Einführung in die Funktionalanalysis

(Sommersemester 2011)

Raimar Wulkenhaar

In der Funktionalanalysis werden unendlich-dimensionale topologischen Vektorräume sowie lineare Abbildungen zwischen diesen studiert. Dazu werden Methoden aus Analysis, Topologie und linearer Algebra verknüpft. Die historischen Wurzeln der Funktionalanalysis sind die Fourier-Transformation und die Theorie der Integralgleichungen; wichtige Motivation kam später aus der Quantenphysik. Die grundlegenden Sätze der Funktionalanalysis sind in vielen Bereichen der theoretischen und angewandten Mathematik unverzichtbar.

Inhalt

0 Einleitung . . . 1

I Normierte Vektorräume **3**

1 Topologische Vektorräume und normierte Räume . . . 3
2 Beispiele für normierte Vektorräume . . . 8
3 Hahn-Banach-Sätze . . . 17
4 Schwache Topologien und reflexive Räume . . . 24
5 Gleichmäßig konvexe Räume . . . 33
6 Der Satz von Baire und seine Konsequenzen . . . 37

II Lokalkonvexe Vektorräume **45**

7 Netze . . . 45
8 Systeme von Halbnormen . . . 47
9 Konvexität . . . 51
10 Der Satz von Stone-Weierstraß . . . 55
11 Metrisierbarkeit und Vollständigkeit . . . 63

III Hilbert-Räume **67**

12 Grundlegende Eigenschaften . . . 67
13 Summierbarkeit und Orthonormalbasen . . . 70
14 Operatoren auf Hilbert-Räumen . . . 75
15 Das Spektrum . . . 83
16 Kompakte Operatoren . . . 90

Literatur

F. Hirzebruch & W. Scharlau, "Einführung in die Funktionalanalysis," Spektrum Akademischer Verlag (1996)

M. Reed & B. Simon, "Methods of Modern Mathematical Physics I: Functional Analysis," Academic Press (1980)

D. Werner, "Funktionalanalysis," Springer-Verlag (2007)

W. Rudin, "Functional Analysis," McGraw-Hill (2006)

H. Heuser, "Funktionalanalysis," Teubner (2006)

0 Einleitung

0.1 Übersicht über die Räume in der Funktionalanalysis

Topologischer Raum
– offene Mengen, abgeschlossen
– Abbildungen: *Stetigkeit*
– Konvergenz (von Folgen, Netzen)
– Trennungsaxiome
– Abzählbarkeit
– Kompaktheit

Topologischer Vektorraum
– Vektor-Operationen stetig
– Abbildungen: *linear+stetig*
– Unterräume, Quotientenräume

Metrischer Raum
– Metrik (Abstand) d
– offene Kugeln
– Topologie "gutartig"
– *Vollständigkeit*
(Cauchy-Folgen konvergent)
– Bsp: Sphäre, Torus
– Differentialgeometrie

Lokalkonvexer Raum
– System von Halbnormen
(metrisierbar)
– Vollständigkeit: *Fréchet-Raum*
– Bsp: $C^\infty(X)$
– Abbildungen: z.B. *Differenzierbarkeit*
– schwache Topologien

Normierter Vektorraum
– Norm $\| \ \|$
– Vollständigkeit: *Banach-Raum*
– Bsp: $C(X)$, $L^p(X)$, $\mathcal{L}(X,Y)$
– Abbildungen: *linear+beschränkt*
(sehr reichhaltige Strukturen,
Topologie kann "wild" sein)
– Dualräume
– Fortsetzungssatz von Hahn-Banach
(lineare Funktionale mit
Zusatzeigenschaften,
erfordert Lemma von Zorn)
– Bairescher Kategoriensatz
(gleichmäßige Beschränktheit,
Satz von der offenen Abbildung)
– Spektraltheorie

Unitärer Vektorraum
– Skalarprodukt $\langle \ , \ \rangle$, Cauchy-Schwarz
– Vollständigkeit: *Hilbert-Raum*
– Bsp: ℓ^2, $L^2(X)$
– Orthonormalbasen
– Abbildungen: *linear+beschränkt*
(+normal, unitär, selbstadjungiert,
kompakt, Fredholm)
– Spektralsatz
– Funktionalkalkül
– Unbeschränkte Operatoren

Operatoralgebren
– C^*-Algebra
– Von–Neumann-Algebra

0.2 Meilensteine

Historische Notizen findet man in
– Heuser, Kapitel XIX, und
– Werner, am Ende jedes Kapitels

- 1900: Fredholm – Theorie der Integralgleichungen
- 1904…1910: Hilbert-Schmidt: Spektraltheorie von Operatoren
- 1910: Riesz – L^p-Räume
- 1920: Banach – normierter Raum
- 1927: Hahn – Fortsetzungssatz
- 1929: von Neumann – Hilbert-Raum, unbeschränkte Operatoren
- 1938: Sobolev – verallgemeinerte Ableitungen
- 1943: Gelfand-Naimark – Operatoralgeben
- 1948: Schwartz – Distributionen

Teil I
Normierte Vektorräume

1 Topologische Vektorräume und normierte Räume

In dieser Vorlesung ist $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Definition und grundlegende Eigenschaften topologischer Räume werden vorausgesetzt. Wir erinnern dennoch an den Begriff der Stetigkeit:

Definition 1.1 Seien X, Y topologische Räume. Eine Abbildung $f : X \to Y$ heißt

i) *stetig*, wenn das Urbild $f^{-1}(\mathcal{O})$ jeder offenen Teilmenge $\mathcal{O} \subset Y$ offen in X ist, und

ii) *stetig im Punkt* $x \in X$, wenn es zu jeder Umgebung $V \subset Y$ von $f(x)$ eine Umgebung $U \subset X$ von x gibt mit $f(U) \subset V$.

Eine *Umgebung* eines Punktes $x \in X$ ist eine Teilmenge $U \subset X$, für die es eine offene Teilmenge $\mathcal{O}$ gibt mit $x \in \mathcal{O} \subset U$. Eine Umgebung muß selbst nicht offen sein. Eine offene Teilmenge ist Umgebung für jeden ihrer Punkte.

Lemma 1.2 *Seien X, Y topologische Räume. Für eine Abbildung $f : X \to Y$ sind äquivalent:*

i) *f ist stetig.*

ii) *f ist in jedem Punkt $x \in X$ stetig.*

Beweis. i)⇒ii) ist klar, wenn man die offenen Teilmengen der Umgebungen nach i) wählt.

ii)⇒i) $V \subset Y$ offen ist Umgebung jedes Punktes $y \in V$. Sei $U := f^{-1}(V)$. Für jedes $x \in U$ gibt es eine Umgebung U_x und damit eine offene Teilmenge $\mathcal{O}_x \ni x$ mit $f(\mathcal{O}_x) \subset V$. Dann ist $U = \bigcup_{x \in U} \mathcal{O}_x = f^{-1}(V)$ offen. □

Definition 1.3 Es sei X ein Vektorraum über $\mathbb{K}$ und $\mathcal{T}$ eine Topologie auf X. Dann heißt $(X, \mathcal{T})$ *topologischer Vektorraum*, wenn gilt

i) die Addition $X \times X \ni (x, y) \mapsto x + y \in X$ ist stetig,

ii) die Multiplikation mit Skalaren $\mathbb{K} \times X \ni (\lambda, x) \mapsto \lambda x \in X$ ist stetig.

Bemerkung 1.4 Hier sind $X \times X$ und $\mathbb{K} \times X$ wieder topologische Räume, versehen mit der Produkttopologie. Die Topologie auf $\mathbb{K}$ wird durch die Metrik $d(\lambda, \mu) = |\lambda - \mu|$ induziert.

Stetigkeit der Addition in (x_0, y_0) heißt, daß es zu jeder Umgebung U von $x_0 + y_0$ Umgebungen U_1 von x_0 und U_2 von y_0 gibt mit $U_1 + U_2 := \{x + y \mid x \in U_1\ ,\ y \in U_2\} \subset U$. Analog für λx.

Rudin fordert noch iii) jeder Punkt von X ist abgeschlossen. Damit wird jeder topologische Vektorraum ein Hausdorff-Raum (T_2).

Lemma 1.5 *Sei X ein topologischer Vektorraum, $a \in X$ und $0 \neq \lambda \in \mathbb{K}$ fest vorgegeben. Dann sind die Abbildungen $T_a : X \ni x \mapsto a + x \in X$ und $M_\lambda : X \ni x \mapsto \lambda x \in X$ Homöomorphismen.*

Beweis. Die Abbildungen sind invertierbar, $(T_a)^{-1} = T_{-a}$ und $(M_\lambda)^{-1} = M_{\frac{1}{\lambda}}$. Stetigkeit der usprünglichen und damit der invertierten Abbildungen folgt aus der Definition. □

Ist $V \subset X$ eine Umgebung von 0, dann ist $T_x(V) = x + V$ eine Umgebung von $x \in X$, und umgekehrt ist für eine Umgebung U von x die Teilmenge $T_{-x}(U)$ eine Umgebung von 0. Ist $A \subset X$ eine beliebige Teilmenge, so ist der topologische Abschluß $\bar{A} = A \cup A'$ erklärt. Dabei ist A' die Menge der Häufungspunkte von A, d.h. jener Punkte $y \in X$, $y \notin A$, so daß für eine beliebige Umgebung $y + V$ von y gilt $(y+V) \cap A \neq \varnothing$. Folglich gilt $x \in \bar{A} \quad \Leftrightarrow \quad x \in A + V$ für jede Umgebung V von 0, somit $\bar{A} = \bigcap_V (A + V)$.

Lemma 1.6 *Sei X ein topologischer Vektorraum und $W \subset X$ Untervektorraum. Dann ist auch $\bar{W}$ (die abgeschlossene Hülle von W) ein Untervektorraum.*

Beweis. Seien $x_0, y_0 \in \bar{W}$. Zu zeigen ist: Zu jeder Umgebung V von $x_0 + y_0$ gibt es ein $w \in W \cap V$. Nach Stetigkeit der Addition gibt es Umgebungen U_1 von x_0 und U_2 von y_0 mit $U_1 + U_2 \subset V$. Somit gibt es $x \in U_1 \cap W$ und $y \in U_2 \cap W$. Es folgt $x + y =: w \in W \cap V$. □

Satz 1.7 *Seien X, Y topologische Vektorräume. Dann sind für eine lineare Abbildung $f : X \to Y$ äquivalent:*

i) *f ist stetig.*

ii) *f ist stetig in 0.*

iii) *f ist stetig in einem beliebigen Punkt $x \in X$.*

Beweis. iii)⇒i)⇒ii) ist klar. Für festes $x \in X$ gilt

$$f(x) = f(x+0) = f(x) + f(0) = T^Y_{f(x)}(f(0)) = \big(T^Y_{f(x)} \circ f \circ T^X_{-x}\big)(x)$$

Die Komposition stetiger Abbildungen ist stetig, somit ii)⇒iii). □

Bezeichnung 1.8 Seien X, Y topologische $\mathbb{K}$-Vektorräume. Wir bezeichen mit $\mathcal{L}(X, Y)$ die Menge der linearen stetigen Abbildungen von X nach Y. Mit den "punktweise" erklärten Operationen $(f_1 + f_2)(x) := f_1(x) + f_2(x) = T_{f_1(x)}(f_2(x))$ und, für $\lambda \neq 0$, $(\lambda f)(x) := \lambda(f(x)) = M_\lambda(f(x))$ wird $\mathcal{L}(X, Y)$ zu einem $\mathbb{K}$-Vektorraum. Speziell heißt

$$X' := \mathcal{L}(X, \mathbb{K})$$

der *Dualraum* von X. Mit der Komposition als Multiplikation wird $\mathcal{L}(X, X)$ zu einer $\mathbb{K}$-Algebra. Elemente aus $\mathcal{L}(X, X)$ heißen *lineare stetige Operatoren auf X*.

