$$\mathscr{F}(D^\alpha f)(\xi) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} (D^\alpha f)(x) e^{-ix\xi} dx$$

$$= \frac{(-1)^{|\alpha|}}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} f(x) D_x^\alpha e^{-ix\xi} dx = \xi^\alpha \mathscr{F} f(\xi).$$

(iii) Nach (i) gilt $\mathscr{F} f \in C^\infty(\mathbb{R}^n)$. Außerdem gilt

$$x^\alpha D^\beta (\mathscr{F} f)(\xi) = (-1)^{|\beta|} \mathscr{F} (\underbrace{D^\alpha x^\beta f}_{\in \mathscr{S}(\mathbb{R}^n)})(\xi) \to 0 \quad (|\xi| \to \infty),$$

wobei wir Satz 10.4 verwendet haben. □

10.8 Lemma. *Für* $\gamma(x) := \exp(-\frac{|x|^2}{2})$ *gilt* $\mathscr{F}\gamma = \gamma$.

Beweis. (i) Sei $n = 1$. Die Funktion γ ist die eindeutige Lösung des Anfangswertproblems

$$y' + xy = 0, \quad y(0) = 1. \tag{10-1}$$

Damit gilt

$$0 = \mathscr{F}(\gamma' + x\gamma) = i\xi(\mathscr{F}\gamma) + i(\mathscr{F}\gamma)'.$$

Wegen

$$(\mathscr{F}\gamma)(0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = 1$$

löst $\mathscr{F}\gamma$ ebenfalls das Anfangswertproblem (10-1). Also gilt $\gamma = \mathscr{F}\gamma$.

(ii) Für $n > 1$ schreiben wir

$$\begin{aligned}(\mathscr{F}\gamma)(\xi) &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \Big(\prod_{j=1}^n e^{-x_j^2/2} \cdot \prod_{j=1}^n e^{-ix_j\xi_j} \Big) dx \\ &= \prod_{j=1}^n \Big(\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-x_j^2/2} e^{-ix_j\xi_j} dx_j \Big) \\ &= \prod_{j=1}^n e^{-\xi_j^2/2} = \gamma(\xi).\end{aligned}$$

□

10.9 Lemma. *a) Für* $f, g \in \mathscr{S}(\mathbb{R}^n)$ *gilt*

$$\int_{\mathbb{R}^n} f(x) \mathscr{F} g(x) dx = \int_{\mathbb{R}^n} \mathscr{F} f(x) g(x) dx,$$

d.h. $\langle f, \overline{\mathscr{F} g} \rangle_2 = \langle \mathscr{F} f, \overline{g} \rangle_2$.

b) Für $f \in \mathscr{S}(\mathbb{R}^n)$ *gilt* $\mathscr{F}^2 f(x) = f(-x) \quad (x \in \mathbb{R}^n)$.

Beweis. a) folgt sofort aus dem Satz von Fubini, da $(x, y) \mapsto f(x)g(y)e^{-ixy} \in L^1(\mathbb{R}^{2n})$.

b) Sei $\gamma_a(x) := \gamma(ax)$ für $a > 0$ mit γ wie in Lemma 10.8. Nach Lemma 10.3 gilt

$$(\mathscr{F}\gamma_a)(\xi) = a^{-n}(\mathscr{F}\gamma)\Big(\frac{\xi}{a}\Big).$$

Sei $g(x) := e^{-ix\xi_0}\gamma_a(x)$ für $\xi_0 \in \mathbb{R}^n$ fest und $a > 0$ fest. Dann ist $g \in \mathscr{S}(\mathbb{R}^n)$ und

$$(\mathscr{F}g)(\xi) = (\mathscr{F}\gamma_a)(\xi + \xi_0) = a^{-n}\gamma\Big(\frac{\xi + \xi_0}{a}\Big).$$

Unter Verwendung von a) erhalten wir

$$\begin{aligned}\frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n}(\mathscr{F}f)(x)g(x)dx &= \frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n} f(x)(\mathscr{F}g)(x)dx \\ &= \frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n} f(x)a^{-n}(\mathscr{F}\gamma)\Big(\frac{x + \xi_0}{a}\Big)dx \\ &= \frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n} f(au - \xi_0)\gamma(u)du. \qquad (10\text{-}2)\end{aligned}$$

Bei der letzten Gleichheit haben wir die Substitution $u = \frac{x+\xi_0}{a}$ und die Identität $\mathscr{F}\gamma = \gamma$ verwendet.

Wir nehmen den Grenzwert $a \to 0$. Es gilt dann $\gamma(au) \to 1$ punktweise und $g(x) \to e^{-ix\xi_0}$ punktweise. Wegen $\mathscr{F}f \in L^1(\mathbb{R}^n)$ können wir majorisierte Konvergenz anwenden und erhalten

$$\frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n}(\mathscr{F}f)(x)g(x)dx \to (\mathscr{F}^2 f)(\xi_0).$$

Um den Grenzwert für die rechte Seite von (10-2) zu berechnen, verwenden wir $f(au - \xi_0) \to f(-\xi_0)$ punktweise. Da $\|f\|_\infty \cdot \gamma$ eine integrierbare Majorante ist, erhalten wir

$$\frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n} f(au - \xi_0)\gamma(u)du \to f(-\xi_0).$$