Definition 1.9 Sei X ein Vektorraum über $\mathbb{K}$. Eine Abbildung $\| \ \| : X \ni x \mapsto \|x\| \in \mathbb{R}$ heißt *Norm*, falls für beliebige $x, y \in X$ und $\lambda \in \mathbb{K}$ gilt

i) $\|x\| \geq 0$ und $\|x\| = 0 \ \Leftrightarrow \ x = 0$,

ii) $\|\lambda x\| = |\lambda| \|x\|$,

iii) $\|x + y\| \leq \|x\| + \|y\|$ (Dreiecksungleichung).

Das Paar $(X, \| \ \|)$ heißt normierter Vektorraum. Gilt statt i) nur i') $\|x\| \geq 0$, so heißt $\| \ \|$ eine *Halbnorm*.

Ein normierter Vektorraum $(X, \| \ \|)$ wird durch $d(x, y) := \|x - y\|$ zu einem metrischen Raum (X, d).

Satz 1.10 *Die durch die Metrik definierte Topologie $\mathcal{T}$ macht $(X, \mathcal{T})$ zu einem topologischen Vektorraum.*

Beweis. Nach Dreiecksungleichung gilt $\|(x+y)-(x_0+y_0)\| \leq \|(x-x_0)\|+\|y-y_0\|$, d.h. die Bilder der $\frac{\epsilon}{3}$-Umgebungen von x_0, y_0 unter der Addition sind enthalten in der ϵ-Umgebung von $x_0 + y_0$. Die Addition ist sogar gleichmäßig stetig.

Für die Multiplikation findet man $\|\lambda x - \lambda_0 x_0 \ \| = \|\lambda(x - x_0) + (\lambda - \lambda_0)x_0 \ \| \leq |\lambda|\|x - x_0\| + |\lambda_0 - \lambda|\|x_0\|$, also (nicht gleichmäßige) Stetigkeit. □

Definition 1.11 Ein normierter Vektorraum $(X, \| \ \|)$, der bezüglich der kanonischen Metrik vollständig ist, heißt *Banach-Raum*.

Folgendes Vollständigkeitskriterium wird oft nützlich sein:

Lemma 1.12 *In einem normierten Vektorraum $(X, \| \ \|)$ sind äquivalent:*

i) *X ist vollständig.*

ii) *Für jede Folge (x_k) in X mit $\sum_{k=0}^{\infty} \|x_k\| < \infty$ gibt es ein $x \in X$ mit $\lim_{N\to\infty} \| \sum_{k=0}^{N} x_k - x\| = 0$, d.h. jede absolut konvergente Reihe ist konvergent.*

Beweis. i)⇒ii) Sei $\sum_{k=0}^{\infty} \|x_k\| < \infty$. Zu $\epsilon > 0$ gibt es ein $N \in \mathbb{N}$, so daß $\sum_{k=n}^{\infty} \|x_k\| < \epsilon$ für alle $n \geq N$. Dann gilt für $n \geq m \geq N$ die Abschätzung

$$\Big\|\Big(\sum_{k=0}^{n} x_k\Big) - \Big(\sum_{k=0}^{m} x_k\Big)\Big\| = \Big\| \sum_{k=m+1}^{n} x_k\Big\| \leq \sum_{k=m+1}^{n} \|x_k\| \leq \sum_{k=m+1}^{\infty} \|x_k\| < \epsilon \ ,$$

d.h. $\Big(\sum_{k=0}^{n} x_k\Big)_{n\in\mathbb{N}}$ ist Cauchy-Folge. Ist X vollständig, so konvergiert diese gegen ein $x \in X$.

ii)⇒i) Sei (x_n) eine Cauchy-Folge in X. Zu $\epsilon = \frac{1}{2^k}$ gibt es ein $N_k \in \mathbb{N}$, so daß $\|x_{m_k} - x_{n_k}\| < \frac{1}{2^k}$ für alle $n_k, m_k \geq N_k$. Insbesondere gibt es eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$

mit $\|x_{n_{k+1}} - x_{n_k}\| \leq \frac{1}{2^k}$. Für $y_k := x_{n_{k+1}} - x_{n_k}$ folgt dann $\sum_{k=0}^\infty \|y_k\| \leq 2 < \infty$. Nach Voraussetzung gibt es ein $y \in X$ mit

$$\lim_{N\to\infty} \Big\| \sum_{k=0}^N y_k - y \Big\| = \lim_{N\to\infty} \Big\| x_{n_{N+1}} - x_{n_0} - y \Big\| = 0 \ .$$

d.h. x_{n_k} ist konvergent gegen $x_{n_0} + y$. Ist $\|x_n - x_m\| < \epsilon$ für $n, m \geq N$, so gibt es auch ein $n_k \geq N$ mit $\|x_{n_k} - (x_{n_0} + y)\| < \epsilon$, so daß $\|x_n - (x_{n_0} + y)\| < 2\epsilon$. Damit ist auch x_n konvergent gegen $x_{n_0} + y$. □

Satz 1.13 *Seien X, Y normierte Vektorräume. Dann sind für eine lineare Abbildung $f : X \to Y$ äquivalent:*

i) *f ist stetig.*

ii) *f ist beschränkt, d.h. es gibt ein $C > 0$, so daß für alle $x \in X$ gilt $\|f(x)\| \leq C\|x\|$.*

Beweis. Da nach Satz 1.7 der Nachweis der Stetigkeit in 0 genügt, ist ii)⇒i) klar. Zu $\epsilon = 1$ existiert ein $\delta > 0$ mit $\|f(x)\| < 1$ für alle $\|x\| < \delta$. Ist $y \neq 0$, dann

$$\|f(y)\| = \Big\| f\Big(\frac{2\|y\|}{\delta} \underbrace{\frac{\delta}{2} \frac{y}{\|y\|}}_{=:x\ ,\ \|x\|<\delta} \Big) \Big\| \leq \frac{2\|y\|}{\delta} \|f(x)\| \leq \underbrace{\frac{2}{\delta}}_{=:C} \|y\| \ .$$

Für $y = 0$ ist nichts zu zeigen. Somit i)⇒ii). □

Satz/Definition 1.14 *Es seinen X, Y normierte Vektorräume. Dann ist $\mathcal{L}(X, Y)$ bezüglich der* Operatornorm

$$\|f\| := \inf\{C \ : \ \|f(x)\| \leq C\|x\|\} \ , \qquad f \in \mathcal{L}(X, Y)$$

ein normierter Vektorraum. Ist Y vollständig, so auch $(\mathcal{L}(X, Y), \|\ \|)$.

Offenbar gilt für die Operatornorm

$$\|f\| = \sup_{x\neq 0} \frac{\|f(x)\|}{\|x\|} = \sup_{\|x\|=1} \|f(x)\| = \sup_{\|x\|\leq 1} \|f(x)\| \qquad \text{sowie} \quad \|f(x)\| \leq \|f\|\|x\| \ .$$

Beweis von Satz 1.14. Die Normaxiome i) und ii) in Definition 1.9 folgen aus der Norm in Y, außerdem

$$\begin{aligned} \|f + g\| &= \sup_{\|x\|=1} \|(f + g)(x)\| \leq \sup_{\|x\|=1} (\|f(x)\| + \|g(x)\|) \\ &\leq \sup_{\|x_1\|=1} \|f(x_1)\| + \sup_{\|x_2\|=1} \|g(x_2)\| = \|f\| + \|g\| \ . \end{aligned}$$

Sei $(Y, \| \ \|)$ vollständig und $\{f_n\}$ eine Cauchy-Folge in $\mathcal{L}(X,Y)$. Für beliebiges $x \in X$ ist die Bildfolge $(f_n(x))$ wegen $\|f_n(x) - f_m(x)\| \leq \|f_n - f_m\|\|x\|$ eine Cauchy-Folge in Y, die wegen der Vollständigkeit einen Grenzwert hat, den wir $f(x)$ nennen: $\lim_{n\to\infty} f_n(x) =: f(x)$. Zu zeigen ist: Die so definierte Abbildung $f : X \to Y$ ist i) linear und ii) beschränkt:

i)
$$\begin{aligned} f(\lambda_1 x_1 + \lambda_2 x_2) &= \lim_{n\to\infty} \big(f_n(\lambda_1 x_1 + \lambda_2 x_2)\big) = \lim_{n\to\infty} \big(\lambda_1 f_n(x_1) + \lambda_2 f_n(x_2)\big) \\ &= \lambda_1 \lim_{n\to\infty} f_n(x_1) + \lambda_2 \lim_{n\to\infty} f_n(x_2) = \lambda_1 f(x_1) + \lambda_2 f(x_2)\ . \end{aligned}$$

ii) Zu $\epsilon > 0$ gibt es ein $N \in \mathbb{N}$ mit $\|f_m - f_n\| < \frac{\epsilon}{2}$ für für alle $m, n \geq N$. Zu beliebigem $x \in X$ mit $\|x\| \leq 1$ gibt es dann ein $m_x \geq N$ mit $\|f(x) - f_{m_x}(x)\| < \frac{\epsilon}{2}$. Nach Dreiecksungleichung folgt dann $\|f(x) - f_n(x)| < \epsilon$ für alle $n \geq N$, so daß

$$\|f\| = \sup_{\|x\|\leq 1} \|f(x)\| \leq \sup_{\|x\|\leq 1} \big(\|f(x) - f_N(x)\| + \|f_N(x)\|\big) \leq \epsilon + \|f_N\|$$

sowie
$$\|f - f_n\| = \sup_{\|x\|\leq 1} \|f(x) - f_n(x)\| \leq \epsilon\ .$$

Also ist f beschränkt, und (f_n) konvergiert in der Operatornorm gegen f. $\square$

Definition 1.15 Sei $(X, \| \ \|)$ normierter Vektorraum und $A \subset X$ eine Teilmenge. Dann heißt
$$d(x, A) := \inf_{y\in A} \|x - y\|$$
der Abstand eines Punktes $x \in X$ zu A.

Die Topologie metrischer Räumen liefert dann $d(x, A) = 0 \quad \Leftrightarrow \quad x \in \bar{A}$.

Satz 1.16 *Sei $(X, \| \ \|)$ ein normierter Vektorraum und $M \subset X$ ein Untervektorraum. Zu $x \in X$ sei $[x] = x + M \in X/M$ die Äquivalenzklasse. Dann gilt*

i) $\|[x]\| := d(x, M)$ *definiert eine Halbnorm auf X/M.*

ii) *Ist M abgeschlossen, so ist $x \mapsto \|[x]\|$ eine Norm.*

iii) *Die kanonische Abbildung $\iota : X \ni x \mapsto [x] \in X/M$ ist linear und beschränkt mit $\|\iota\| \leq 1$, also stetig. Sie bildet offene Mengen auf offene Mengen ab, d.h. ι ist offen.*

iv) *Ist X vollständig und M abgeschlossen, so ist X/M ein Banach-Raum.*

v) *Ist X vollständig und M abgeschlossen, dann ist M vollständig.*

vi) *Ist M vollständig, so ist M auch abgeschlossen.*

Beweis. i) M/X ist ein Vektoraum mit $[x]+[y]=[x+y]$ und $\lambda[x]=[\lambda x]$. Für $[x]=[y]$ ist $y=x+m$ für ein $m\in M$, also $d(x,M)=d(y,M)$. Somit ist $\|[x]\|$ wohldefiniert (d.h. unabhängig von der Wahl des Repräsentanten). Ist $\lambda\neq 0$, so gilt

$$\|\lambda[x]\| = d(\lambda x, M) = \inf_{m\in M}\|\lambda x-\lambda m\| = |\lambda|\big(\inf_{m\in M}\|x-m\|\big) = |\lambda|\;\|[x]\|\;.$$

Sei $\epsilon>0$ beliebig. Zu $x,y\in X$ gibt es $m_1,m_2\in M$ mit $\|x-m_1\|<\|[x]\|+\epsilon$ und $\|y-m_2\|<\|[y]\|+\epsilon$. Dann gilt

$$\|[x]+[y]\| = \|[x+y]\| \leq \|x+y-m_1-m_2\| \leq \|x-m_1\|+\|y-m_2\| \leq \|[x]\|+\|[y]\|+2\epsilon\;.$$

Da $\epsilon>0$ beliebig, folgt die Dreiecksungleichung.

ii) $\|[x]\|=0 \quad\Rightarrow\quad x\in\bar{M}=M \quad\Rightarrow\quad [x]=0$

iii) $\|[x]\|\leq\|x\|$ ist klar. Sei $V\subset X$ offen und $x\in V$. Dann enthält V auch eine ϵ-Umgebung $U_\epsilon(x)=\{y\in X \;:\; \|x-y\|<\epsilon\}\subset V$ von x. Wir zeigen $U_\epsilon([x])\subset V+M$, d.h. $V+M=\{[x]\;:\;x\in V\}=\iota(V)$ ist offen. Sei $[z]\in U_\epsilon([x])$ beliebig. Dann gibt es ein $m\in M$ mit $\|x-z+m\|<\epsilon$, denn das Gegenteil $\|x-z+m\|\geq\epsilon$ für alle $m\in M$ wäre im Widerspruch zu $\|[x]-[z]\|<\epsilon$. Also ist $z-m\in U_\epsilon(x)$ und $U_\epsilon([x])\subset\iota(V)$.

iv) Wir nutzen Lemma 1.12. Sei $\sum_{k=0}^\infty\|[x_k]\|<\infty$. Zu $[x_k]$ gibt es einen Repräsentanten $y_k=x_k+m_k$ mit $\|y_k\|<\|[x_k]\|+\frac{1}{2^k}$. Dann ist auch $\sum_{k=0}^\infty\|y_k\|<\infty$. Da X vollständig ist, konvergiert $\Big(\sum_{k=0}^n y_k\Big)_{n\in\mathbb{N}}$ gegen ein $x\in X$. Für die zugehörige Klasse $[x]\in X/M$ gilt

$$\Big\|[x]-\sum_{k=0}^n[x_k]\Big\| = \Big\|\Big[x-\sum_{k=0}^n y_k\Big]\Big\| \overset{\text{iii)}}{\leq} \Big\|x-\sum_{k=0}^n y_k\Big\| \to 0$$

Folglich ist $\Big(\sum_{k=0}^n[x_k]\Big)_{n\in\mathbb{N}}$ konvergent im (nach ii) normierten) Vektorraum X/M gegen $[x]\in X/M$. Nach Lemma 1.12 ist X/M vollständig.

v) und vi) folgen aus der Topologie metrischer Räume. □

2 Beispiele für normierte Vektorräume

Beispiel 2.1 (endlich-dimensionale Vektorräume) $\mathbb{K}^n\ni x=(x_1,\dots,x_n)$ ist normierter Vektorraum bezüglich

i) der p-Norm $\|x\|_p:=\sqrt[p]{|x_1|^p+\dots+|x_n|^p}$, für $1\leq p<\infty$,

ii) der Maximum-Norm $\|x\|_\infty:=\max_{i\in\{1,\dots,n\}}|x_i|$.

Die Dreiecksungleichung in $(\mathbb{K}^n,\|\;\|_p)$ heißt *Minkowskische Ungleichung.* Man beweist sie über die *Höldersche Ungleichung*

$$\Big|\sum_{i=1}^n x_iy_i\Big| \leq \|x\|_p\|y\|_q \qquad \text{falls } \frac{1}{p}+\frac{1}{q}=1\;.$$

Sie überträgt sich sinngemäß für $p = 1$, $q = \infty$. $\triangleleft$

Diese endlich-dimensionalen Beispiele sind eher uninteressant. In der Analysis wird bewiesen, daß je zwei beliebige Normen $\| \ \|_a$ und $\| \ \|_b$ auf einem endlich-dimensionalen Vektorraum X äquivalent sind, $c\|x\|_a \leq \|x\|_b \leq C\|x\|_b$ für alle $x \in X$ und universelle Konstanten $0 < c \leq C$. Außerdem ist jeder endlich-dimensionale normierte Vektorraum vollständig.

Beispiel 2.2 (beschränkte Funktionen) Sei M eine beliebige Menge und $\ell^\infty(M)$ die Menge der beschränkten Funktionen $f : M \to \mathbb{K}$. Mit der punktweisen Addition $(f_1 + f_2)(m) := f_1(m) + f_2(m)$ und $(\lambda f)(m) := \lambda f(m)$ wird $\ell^\infty(M)$ zu einem Vektorraum, auf dem wir die Supremums-Norm

$$\|f\|_\infty := \sup_{m \in M} |f(m)|$$

einführen. $\triangleleft$

Satz 2.3 $\big(\ell^\infty(M), \| \ \|_\infty\big)$ *ist ein Banach-Raum.*

Beweis. Die Normeigenschaften ergeben sich wie in Satz 1.14 für $\mathcal{L}(X, Y)$, wo die Vektorraum-Struktur von X gar nicht eingeht.

Zum Beweis der Vollständigkeit betrachte eine Cauchy-Folge (f_n) in $\ell^\infty(M)$. Dann gilt $|f_n(m) - f_k(m)| \leq \sup_{m \in M} |f_n(m) - f_k(m)| = \|f_n - f_k\|_\infty$ für beliebiges $m \in M$, so daß $(f_n(m))$ eine Cauchy-Folge in $\mathbb{K}$ ist, die punktweise einen Grenzwert $f(m) := \lim_{n \to \infty} f_n(m)$ besitzt. Analog zu Beweis in Satz 1.14 zeigt man die Beschränktheit von f. $\square$

Konvergenz einer Folge bezüglich der Supremums-Norm bedeutet *gleichmäßige Konvergenz* der Folge.

Satz 2.4 (stetige beschränkte Funktionen) *Sei X ein topologischer Raum und $\mathcal{C}_b(X) \subset \ell^\infty(X)$ der Untervektorraum der stetigen beschränkten Funktionen auf X. Dann ist $(\mathcal{C}_b(X), \| \ \|_\infty)$ ein Banach-Raum bezüglich der aus $\ell^\infty(X)$ geerbten Supremums-Norm.*

Beweis. Nach Satz 1.16.v) ist nur zu zeigen, daß $\mathcal{C}_b(X)$ abgeschlossen ist, d.h. daß für jede Folge stetiger beschränkter Funktionen, die in der Norm $\| \ \|_\infty$ gegen eine Funktion $f \in \ell^\infty(X)$ konvergiert, gilt $f \in C_b(X)$. Das ist ein Satz der Analysis: die Grenzfunktion einer gleichmäßig konvergenten Folge stetiger Funktionen ist stetig ($\frac{\epsilon}{3}$-Argument). $\square$

Für gutartige topologische Räume X kann man sich auf Untervektorräume von $\mathcal{C}_b(X)$ einschränken. Ist Y ein kompakter Hausdorff-Raum, dann ist jede stetige Funktion beschränkt, so daß $(\mathcal{C}(Y), \| \ \|)$ ein Banach-Raum wird. Ist Y ein lokal-kompakter Hausdorff-Raum (d.h. jeder Punkt $y \in Y$ besitzt eine kompakte Umgebung), dann ist

$$\mathcal{C}_0(Y) := \big\{ f \in \mathcal{C}(Y) \ : \ \text{für jedes } \epsilon > 0 \text{ ist } \{y \ : \ |f(y)| \geq \epsilon\} \text{ kompakt} \big\}$$

ein Banach-Raum bezüglich $\| \ \|$.

Beispiel 2.5 Auch auf dem Vektorraum der k-mal stetig differenzierbaren Funktionen $\mathcal{C}^k(T)$, mit $T \subset \mathbb{R}^n$ offen, ist $\| \ \|_\infty$ eine Norm, jedoch ist $(\mathcal{C}^k(T), \| \ \|)$ nicht mehr vollständig. Die Grenzfunktionen einer Folge differenzierbarer Funktionen ist i.a. nur stetig, nicht mehr differenzierbar. Eine andere Norm, bezüglich der $\mathcal{C}^k(T)$ ein Banach-Raum wird, wird als Übungsaufgabe untersucht. $\triangleleft$

Beispiel 2.6 (Folgenräume) Der Vektorraum der beschränkten Zahlenfolgen $(\ell^\infty := \ell^\infty(\mathbb{N}), \| \ \|_\infty)$ ist nach Satz 2.2 ein Banach-Raum. Die Untervektorräume ℓ_c der konvergenten Folgen und ℓ_0 der Nullfolgen sind abgeschlossen und damit ebenfalls Banach-Räume bezüglich $\| \ \|_\infty$. $\triangleleft$

Satz 2.7 (ℓ^p-Räume) *Für $1 \leq p < \infty$ sei*

$$\ell^p := \Big\{ x = (x_n)_{n\in\mathbb{N}} \ : \ \sum_{n=0}^{\infty} |x_n|^p < \infty \Big\}$$

der Vektorraum der p-summierbaren Zahlenfolgen. Dann ist ℓ^p ein Banach-Raum bezüglich der p-Norm

$$\|x\|_p := \Big(\sum_{n=0}^{\infty} |x_n|^p \Big)^{\frac{1}{p}} .$$

Beweis. Zu zeigen ist zunächst, daß ℓ^p ein Vektorraum ist:

$$\begin{aligned} \sum_{n=0}^{\infty} |x_n + y_n|^p &\leq \sum_{n=0}^{\infty} (|x_n| + |y_n|^p) \leq \sum_{n=0}^{\infty} (2\max(|x_n|, |y_n|))^p \\ &\leq 2^p \sum_{n=0}^{\infty} (|x_n|^p + |y_n|^p) < \infty . \end{aligned}$$

Die Dreiecksungleichung (Minkowskische Ungleichung) folgt wieder auf der Hölderschen Ungleichung (für Reihen statt Summen): Ist $x \in \ell^p$ und $y \in \ell^q$ mit $\frac{1}{p} + \frac{1}{q} = 1$, dann gilt für $xy = (x_n y_n)_{n\in\mathbb{N}}$

$$\|xy\|_1 \leq \|x\|_p \|y\|_q .$$

Sie überträgt sich sinngemäß auf $p = 1, q = \infty$.