Also gilt $(\mathscr{F}^2 f)(\xi_0) = f(-\xi_0)$. □

10.10 Definition und Satz. *a) Die Fourier-Transformation $\mathscr{F} : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ ist eine Bijektion mit*

$$(\mathscr{F}^{-1}g)(x) = \frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n} g(\xi)e^{ix\xi}d\xi.$$

b) (Satz von Plancherel.) Es gilt

$$\|f\|_{L^2(\mathbb{R}^n)} = \|\mathscr{F}f\|_{L^2(\mathbb{R}^n)} \quad (f \in \mathscr{S}(\mathbb{R}^n)).$$

Somit ist $\mathscr{F}|_{\mathscr{S}(\mathbb{R}^n)}$ eine Isometrie und damit eindeutig zu einem isometrischen Isomorphismus $\mathscr{F}_2 : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ fortsetzbar, der ebenfalls Fourier-Transformation genannt wird. Insbesondere gilt

$$\|\mathscr{F}_2 f\|_2 = \|f\|_2 \quad (f \in L^2(\mathbb{R}^n)).$$

Beweis. a) Nach Lemma 10.9 gilt $\mathscr{F}^2 f(x) = f(-x)$, d.h. $\mathscr{F}^4 = \mathrm{id}_{\mathscr{S}(\mathbb{R}^n)}$. Damit gilt

$$(\mathscr{F}^{-1}g)(x) = (\mathscr{F}^3 g)(x) = (\mathscr{F}g)(-x) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} g(\xi) e^{ix\xi} d\xi.$$

b) Im Beweis von Lemma 10.9 sahen wir

$$\langle \mathscr{F}f, \overline{g} \rangle_{L_2(\mathbb{R}^n)} = \langle f, \overline{\mathscr{F}g} \rangle_{L_2(\mathbb{R}^n)}.$$

Setze nun $g := \mathscr{F}^{-1}\overline{g}$, d.h.

$$\overline{g} = \overline{\mathscr{F}^{-1}\overline{f}} = (2\pi)^{-n/2} \overline{\int \overline{f(x)} e^{ix\xi} dx} = \mathscr{F}f.$$

Damit folgt $\|\hat{f}\|_{L^2} = \|f\|_{L^2}$.

Da $\mathscr{S}(\mathbb{R}^n) \subset L_2(\mathbb{R}^n)$ dicht ist, ist die Fortsetzung $\mathscr{F}_2 : L_2(\mathbb{R}^n) \to L_2(\mathbb{R}^n)$ wohldefiniert und ebenfalls isometrisch. Zu zeigen ist noch, dass $\mathscr{F}_2$ surjektiv ist. Nach Lemma 6.12 ist $R(\mathscr{F}_2)$ abgeschlossen. Wegen $R(\mathscr{F}_2) \subset \mathscr{F}_2(\mathscr{S}(\mathbb{R}^n)) = \mathscr{S}(\mathbb{R}^n)$ ist $R(\mathscr{F}_2)$ dicht in $L^2(\mathbb{R}^n)$ und damit $\mathscr{F}_2$ surjektiv. □

10.11 Definition und Satz. *Für $T \in \mathscr{S}'(\mathbb{R}^n)$ definiere $\mathscr{F}T(\varphi) := T(\mathscr{F}\varphi)$ $(\varphi \in \mathscr{S}(\mathbb{R}^n))$. Dann ist $\mathscr{F}\colon \mathscr{S}'(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ linear, stetig und bijektiv mit stetigem Inversen.*

Literatur

[1] Adams, R. A.: Sobolev spaces. Academic Press, New York 1975.

[2] Denk, R., Racke, R.: Kompendium der Analysis. Band 2 Springer-Vieweg, Wiesbaden 2012.

[3] Diestel, J., Uhl, J. J.: Vector measures. Amer. Math. Soc., Providence, 1977.

[4] Dunford, N., Schwartz, J. T.: Linear Operators. I. General Theory. Interscience Publishers, New York 1963.

[5] Edwards, R. E.: Functional analysis. Theory and applications. Dover, New York 1995.

[6] Gohberg, I., Goldberg, S.: Basic operator theory. Birkhäuser Boston 1981.

[7] Gohberg, I., Goldberg, S., Kaashoek, M. A.: Classes of linear operators. Vol. I. Birkhäuser Basel 1990.

[8] Halmos, P.: A Hilbert Space Problem Book. Springer-Verlag, Berlin 1974.

[9] Heuser, H.: Funktionalanalysis. B.G. Teubner Verlag, Stuttgart 1986.

[10] Hirzebruch, F., Scharlau, W.: Einführung in die Funktionalanalysis. Bibliogr. Inst. Mannheim 1971.

[11] Kato, T.: Perturbation theory for linear operators. Springer Berlin 1976.

[12] Pazy, A.: Semigroups of Linear Operators and Applications to Partial Differential Equations. Springer, Berlin et al. 1983.

[13] Reed, M., Simon, B.: Methods of modern mathematical physics. I. Functional analysis. Academic Press San Diego 1980.

[14] Rudin, W.: Functional analysis. McGraw-Hill New York 1973.

[15] Rudin, W.: Real and complex analysis. McGraw-Hill New York 1986.

[16] Schaefer, H.H., Wolff, M.P.: Topological vector spaces. Springer Berlin, 1999.

[17] Schröder, H.: Funktionalanalysis. Akademie Verlag Berlin, 1997.

[18] Werner, D.: Funktionalanalysis (5., erw. Aufl.). Springer Berlin 2005.

[19] Yosida, K.: Functional Analysis. Springer-Verlag, Berlin et al. 1965.

Index