Verbleibt die Vollständigkeit (verbunden mit einem Bezeichnungsproblem). Sei $(x_{(k)})_{k\in\mathbb{N}}$ eine Cauchy-Folge von Zahlenfolgen $x_{(k)} = (x_{(k)n})_{n\in\mathbb{N}} \in \ell^p$. Für das n-te Folgenglied gilt $|x_{(k)n}| \leq \|x_{(k)}\|_p$, d.h. für jedes n ist $(x_{(k)n})_{k\in\mathbb{N}}$ eine Cauchy-Folge in $\mathbb{K}$, die somit gegen eine Grenzfolge $x = (x_n)_{n\in\mathbb{N}}$ konvergiert, $x_n := \lim_{k\to\infty} x_{(k)n}$. Es verbleibt, wie üblich, zu zeigen, daß $x = (x_n) \in \ell^p$ mit $\|x - x_{(k)}\|_p \to 0$. Zu $\epsilon > 0$ gibt es ein $N \in \mathbb{N}$, so daß $\|x_{(k)} - x_{(l)}\|_p < \epsilon$ für alle $k, l \geq N$ (Voraussetzung der Cauchy-Folge). Insbesondere

$$\Big(\sum_{n=0}^{K} |x_{(k)n} - x_{(l)n}|^p\Big)^{\frac{1}{p}} < \epsilon \qquad \text{für alle } K \in \mathbb{N}\;,\; k,l \geq N\;.$$

Da für festes $n \in \mathbb{N}$ die Folge $(x_{(l)n})_{l\in\mathbb{N}}$ konvergent ist, ergibt sich im Limes $l \to \infty$

$$s_K := \lim_{l\to\infty} \Big(\sum_{n=0}^{K} |x_{(k)n} - x_{(l)n}|^p\Big)^{\frac{1}{p}} = \Big(\sum_{n=0}^{K} |x_{(k)n} - x_n|^p\Big)^{\frac{1}{p}} \leq \epsilon\;.$$

Für festes $k \geq N$ ist die Folge der Partialsummen (s_K) monoton und beschränkt, so daß im Limes $K \to \infty$ folgt $\|x_{(k)} - x\|_p \leq \epsilon$. $\square$

Die Verallgemeinerung der p-Normen auf Integrale (z.B. über Riemannsche Summen) ist mit Schwierigkeiten verbunden. Zwar ist $\mathcal{C}([a,b])$ ein normierter Vektorraum bezüglich des Riemann-Integrals

$$\|f\|_1 := \int_a^b dx\, |f(x)|\;,$$

aber *nicht* vollständig. Für denkbare Erweiterungen der Funktionsklasse ist dann $\|\;\|_1$ keine Norm mehr. Der Ausweg besteht im Übergang zum Lebesgue-Integral. Wir wiederholen die wichtigsten Schritte der Konstruktion:

- Sei $(X, \mathcal{T})$ ein topologischer Raum, dann erzeugt $\mathcal{T}$ die *Borel-Algebra* $\mathcal{B}(X)$ von X, d.h. das kleinste System von Teilmengen von X, welches $\mathcal{T}$ enthält und abgeschlossen unter Komplementbildung und abzählbaren Vereinigungen und Durchschnitten ist. Elemente $A \in \mathcal{B}(X)$ heißen *Borel-Mengen*.
- Ein *Maß* auf X ist eine Abbildung $\mu : \mathcal{B}(X) \to [0, \infty]$ mit $\mu(\varnothing) = 0$ und $\mu(\bigcup_{n=0}^{\infty} E_n) = \sum_{n=0}^{\infty} \mu(E_n)$ für paarweise disjunkte $E_n \in \mathcal{B}(X)$. $(X, \mathcal{B}(X), \mu)$ bzw. kurz (X, μ) heißt *Maßraum*. Auf $\mathbb{R}^n$ gibt es genau ein normiertes translationsinvariantes Maß, das Lebesgue-Maß λ.
- Eine Funktion $f : X \to \mathbb{K}$ heißt *meßbar*, falls das Urbild $f^{-1}(U)$ einer beliebigen offenen Teilmenge $U \subset \mathbb{K}$ Borel-Menge in X ist.
- Ist $A \in \mathcal{B}(X)$, dann heißt $1_A(x) = \begin{cases} 1 & \text{für } x \in A \\ 0 & \text{für } x \notin A \end{cases}$ die *charakteristische Funktion* von A; diese ist meßbar. Funktionen $\phi = \sum_{i=1}^{m} \phi_i 1_{A_i}$ mit $A_i \in \mathcal{B}(X)$ und $\phi_i \in \mathbb{K}$ heißen *Treppenfunktionen*. Ist μ ein Maß auf X und

$\mu(A_i) < \infty$, so heißt die Treppenfunktion $\phi = \sum_{i=1}^m \phi_i 1_{A_i}$ *endlich.* Das *Integral* einer endlichen Treppenfunktion (bezüglich μ) ist erklärt als

$$\int_X d\mu\phi := \sum_{i=1}^m \phi_i \mu(A_i) \ .$$

- Ist $f : X \to \mathbb{K}$ meßbar, dann heißt

  $$\|f\|_1 := \sup \left\{ \int_X d\mu \ \phi \ : \ \phi \text{ ist endliche Treppenfunktion mit } 0 \le \phi \le |f| \right\}$$

  die $\mathcal{L}^1$*-Halbnorm* von f. Eine meßbare Funktion $f : X \to \mathbb{K}$ heißt μ-integrierbar, falls es eine Folge $(\phi_n)_{n\in\mathbb{N}}$ endlicher Treppenfunktionen gibt mit $\lim_{n\to\infty} \|f - \phi_n\|_1 = 0$. Dann ist $\left(\int_X d\mu \ \phi_n \right)$ eine Cauchy-Folge, und

  $$\int_X d\mu \ f := \lim_{n\to\infty} \int_X d\mu \ \phi_n$$

  definiert das *(Lebesgue-) Integral* (bezüglich μ) von f.

- Eine Teilmenge $N \subset X$ heißt μ*-Nullmenge*, falls es zu jedem $\epsilon > 0$ eine Menge $A \subset \mathcal{B}(X)$ gibt mit $N \subset A$ und $\mu(A) < \epsilon$. Sei $\mathcal{N} := \{N \subset X \ : \ N \text{ ist Nullmenge}\}$. Aus technischen Gründen wird $\mathcal{B}(X)$ vervollständigt durch Hinzunahme aller Mengen $B \subset X$, für die es $A \in \mathcal{B}(X)$ und Nullmengen N_1, N_2 gibt mit $A \setminus N_1 \subset B \subset A \cup N_2$. Man definiert dann $\mu(B) := \mu(A)$. Eine Eigenschaft E gilt μ*-fast-überall* auf X, wenn die Menge der Punkte, für die E nicht gilt, eine μ-Nullmenge ist.

Man zeigt folgende Eigenschaften:

i) Für f, g meßbar sind auch $f + g$, fg, $|f|$, $|f|^p$, λf sowie (falls definiert) f/g, $\max(f, g)$ und $\min(f, g)$ meßbar. Ist f_n meßbar für alle $n \in \mathbb{N}$ und existiert $\lim_{n\to\infty} f_n(x)$ für alle $x \in X$, so ist f meßbar.

ii) Für eine meßbare Funktion $f : X \to \mathbb{K}$ gilt:

$$f \text{ integrierbar} \quad \Leftrightarrow \quad |f| \text{ integrierbar} \quad \Leftrightarrow \quad \|f\|_1 < \infty \ ,$$

und dann gilt $\left| \int_X d\mu \ f \right| \le \int_X d\mu \ |f| = \|f\|_1$.

iii) Das Integral ist linear und monoton: $f \le g \quad \Rightarrow \quad \int_X d\mu \ f \le \int_X d\mu \ g.$

iv) (Majorantenkriterium) Ist $f : X \to \mathbb{K}$ meßbar, $g : X \to \mathbb{R}_+$ integrierbar mit $|f| \le g$, dann ist auch f integrierbar, mit $\left| \int_X d\mu \ f \right| \le \int_X d\mu \ g.$

v) (monotone Konvergenz – Beppo Levi) Ist (f_n) monoton wachsende Folge integrierbarer Funktionen $f_n : X \to \mathbb{R}$ mit beschränkter Folge der Integrale, dann ist die punktweise gebildete Grenzfunktion $f(x) := \lim_{n\to\infty} f_n(x)$ meßbar und integrierbar, mit $\displaystyle\int_X d\mu\, f = \lim_{n\to\infty} \int_X d\mu\, f_n$.

vi) (dominierte Konvergenz – Lebesgue) Sei (f_n) Folge integrierbarer Funktionen mit $f_n(x) \to f(x)$ μ-fast-überall. Es gebe eine integrierbare Funktion g mit $|f_n(x)| \leq g(x)$ μ-fast-überall. Dann ist f integrierbar, und es gilt $\displaystyle\int_X d\mu\, f = \lim_{n\to\infty} \int_X d\mu\, f_n$.

vii) $\|f\|_1 = 0 \quad\Leftrightarrow\quad f(x) = 0$ μ-fast-überall,
$f(x) = g(x)$ μ-fast-überall $\quad\Rightarrow\quad \displaystyle\int_X d\mu\, f = \int_X d\mu\, g$.

viii) Ist $f : X \to \mathbb{K}$ integrierbar, dann ist $|f(x)| < \infty$ μ-fast-überall.

Definition 2.8 Sei (X,μ) ein Maßraum. Für $1 \leq p < \infty$ werde der folgende Funktionenraum definiert:

$$\mathcal{L}^p(X,\mu) := \{ f : X \to \mathbb{K} \;:\; f \text{ ist meßbar und } \int_X d\mu\, |f|^p < \infty \} .$$

Für $p = 1$ fällt diese Definition mit jener der μ-integrierbaren Funktionen zusammen.

Satz 2.9 *$\mathcal{L}^p(X,\mu)$ ist ein Vektorraum, auf dem durch $\|f\|_p := \Big(\int_X d\mu\ |f|^p\Big)^{\frac{1}{p}}$ für $f \in \mathcal{L}^p(X,\mu)$ eine Halbnorm definiert wird.*

Mit f, g sind auch $|f|^p, f+g$ meßbar. Dann folgt die Vektorraum-Struktur von $\mathcal{L}^p(X,\mu)$ in Analogie zu Satz 2.7. Der weitere Beweis wird in Teilschritten aus eigenen Sätzen geführt.

Satz 2.10 (Höldersche Ungleichung) *Sei $1 < p, q < \infty$ mit $\frac{1}{p} + \frac{1}{q} = 1$. Sind $f \in \mathcal{L}^p(X,\mu)$ und $g \in \mathcal{L}^q(X,\mu)$, dann ist $fg \in \mathcal{L}^1(X,\mu)$, und es gilt*

$$\|fg\|_1 \leq \|f\|_p \|g\|_q .$$

Beweis. Konvexität der Funktion $-\ln x$ liefert die Ungleichung $a^{\frac{1}{p}} \cdot b^{\frac{1}{q}} \leq \frac{a}{p} + \frac{b}{q}$ für alle $a, b \geq 0$. Wir können $\|f\|_p \neq 0$ und $\|g\|_p \neq 0$ annehmen; ansonsten ist $|f|^p = 0$ oder $|g|^q = 0$ μ-fast-überall, damit $f = 0$ oder $g = 0$ μ-fast-überall, somit $fg = 0$ μ-fast-überall. Außerhalb einer μ-Nullmenge N ist $|f|, |g| < \infty$. Für $x \in X \setminus N$ setze $a = \frac{|f|^p(x)}{\|f\|_p^p}$ und $b = \frac{|g|^q(x)}{\|g\|_q^q}$, dann folgt

$$\frac{|fg|(x)}{\|f\|_p \|g\|_q} \leq \frac{1}{p}\frac{|f|^p(x)}{\|f\|_p^p} + \frac{1}{q}\frac{|g|^q(x)}{\|g\|_q^q} .$$

Alle auftretenden Funktionen sind meßbar, und das μ-Integral liefert

$$\frac{1}{\|f\|_p\|g\|_q}\int_X d\mu\;|fg| \leq \frac{1}{p}+\frac{1}{q}=1\;.$$

Die Behauptung folgt dann aus dem Majorantenkriterium. $\square$

Satz 2.11 (Minkowskische Ungleichung) *Es gilt* $\|f+g\|_p \leq \|f\|_p+\|g\|_p$ *für alle* $f,g\in\mathcal{L}^p(X,\mu)$.

Beweis. Für $p=1$ folgt die Dreiecksungleichung aus der Monotonie des Integrals. Sei also $1<p<\infty$ und q definiert durch $\frac{1}{p}+\frac{1}{q}=1$, d.h. $pq=p+q$. Setze $h:=|f+g|^{p-1}$, dann ist $|h|^q=|f+g|^p$ und deshalb $h\in\mathcal{L}^q(X,\mu)$ sowie $\|h\|_q^q=\|f+g\|_p^p$. Damit gilt

$$\begin{aligned}\|f+g\|_p^p=\int_X d\mu\;|f+g|^p=\int_X d\mu\;|h||f+g|\leq\int_X d\mu\;|h||f|+\int_X d\mu\;|h||g|\\ \leq\|h\|_q\|f\|_p+\|h\|_q\|g\|_p=\|f+g\|_p^{\frac{p}{q}}\big(\|f\|_p+\|g\|_p\big)\;.\end{aligned}$$

Für $\|f+g\|_p=0$ ist nichts zu zeigen, ansonsten folgt die Behauptung aus $p-\frac{p}{q}=1$. $\square$

Bemerkung 2.12 Ist $f\in\mathcal{L}^p(X,\mu)$ mit $\|f\|_p=0$, dann ist $|f|^p=0$ μ-fast-überall, somit $f=0$ μ-fast-überall:

$$\|f\|_p=0\quad\Leftrightarrow\quad f\in\mathcal{N}(X,\mu):=\{g:X\to\mathbb{K}\text{ meßbar}, g=0\;\mu\text{-fast-überall}\}$$

Insbesondere ist $\mathcal{N}(X,\mu)\subset\mathcal{L}^p(X,\mu)$ Untervektorraum für alle $1\leq p<\infty$.

Folgerung 2.13 *Für* $1\leq p<\infty$ *ist* $L^p(X,\mu):=\mathcal{L}^p(X,\mu)/\mathcal{N}(X,\mu)$ *ein normierter Vektorraum bezüglich* $\|[f]\|_p:=\|f\|_p$.

Beweis. Sind $f\in\mathcal{L}^p(X,\mu)$ und $h\in\mathcal{N}(X,\mu)$, dann ist $\|f\|_p=\|f+h\|_p$ nach Modifikationssatz, die Norm auf $L^p(X,\mu)$ somit wohldefiniert. $\square$

Satz 2.14 (Riesz-Fischer) *$(L^p(X,\mu),\|\;\|_p)$ ist ein Banach-Raum.*

Beweis. Zu einer Cauchy-Folge $([f_n])$ in $L^p(X,\mu)$ seien beliebige Repräsentanten $f_n\in\mathcal{L}^p(X,\mu)$ gewählt. Dann ist auch (f_n) Cauchy-Folge in $\mathcal{L}^p(X,\mu)$ im Sinne von $\|f_n-f_m\|_p<\epsilon$ für $n,m\geq N_0$. Wegen $|\|f_n\|_p-\|f_m\|_p|\leq\|f_n-f_m\|_p$ ist dann auch $(\|f_n\|_p)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $\mathbb{R}$. Insbesondere ist $\|f_n\|_p\leq M$ beschränkt für alle n, und damit ist $|f_n(x)|<\infty$ $\mu-$fast-überall. Die Vereinigung abzählbar vieler Nullmengen ist Nullmenge, somit gibt es eine gemeinsame μ-Nullmenge N mit $|f_n(x)|<\infty$ für alle $x\in X\setminus N$ und alle $n\in\mathbb{N}$.

Wir zeigen: Es gibt es eine Funktion $f\in\mathcal{L}^p(X,\mu)$ mit

i) $\lim_{n\to\infty} \|f_n - f\|_p = 0$

ii) Eine Teilfolge $(f_{n_k})_{k\in\mathbb{N}}$ konvergiert μ-fast-überall gegen f.

Nach Übergang zu Äquivalenzklassen konvergiert $([f_n])$ dann in der $\| \ \|_p$-Norm gegen $[f] \in L^p(X,\mu)$.

Es gibt eine monoton wachsende Folge $(n_k)_{k\in\mathbb{N}}$ natürlicher Zahlen mit $\|f_{m_k} - f_{n_k}\|_p < \frac{1}{2^k}$ für alle $m_k \geq n_k$. Insbesondere ist $\|f_{n_{k+1}} - f_{n_k}\|_p < \frac{1}{2^k}$ für alle $k \in \mathbb{N}$. Sei $g_k := f_{n_{k+1}} - f_{n_k}$ außerhalb von N und $g_k(x) := 0$ für $x \in N$, ferner $G_n(x) := \sum_{k=1}^n |g_k(x)|$. Da $G_n \in \mathcal{L}^p(X,\mu)$, gilt

$$\int_X d\mu\ (G_n(x))^p = \|G_n\|_p^p \leq \Big(\sum_{k=1}^n \|g_k\|_p\Big)^p \leq 1 \qquad \text{für alle } n \in \mathbb{N}\ .$$

Nach dem Satz von Beppo Levi ist die punktweise gebildete Grenzfunktion $H(x) := \lim_{n\to\infty}(G_n(x))^p$ meßbar und integrierbar. Somit ist $G(x) := \lim_{n\to\infty} G_n(x)$ μ-fast-überall endlich. Nach Majorantenkriterium in $\mathbb{K}$ existiert außerhalb einer μ-Nullmenge (wieder bezeichnet mit) N der Limes $g(x) := \sum_{k=1}^\infty g_k(x)$ mit absoluter Konvergenz der Reihe. Durch $g(x) := 0$ für $x \in N$ wird g auf ganz X fortgesetzt. Die Borel-Algebra wird als vollständig vorausgesetzt, so daß mit g_k und ihrer Folge der Partialsummen auch g meßbar ist.

Die Folge der Partialsummen $s_n := \Big|\sum_{k=1}^n g_k\Big|^p$ konvergiert μ-fast-überall (außerhalb N) gegen die meßbare Funktion $|g|^p$. Außerdem ist

$$s_n^{\frac{1}{p}} = \Big|\sum_{k=1}^n g_k\Big| \leq \sum_{k=1}^n |g_k| = G_n \qquad \Rightarrow \qquad s_n \leq G_n^p \leq H$$

betragsmäßig beschränkt durch eine integrierbare Funktion H. Nach dem Satz von Lebesgue ist dann $|g|^p$ integrierbar, also $g \in \mathcal{L}^p(X,\mu)$. Definieren wir $f := g + f_{n_1} \in \mathcal{L}^p(X,\mu)$, dann gilt außerhalb der μ-Nullmenge N die Beziehung ii) $f = \lim_{k\to\infty}(g_k + f_{n_1}) = \lim_{k\to\infty} f_{n_{k+1}}$.

Schließlich konvergiert die Folge $(|f - f_{n_k}|^p)_{k\in\mathbb{N}}$ für $k \to \infty$ μ-fast-überall gegen Null. Außerhalb von N gilt für alle $l \geq k$

$$|f_{n_{l+1}} - f_{n_k}| = \Big|\sum_{i=k}^l g_i\Big| \leq \sum_{i=1}^l |g_i| = G_l \qquad \Rightarrow \qquad |f - f_{n_k}|^p \leq \lim_{l\to\infty} G_l^p = H\ .$$

Somit ist $(|f - f_{n_k}|^p)_{k\in\mathbb{N}}$ eine Folge integrierbarer Funktionen, die μ-fast-überall gegen die Funktion 0 konvergiert und μ-fast-überall durch eine integrierbare Funktion H beschränkt ist. Nach dem Satz von Lebesgue ist dann

$$\lim_{k\to\infty}\int_X d\mu |f(x) - f_{n_k}(x)|^p = \int_X d\mu\Big(\lim_{k\to\infty}|f(x) - f_{n_k}(x)|^p\Big) = 0\ .$$

Somit $\lim_{k\to\infty}\|f - f_{n_k}\|_p = 0$ und dann auch $\lim_{n\to\infty}\|f - f_n\|_p = 0$ nach Dreiecksungleichung. $\square$

Bemerkung 2.15 Die Betrachtung von Teilfolgen im Beweis ist wichtig. Da $\| \, \|_p$ keine Norm auf $\mathcal{L}^p(X,\mu)$ ist, folgt aus der Konvergenz einer Teilfolge einer Cauchy-Folge nicht die Konvergenz der Cauchy-Folge.

In Analogie zu den Folgenräumen ℓ^p, ℓ^∞ gibt es auch den Funktionenraum $L^\infty(X,\mu)$:

Definition 2.16 Eine meßbare Funktion $f : X \to \mathbb{K}$ heißt *μ-fast-überall beschränkt*, falls es ein $C > 0$ gibt mit $|f| \leq C$ μ-fast-überall. Dann sei

$$\mathcal{L}^\infty(X,\mu) := \{f : X \to \mathbb{K} \text{ meßbar und } \mu\text{-fast-überall beschränkt}\} .$$

Satz 2.17 *Sei (X,μ) ein Maßraum. Dann gilt:*

i) *$\mathcal{L}^\infty(X,\mu)$ ist ein Vektorraum, auf dem durch $\|f\|_\infty := \inf\{C : |f| \leq C$ μ-fast-überall$\}$ eine Halbnorm definiert wird.*

ii) *$L^\infty(X,\mu) := \mathcal{L}^\infty(X,\mu)/\mathcal{N}(X,\mu)$, mit $\mathcal{N}(X,\mu)$ aus Bemerkung 2.12, ist bezüglich $\| \, \|_\infty$ ein Banach-Raum.*

iii) *Für $f \in \mathcal{L}^\infty(X,\mu)$ und $g \in \mathcal{L}^1(X,\mu)$ ist $fg \in \mathcal{L}^1(X,\mu)$, und es gilt $\|fg\|_1 \leq \|f\|_\infty \|g\|_1$.*

Beweis. Übungsaufgabe. □

Die Funktionenräume $L^p(X,\mu)$ liefern wichtige und vor allem interessante Beispiele für unendlich-dimensionale normierte Vektorräume. Wir werden die Strukturen der Funktionalanalysis oft an den L^p-Räumen demonstrieren. Eine wichtige Tatsache folgt bereits aus der Hölderschen Ungleichung:

Folgerung 2.18 *Sei (X,μ) ein Maßraum, $1 \leq p \leq \infty$ und $\frac{1}{p} + \frac{1}{q} = 1$. Dann definiert jedes $g \in L^q(X,\mu)$ ein lineares beschränktes Funktional $S_g \in (L^p(X,\mu))'$ auf $L^p(X,\mu)$ durch*

$$S_g(f) := \int_X d\mu \, fg \,, \qquad f \in L^p(X,\mu) \,.$$

Für die Norm im Dualraum gilt $\|S_g\| \leq \|g\|_q$.

Bis auf exotische Beispiele ist $L^q(X,\mu) \neq \{0\}$, so daß zumindest für die L^p-Räume der Dualraum interessant ist. Ein exotisches Beispiel ist das folgende:

Beispiel 2.19 Betrachte $(X,\mathcal{T})$ mit $\mathcal{T} = \{\varnothing, X\}$ und $\mu(\varnothing) = 0$, $\mu(X) = \infty$. Dann ist $\mathcal{B}(X) = \{\varnothing, X\}$, und eine Funktion $f : X \to \mathbb{K}$ ist meßbar genau dann, wenn sie konstant ist (da $f^{-1}(\{\lambda\}) \in \{\varnothing, X\}$ für $\lambda \in \mathbb{K}$). Es folgt $L^p(X,\mu) = \{0\}$ für $1 \leq p < \infty$ und $L^\infty(X,\mu) = \mathbb{K}$. ◁

Für $1 < p \leq \infty$ läßt sich sogar mehr zeigen:

Satz 2.20 *Sei* $1 < p \leq \infty$. *Dann ist die in Folgerung 2.18 eingeführte Abbildung* $S : L^q(X,\mu) \ni g \mapsto S_g \in (L^p(X,\mu))'$ *isometrisch, d.h.* $\|S_g\| = \|g\|_q$.

Beweis. Wir können $\|g\|_q = 1$ annehmen. Sei $g \in \mathcal{L}^q(X,\mu)$ ein beliebiger Repräsentant. Dann ist $|g(x)| < \infty$ außerhalb einer μ-Nullmenge N. Betrachte

$$f(x) := \begin{cases} \frac{|g(x)|^q}{g(x)} & \text{für } |g(x)| \notin \{0,\infty\} \\ 0 & \text{für } |g(x)| \in \{0,\infty\} \end{cases}$$

Dann ist $f \in \mathcal{L}^p(X,\mu)$, was klar ist für $(p=\infty, q=0)$ und für $1 < p < \infty$ wie folgt zu sehen ist: Für $|g(x)| \notin \{0,\infty\}$ gilt $|f(x)|^p = |g(x)|^{qp-p} = |g(x)|^q$, was sich auch in Punkte mit $g(x) = 0$ fortsetzt. Insbesondere gilt $\|f\|_p = 1$, und es folgt

$$S_g(f) = \int_X d\mu\; fg = \int_X d\mu\; |g(x)|^q = 1\;.$$

Somit $\|S_g\| = \sup_{\|f\|_p=1} |S_g(f)| \geq 1$. □

Bemerkung 2.21 Wir werden im Laufe der Vorlesung zeigen, daß für $1 < p < \infty$ sogar gilt $(L^p(X,\mu))' = L^q(X,\mu)$ (Rieszscher Darstellungssatz). Dazu ist die Surjektivität von S zu beweisen (Injektivität folgt bereits aus $\|S_g\| = \|g\|$), was sich aus einem Konvexitätsargument ergeben wird. Insbesondere gilt $(L^p(X,\mu))'' = L^p(X,\mu)$ für $1 < p < \infty$, d.h. $L^p(X,\mu)$ ist reflexiv. Für $p = 1$ gilt Satz 2.20 i.a. nicht; kann aber gezeigt werden für σ-endliche Maßräume (es gibt eine Folge (A_n) von Teilmengen $A_n \subset X$ mit $\mu(A_n) < \infty$ und $\bigcup_{n=0}^\infty A_n = X$). In diesem Fall läßt sich auch $(L^1(X,\mu))' = L^\infty(X,\mu)$ zeigen. Dagegen ist üblicherweise $(L^\infty(X,\mu))'$ viel größer als $L^1(X,\mu)$.

3 Hahn-Banach-Sätze

Die Hahn-Banach-Sätze liefern wichtige Informationen über den Dualraum X' eines Banach-Raums. Sie zeigen, daß X und X' eng verknüpft sind. Das wird unter anderem genutzt, um schwache Topologien auf X' einzführen. Der Einheitsball in X' ist kompakt in der schwachen Topologie, falls $X = X''$ reflexiv ist.

Der Beweis des ersten Hahn-Banach-Satzes basiert auf einem Induktionsargument, für das das Lemma von Zorn benötigt wird.

Definition 3.1 Eine Menge $M \neq \varnothing$ heißt *geordnet*, falls eine Relation $\leq$ auf M existiert mit

i) $x \leq x$ für alle $x \in M$

ii) $x \leq y$ und $y \leq x \quad\Rightarrow\quad x = y$

iii) $x \leq y$ und $y \leq z \quad\Rightarrow\quad x \leq z$.

Ist M geordnet, so heißt eine Teilmenge $K \subset M$ eine *Kette*, falls für alle $x, y \in K$ gilt: $x \leq y$ oder $y \leq x$. Ist $K \subset M$ eine Kette, so heißt $z \in M$ eine *obere Schranke* von K, falls $x \leq z$ für alle $x \in K$. Ein Element $z \in M$ heißt *maximales Element* von M, falls für alle $x \in M$ gilt: $(x \geq z \quad \Rightarrow \quad x = z)$.

Lemma 3.2 (Zorn) *Sei $M \neq \varnothing$ eine geordnete Menge, so daß jede Kette $K \subset M$ eine obere Schranke besitzt. Dann besitzt M ein maximales Element.*

Zu beachten ist, daß ein maximales Element kein größtes Element sein muß.

Definition 3.3 Sei X ein $\mathbb{K}$-Vektorraum. Eine Abbildung $p : X \to \mathbb{R}$ heißt *sublinear*, falls

i) $p(\lambda x) = \lambda p(x) \qquad$ für alle $\lambda \geq 0$ und $x \in X$,

ii) $p(x + y) \leq p(x) + p(y) \qquad$ für alle $x, y \in X$.

Im Gegensatz zur (Halb-)Norm wird keine Positivität gefordert.

Beispiel 3.4 Folgende Abbildungen sind sublinear

i) Jede Halbnorm auf X.

ii) $X = \mathbb{R}$ mit $p(x) = \begin{cases} a_1 x & \text{für } x \geq 0 \\ a_2 x & \text{für } x \leq 0 \end{cases} \quad$ mit $a_1 \geq a_2$.

iii) $X = \ell^\infty$ mit $p((x_n)) = \text{limsup}_n \text{Re}(x_n)$.

Satz 3.5 (Hahn-Banach: I. reelle Vektorräume) *Sei X ein reeller Vektorraum, $U \subset X$ ein Untervektorraum und $p : X \to \mathbb{R}$ sublinear. Dann gilt: Zu jeder linearen Abbildung $f : U \to \mathbb{R}$ mit $f(x) \leq p(x)$ für alle $x \in U$ existiert eine lineare Fortsetzung*

$$F : X \to \mathbb{R} \qquad \textit{mit} \quad F\Big|_U = f \ \textit{und} \ F(x) \leq p(x) \quad \forall x \in X .$$

Beweis. Durch Induktion nach $n = \dim(X/U)$ unter Verwendung des Lemmas von Zorn.

i) Sei $n = 1$ und $x_0 \in X/U$. Jedes $x \in X$ hat eine eindeutige Darstellung $x = u + \lambda x_0$ mit $u \in U$ und $\lambda \in \mathbb{R}$. Für jedes $r \in \mathbb{R}$ wird durch $F(x) = f(u) + \lambda r$ eine lineare Abbildung $F : X \to \mathbb{R}$ definiert. Wir zeigen: Es gibt eine Wahl von r, so daß $F(x) \leq p(x)$, d.h.

$$f(u) + \lambda r \leq p(u + \lambda x_0) . \tag{*}$$

Die Gleichung (*) ist für $\lambda = 0$ automatisch erfüllt. Je nach Vorzeichen von $\lambda \neq 0$ erhalten wir aus der Sublinearität

$$\lambda > 0 : \qquad r \overset{!}{\leq} p\Big(\frac{u}{\lambda} + x_0\Big) - f\Big(\frac{u}{\lambda}\Big)$$

$$\lambda < 0 : \qquad -r \overset{!}{\leq} p\Big(\frac{u}{-\lambda} - x_0\Big) - f\Big(\frac{u}{-\lambda}\Big)$$

und damit

$$\sup_{v\in U}\big(f(v)-p(v-x_0)\big) \overset{!}{\leq} r \overset{!}{\leq} \inf_{w\in U}\big(p(w+x_0)-f(w)\big) .$$

Diese Ungleichung besitzt Lösungen $r \in \mathbb{R}$ wegen

$$f(v)+f(w) = f(v+w) \leq p(v+w) = p(v-x_0+w+x_0) \leq p(v-x_0)+p(w+x_0) .$$

Somit gibt es für $\dim(X/U) = 1$ eine Fortsetzung F mit den geforderten Eigenschaften.

ii) Induktiv läßt sich f fortsetzen auf Untervektorräume $U \subset V_n \subset X$ mit $\dim(V_n/U) = n$. Unter Verwendung des Lemmas von Zorn wird nun gezeigt, daß X erreicht wird. Dazu sei

$$\begin{aligned} M := \{(V,F_V)\ &: V \subset X \text{ Untervektorraum mit } U \subset V, \\ &F_V : V \to \mathbb{R} \text{ linear mit } F_V \leq p|_V\ , F_V\big|_U = f\} . \end{aligned}$$

Wegen $(U,f) \in M$ ist $M \neq \varnothing$. Wir definieren auf M eine Relation $\leq$ durch

$$(V_1,F_{V_1}) \leq (V_2,F_{V_2}) \qquad \Leftrightarrow \qquad V_1 \subset V_2 \text{ und } F_{V_2}\Big|_{V_1} = F_{V_1} .$$

Jede Kette $K = (V_i, L_{V_i})_{i\in I}$ besitzt eine obere Schranke (V_K, L_{V_K}) mit

$$V_K := \bigcup_{i\in I} V_i\ , \qquad F_{V_K}(x) = F_{V_i}(x) \text{ für } x \in V_i .$$

Somit gibt es ein maximales Element $m = (X_0, F_{X_0})$. Wäre $X_0 \neq X$, so gibt es ein $x_0 \in X/X_0$, und nach i) ließe sich m fortsetzen, im Widerspruch zur Maximalität von m. Also ist $X = X_0$ und $F = F_{X_0}$. □

Zu beachen ist, daß die Ungleichung für r im allgemeinen viele Lösungen hat, so daß auch die Fortsetzung F nicht eindeutig ist.

Satz 3.6 (Hahn-Banach: II. komplexe Vektorräume) *Sei X ein komplexer Vektorraum, $U \subset X$ ein Untervektorraum und $p : X \to \mathbb{R}$ sublinear. Dann gilt: Zu jeder linearen Abbildung $f : U \to \mathbb{C}$ mit $\mathrm{Re}(f(x)) \leq p(x)$ für alle $x \in U$ existiert eine lineare Fortsetzung*

$$F : X \to \mathbb{C} \qquad \textit{mit} \quad F\Big|_U = f \textit{ und } \mathrm{Re}(F(x)) \leq p(x) \quad \forall x \in X .$$

Beweis. Fassen wir X als reellen Vektorraum auf, dann gibt es nach Satz 3.5 ein $\mathbb{R}$-lineares Funktional $\tilde{F} : X \to \mathbb{R}$, welches $\mathrm{Re}(f)$ fortsetzt. Setze $F(x) = \tilde{F}(x) - \mathrm{i}\tilde{F}(\mathrm{i}x)$, dann ist $\mathrm{Re}\, F = \tilde{F}$, und F ist $\mathbb{C}$-linear:

$$\begin{aligned} F((\lambda+\mathrm{i}\mu)x) &= \tilde{F}((\lambda+\mathrm{i}\mu)x) - \mathrm{i}\tilde{F}(\mathrm{i}(\lambda+\mathrm{i}\mu)x) \\ &= \lambda\tilde{F}(x) + \mu\tilde{F}(\mathrm{i}x) - \mathrm{i}\lambda\tilde{F}(\mathrm{i}x) + \mathrm{i}\mu\tilde{F}(x) = (\lambda+\mathrm{i}\mu)F(x) . \end{aligned}$$

Ebenso ist $f = \mathrm{Re}(f) - \mathrm{i}\,\mathrm{Re}(\mathrm{i}f)$ eindeutig durch $\mathrm{Re}(f)$ definiert, so daß F die behaupteten Eigenschaften bezüglich f hat. $\square$

Theorem 3.7 (Hahn-Banach: III. Fortsetzungssatz) *Sei $(X, \|\ \|)$ ein normierter Vektorraum und $U \subset X$ ein Untervektorraum. Dann gilt: Zu jedem linearen beschränkten Funktional $u' : U \to \mathbb{K}$ gibt es ein lineares beschränktes Funktional $x' : X \to \mathbb{K}$ mit $x'\Big|_U = u'$ und $\|x'\| = \|u'\|$.*

Beweis. i) Sei X reeller Vektorraum. Die Abbildung $p(x) := \|u'\|\|x\|$ ist sublinear, so daß es nach Satz 3.5 eine lineare Abbildung $x' : X \to \mathbb{R}$ gibt mit $x'(x) \leq p(x)$ und $x'\Big|_U = u'$. Andererseits gilt $x'(-x) = -x'(x) \leq p(-x) = p(x)$ und somit $|x'(x)| \leq \|u'\|\|x\|$ für alle $x \in X$. Insbesondere ist $\|x'\| \leq \|u'\|$. Die Gleichheit folgt aus

$$\|u'\| = \sup_{u\in U\,,\ \|u\|\leq 1} |u'(u)\| = \sup_{u\in U\,,\ \|u\|\leq 1} |x'(u)\| \leq \sup_{x\in X\,,\ \|x\|\leq 1} |x'(x)\| = \|x'\|\ .$$

ii) Ist X komplexer Vektorraum, dann gibt es nach Satz 3.6 und i) ein $\mathbb{C}$-lineares Funktional $x' : X \to \mathbb{C}$ mit $x'\Big|_U = u'$ und $\|\mathrm{Re}(x')\| = \|u'\|$. Setzt man $x'(x) = e^{\mathrm{i}\alpha(x)}|x'(x)|$, dann ist

$$|x'(x)| = |x'(e^{-\mathrm{i}\alpha(x)}x)| = \big|(\mathrm{Re}\,x')(e^{-\mathrm{i}\alpha(x)}x)\big| \leq \|u'\|\|x\|\ .$$

Insgesamt folgt $\|x'\| = \|u'\|$. $\square$

Da endlich-dimensionale Untervektorräume $U \subset X$ lineare Funktionale $u' \neq 0$ besitzen, ist auch $X' \neq \{0\}$. Wir werden sehen, daß X' groß genug ist, um wichtige Eigenschaften aus X zu erfassen.

Folgerung 3.8 *Sei $(X, \|\ \|)$ normierter Vektorraum. Zu jedem $0 \neq x \in X$ gibt es ein $x' \in X'$ mit $\|x'\| = 1$ und $x'(x) = \|x\|$. Insbesondere trennt X' die Punkte aus X, d.h. zu $x_1, x_2 \in X$ mit $x_1 \neq x_2$ gibt es ein $x' \in X'$ mit $x'(x_1) \neq x'(x_2)$.*

Beweis. Betrachte $U = \mathbb{K}x$ und $u'(\lambda x) := \lambda\|x\|$. Dann ist $\|u'\| = 1$, und nach Theorem 3.7 gibt es eine norm-gleiche Fortsetzung $x' \in X'$. Die Trennung der Punkte ergibt sich durch Anwendung auf $x = x_1 - x_2 \neq 0$. $\square$

Bezeichnung 3.9 Ist $(Y, \|\ \|)$ ein normierter Vektorraum, dann bezeichnet $B_Y := \{y \in Y\ ,\ \|y\| \leq 1\}$ den *Einheitsball* in Y und $S_Y := \{y \in Y\ ,\ \|y\| = 1\}$ die *Einheitssphäre*.

Folgerung 3.10 *In einem normierten Vektorraum $(X, \|\ \|)$ gilt $\|x\| = \sup_{x'\in B_{X'}} |x'(x)|$. Das Supremum wird angenommen.*

Beweis. Einerseits ist $|x'(x)| \leq \|x'\|\|x\| \leq \|x\|$ für $x' \in B_{X'}$, andererseits gibt es nach Folgerung 3.8 ein $x' \in B_{X'}$ mit $|x'(x)| = \|x\|$. □

Folgerung 3.11 *Sei $(X, \| \|)$ normierter Vektorraum und $U \subset X$ ein abgeschlossener Untervektorraum. Dann gibt es zu $x \in X$, $x \notin U$, ein $x' \in X'$ mit $x'\big|_U = 0$ und $x'(x) \neq 0$.*

Beweis. Betrachte die Quotientenabblidung $\iota : X \to X/U$, die nach Satz 1.16 linear und stetig ist. Es gilt $\iota(u) = 0 \in X/U$ für alle $u \in U$ und $\iota(x) \neq 0 \in X/U$. Nach Satz 1.16 ist X/U ein normierter Vektorraum (da U abgeschlossen), und nach Folgerung 3.8 gibt es dann ein $f \in (X/U)'$ mit $f(\iota(x)) = f([x]) = \|[x]\| \neq 0$. Setze $x' := f \circ \iota$. □

Folgerung 3.12 *Sei $(X, \| \|)$ normierter Vektorraum, $U \subset X$ Untervektorraum. Dann sind äquivalent:*

i) *U ist dicht in X, d.h. $\bar{U} = X$.*

ii) *Falls $x' \in X'$ mit $x'\big|_U = 0$, so gilt $x' = 0$.*

Beweis. i)⇒ii) Zu jedem $x \in X$ gibt es eine gegen x konvergente Folge (u_n) in U. Aus der Stetigkeit von x' folgt $x'(x) = 0$.

ii)⇒i) Zunächst ist auch $x'\big|_{\bar{U}} = 0$. Wäre $\bar{U} \neq X$, so gibt es nach Folgerung 3.11 für $x \in X/\bar{U}$ ein $x' \in X$ mit $x'(x) \neq 0$, Widerspruch. □

Als Anwendung erarbeiten wir eine Eigenschaft der Normtopologie, die grundlegend verschieden von den im nächsten Abschnitt zu betrachtenden schwachen Topologien ist:

Satz 3.13 *Sei $(X, \| \|)$ ein normierter Vektorraum. Dann sind äquivalent:*

i) *B_X ist kompakt.*

ii) *X ist endlich-dimensional.*

Beweis. ii)⇒i) ist klar. Sei $B := B_X$ kompakt. Da $B \subset \bigcup_{x \in B} U_{\frac{1}{2}}(x)$ eine offene Überdeckung ist, gibt es $\{x_1, \ldots, x_l\} \in B$ mit $B \subset \bigcup_{i=1}^{l} U_{\frac{1}{2}}(x_i)$. Betrachte den von den x_i aufgespannten Untervektorraum $Y := \operatorname{span}(x_1, \ldots, x_l) \subset X$. Wir zeigen: $Y = X$. Es gilt $U_{\frac{1}{2}}(x_i) = x_i + U_{\frac{1}{2}}(0) \subset x_i + \frac{1}{2}B$; dabei ist rB der Ball in X vom Radius r. Somit gibt es zu jedem $b \in B$ ein $i \in \{1, \ldots, l\}$ mit $b \in x_i + \frac{1}{2}B \subset Y + \frac{1}{2}B$, und induktiv erhalten wir

$$B \subset Y + \tfrac{1}{2}B \subset Y + \tfrac{1}{2}(Y + \tfrac{1}{2}B) = Y + \tfrac{1}{4}B \subset \cdots \subset Y + \tfrac{1}{2^k}B \subset \bigcap_{k \in \mathbb{N}} (Y + \tfrac{1}{2^k}B)\ .$$

Da Punkte $x_0 \in \left(\bigcap_{k \in \mathbb{N}} (Y + \frac{1}{2^k}B) \right) \setminus Y$ gerade die Häufungspunkte von Y sind, ist $\bar{Y} = \bigcap_{k \in \mathbb{N}} (Y + \frac{1}{2^k}B)$ der Abschluß und $B \subset \bar{Y}$.

Angenommen, es gibt ein $x_0 \in \bar{Y} \setminus Y$. Betrachte dann den $(l+1)$-dimensionalen Vektorraum $Y_1 := Y + \mathbb{K}x_0$ sowie auf Y_1 das lineare stetige Funktional $f(y + \lambda x_0) := \lambda$, mit $y \in Y$. Dieses setzt sich nach Hahn-Banach zu einem Funktional $\bar{f} \in \bar{Y}'$ fort (nach Lemma 1.6 ist $\bar{Y}$ ein Vektorraum, und $Y_1 \subset \bar{Y}$ ist Untervektorraum) und erfüllt $\bar{f}(x_0) = 1$ sowie $\bar{f}\big|_Y = 0$. Nach Folgerung 3.12 ist dann $x_0 \notin \bar{Y}$, Widerspruch. Also gilt sogar $Y = \bar{Y}$ und damit $B \subset Y$. Dann ist auch $X = \bigcup_{\lambda > 0}(\lambda B) \subset Y$, d.h. X ist endlich-dimensional. □

Die Trennung von Punkten nach Folgerung 3.8 läßt sich auf geeignete Teilmengen erweitern.

Definition 3.14 Eine Teilmenge $A \subset X$ eines Vektorraums X heißt *konvex*, falls mit $a, b \in A$ auch $\lambda a + (1 - \lambda)b \in A$ für alle $\lambda \in [0, 1]$.

Definition 3.15 Es sei X ein topologischer Vektorraum und $U \subset X$ eine Umgebung von 0. Dann wird durch $p_U(x) := \inf\{\lambda > 0 \ : \ \frac{x}{\lambda} \in U\}$ das *Minkowski-Funktional* $p_U : X \to \mathbb{R}_+$ erklärt.

Lemma 3.16 *Sei $U \subset X$ eine konvexe Umgebung von 0. Dann gilt:*

i) *Das Minkowski-Funktional p_U ist sublinear.*

ii) *Ist U offen, so gilt $U = p_U^{-1}([0, 1[)$.*

Beweis. i) Die Definition liefert $p_U(\lambda x) = \lambda p_U(x)$ für $\lambda > 0$. Zu $x, y \in X$ und $\epsilon > 0$ gibt es $\lambda, \mu > 0$ mit $p_U(x) \leq \lambda \leq p_U(x) + \epsilon$ und $p_U(y) \leq \mu \leq p_U(y) + \epsilon$. Da U konvex ist, ist mit $\frac{x}{\lambda}, \frac{y}{\mu} \in U$ auch

$$\frac{\lambda}{\lambda + \mu}\frac{x}{\lambda} + \frac{\mu}{\lambda + \mu}\frac{y}{\mu} = \frac{x + y}{\lambda + \mu} \in U \ .$$

Somit ist $p(x + y) \leq \lambda + \mu \leq p(x) + p(y) + 2\epsilon$. Da $\epsilon > 0$ beliebig, ist p_U sublinear.

ii) Ist $p_U(x) < 1$, so gibt es ein $0 < \lambda < 1$ mit $\frac{x}{\lambda} \in U$, dann ist auch $x = \lambda\frac{x}{\lambda} + (1 - \lambda)0 \in U$. Ist umgekehrt $p_U(x) \geq 1$, dann ist $\frac{x}{\lambda} \notin U$ für alle $0 < \lambda < 1$. Sei (λ_n) eine gegen 1 konvergierende Folge mit $0 < \lambda_n < 1$, dann konvergiert $\frac{x}{\lambda_n} \in X \setminus U$ gegen x. Da $X \setminus U$ abgeschlossen, ist $x \in X \setminus U$. □

Theorem 3.17 (Hahn-Banach IV: Trennungssatz) *Sei $(X, \| \ \|)$ normierter Vektorraum, $V_1, V_2 \in X$ konvexe disjunkte Teilmengen, mit V_1 offen. Dann gibt es ein $x' \in X'$ mit $\operatorname{Re} x'(v_1) < \operatorname{Re} x'(v_2)$ für alle $v_1 \in V_1$ und $v_2 \in V_2$.*

Beweis. Wie üblich betrachten wir zunächst den Fall reeller Vektorräume. Sei $V := V_1 - V_2 = \{v_1 - v_2 \ : \ v_1 \in V_1 \ , \ v_2 \in V_2\}$. Dann ist V ebenfalls konvex, außerdem offen wegen $V = \bigcup_{v_2 \in V_2}(V_1 - v_2)$, mit $0 \notin V$. Wähle ein $x_0 \in V$, dann ist $U := V - x_0$ offen mit $0 \in U$. Damit gilt nach Lemma 3.16 für das Minkowski-Funktional $p_U(-x_0) \geq 1$.

Betrachte $Y := \mathbb{R}x_0$ und $y'(-\lambda x_0) := \lambda p_U(-x_0)$. Für $\lambda \leq 0$ ist $y'(-\lambda x_0) \leq 0 \leq p_U(-\lambda x_0)$. Für $\lambda > 0$ gilt nach Sublinearität $y'(-\lambda x_0) = p_U(-\lambda x_0)$, insgesamt also $y'(y) \leq p_U(y)$ für alle $y \in \mathbb{R}x_0$. Nach Satz 3.5 gibt es eine lineare Fortsetzung $x' : X \to \mathbb{R}$ (deren Beschränktheit noch zu zeigen ist) mit $x'(x) \leq p_U(x)$ für alle $x \in X$.

Da U offen ist und $0 \in U$, gibt es ein $\epsilon > 0$ mit $U_\epsilon(0) \subset U$. Damit gilt $p_U(x) \leq \frac{1}{\epsilon}\|x\|$ für alle $x \in X$. Somit folgt $x'(x) \leq \frac{1}{\epsilon}\|x\|$ und auch $x'(-x) = -x'(x) \leq \frac{1}{\epsilon}\|x\|$, d.h. x' ist beschränkt.

Schließlich gilt $x'(-x_0) = y'(-x_0) = p_U(-x_0) \geq 1$. Nach Konstruktion läßt sich jedes $v \in V = V_1 - V_2 = x_0 + U$ schreiben als $v = u + x_0$, so daß nach Lemma 3.16 gilt

$$x'(v) = x'(u) - x'(-x_0) \leq x'(u) - 1 < 0\ .$$

Somit ist $x'(v_1) < x'(v_2)$. Für komplexe Vektorräume liefert diese Konstruktion ein reelles Funktional $\operatorname{Re} x' : X \to \mathbb{R}$, welches über $x'(x) = (\operatorname{Re} x')(x) - \mathrm{i}(\operatorname{Re} x')(\mathrm{i}x)$ ein $\mathbb{C}$-linearen Funktional x' mit den geforderten Eigenschaften definiert. □

Bemerkung 3.18 Der Trennungssatz hat eine interessante geometrische Interpretation. Nach Theorem 3.17 gibt es (wieder im reellen Fall) unter diesen Voraussetzungen ein $\alpha \in \mathbb{R}$ mit $x'(v_1) < \alpha < x'(v_2)$. Wir betrachten zu $f \in X'$ und $\alpha \in \mathbb{R}$ die *Hyperebene* $H(f, \alpha) = \{x \in X \ :\ f(x) = \alpha\}$. Ist $x_0 \in H(f, \alpha)$ gegeben, dann ist $H(f, \alpha) = x_0 + \ker f$, und $\ker f$ ist abgeschlossener Untervektorraum der Kodimension 1. (Sei $f(x_0) \neq 0$, dann ist $x = (x - \frac{f(x)}{f(x_0)}x_0) + \frac{f(x)}{f(x_0)}x_0$, d.h. $X/\ker f = \mathbb{R}x_0$.) Der Trennungssatz besagt also, daß disjunkte konvexe Mengen in X, eine davon offen, durch eine Hyperebene in X getrennt werden können.

Ebenso lassen sich Punkte und abgeschlossene Mengen strikt trennen:

Folgerung 3.19 (Hahn-Banach: V. Trennungssatz (2)) *Sei* $(X, \|\ \|)$ *normierter Vektorraum,* $V \subset X$ *eine abgeschlossene konvexe Teilmenge und* $x \in X \setminus V$*. Dann gibt es ein* $x' \in X'$ *und ein* $\delta > 0$ *mit* $\operatorname{Re} x'(x) + \delta \leq x'(v)$ *für alle* $v \in V$*.*

Beweis. $X \setminus V$ ist offen, also gibt es eine ϵ-Umgebung $U_\epsilon(x)$ mit $U_\epsilon(x) \cap V = \varnothing$. Nach Theorem 3.17 gibt es ein $x' \in X'$ mit

$$(\operatorname{Re} x')(x + u) = (\operatorname{Re} x')(x) + (\operatorname{Re} x')(u) < (\operatorname{Re} x')(v)$$

für alle $v \in V$ und $u \in U_\epsilon(0)$. Nach Supremumsbildung über $u \neq 0$ folgt $(\operatorname{Re} x')(x) + \|\operatorname{Re} x'\|\epsilon \leq (\operatorname{Re} x')(v)$, und analog zu Theorem 3.7.ii) darf $\|\operatorname{Re} x'\|$ durch $\|x'\|$ ersetzt werden. Somit liefert $\delta := \epsilon\|x'\| > 0$ die Behauptung. □

4 Schwache Topologien und reflexive Räume

Ist X normierter Vektorraum, so auch sein Dualraum X' und damit auch der Bidualraum $X'' := (X')'$. Jedes $x \in X$ definiert in kanonischer Weise ein $i_X(x) \in X''$ durch $\big(i_X(x)\big)(x') := x'(x)$.

Satz 4.1 *Die kanonische Abbildung $i_X : X \to X''$ ist eine (im allgemeinen nicht surjektive) lineare Isometrie, $\|i_X(x)\| = \|x\|$.*

Beweis. Linearität ist klar, und die Isometrie ist die Folgerung 3.10 $\|x\| = \sup_{x' \in B_{X'}} |x'(x)| = \sup_{x' \in B_{X'}} |(i_X(x))(x')| = \|i_X(x)\|$ aus Hahn-Banach. □

Folgerung 4.2 *Jeder normierte Vektorraum ist isometrisch isomorph zu einem dichten Untervektorraum eines Banach-Raums.*

Beweis. X'' ist immer vollständig. Nach Satz 1.16.v) ist $\overline{i_X(X)}$ vollständig. □

Definition 4.3 Ein normierter Vektorraum X heißt *reflexiv*, falls die kanonische Abbildung $i_X : X \to X''$ surjektiv ist.

Satz 4.4 *Die Folgenräume ℓ^p sind für $1 < p < \infty$ reflexiv, genauer gilt $(\ell^p)' = \ell^q$ für $\frac{1}{p} + \frac{1}{q} = 1$.*

Beweis. Wir betrachten die Abbildung $S : \ell^q \to (\ell^p)'$ aus Folgerung 2.18 für das Zählmaß. Nach Satz 2.20 ist nur die Surjektivität zu beweisen.

Sei $y' \in (\ell^p)'$ beliebig. Es sei $e_n = (\delta_{nk})_{k \in \mathbb{N}}$ die Folge, deren n-tes Glied gleich 1 ist und ansonsten identisch 0. Offenbar ist $e_n \in \ell^p$ für alle p. Für $c_n \in \mathbb{K}$ gilt für endliche Linearkombination $c_{(N)} := \sum_{n=0}^N c_n e_n \in \ell^p$. Sei $x_n := y'(e_n)$ und $x := (x_n)_{n \in \mathbb{N}}$. Wir zeigen: $x \in \ell^q$. Dazu sei

$$t_n := \begin{cases} \frac{|x_n|^q}{x_n} & \text{für } x_n \neq 0 \\ 0 & \text{für } x_n = 0 \end{cases}$$

Dann ist $|t_n|^p = |x_n|^q$, und für alle $N \in \mathbb{N}$ gilt

$$\begin{aligned} \sum_{n=0}^N |x_n|^q &= \sum_{n=0}^N x_n t_n = \sum_{n=0}^N y(e_n) t_n = y\Big(\underbrace{\sum_{n=0}^N t_n e_n}_{t_{(N)}}\Big) \\ &\leq \|y'\|\, \|t_{(N)}\|_p = \|y'\| \Big(\sum_{n=0}^N |t_n|^p\Big)^{\frac{1}{p}} = \|y'\| \Big(\sum_{n=0}^N |x_n|^q\Big)^{\frac{1}{p}} . \end{aligned}$$

Unter Verwendung von $1 - \frac{1}{p} = \frac{1}{q}$ folgt $\big(\sum_{n=0}^N |x_n|^q\big)^{\frac{1}{q}} \leq \|y'\|$, so daß $x \in \ell^q$ mit $\|x\|_q \leq \|y'\|$. Somit ist $S_x \in (\ell^p)'$ nach Folgerung 2.18, und es gilt

$$S_x(e_n) := \sum_{k=0}^\infty x_k \delta_{nk} = x_n = y'(e_n) .$$

Folglich ist $S_x - y' = 0$ auf $U := \operatorname{span}(e_n \ : \ n \in \mathbb{N}) \subset \ell^p$. Da $U \subset \ell^p$ dicht ist, gilt $S_x - y' = 0$ nach Folgerung 3.12. $\square$

Wir werden später zeigen, daß auch $L^p(X, \mu)$ mit $1 < p < \infty$ reflexiv ist für beliebige Maßräume.

Lemma 4.5 *Sei X reflexiver Banach-Raum und $U \subset X$ abgeschlossener Untervektorraum. Dann ist U reflexiv.*

Zu zeigen ist die Surjektivität von i_U. Jedes $u'' \in U''$ definiert eine lineare Abbildung $\tilde{x} : X' \ni x' \mapsto u''(x'\big|_U) \in \mathbb{K}$, für die gilt

$$\big|u''(x'|_U)\big| \leq \|u''\| \big\|x'|_U\big\| \leq \|u''\| \|x'\| \,.$$

Somit ist $\tilde{x} \in X''$, und wegen der Reflexivität gibt es ein $x \in X$ mit $x'(x) = \tilde{x}(x') = u''(x'\big|_U)$ für alle $x' \in X'$. Wäre $x \notin U$, so gibt es nach Folgerung 3.11 (U ist abgeschlossen) ein $x' \in X'$ mit $x'(x) \neq 0$ und $x'\big|_U = 0$, Widerspruch. Verbleibt zu zeigen: $i_U(x) = u''$. Sei $u' \in U'$ beliebig, dann gilt für eine beliebige Hahn-Banach-Fortsetzung x' von u'

$$u''(u') = u''\big(x'|_U\big) = x'(x) = u'(x)$$

wegen $x \in U$. $\square$

Satz 4.6 *Ein Banach-Raum X ist genau dann reflexiv, wenn X' reflexiv ist.*

i) Sei X reflexiv, zu zeigen ist $i_{X'} : X' \to X'''$ surjektiv. Sei $x''' \in X'''$ beliebig. Definiere $x'(x) := x'''(i_X(x))$; dann ist x' linear und stetig. Zu gegebenem $x'' \in X''$ gibt es (eindeutig) ein $x \in X$ mit $x'' = i_X(x)$. Somit gilt

$$x'''(x'') = x'''(i_X(x)) = x'(x) = (i_X(x))(x') = x''(x') \,,$$

d.h. $i_{X'}(x') = x'''$.

ii) Sei X' reflexiv, dann ist X'' reflexiv nach i). Mit X ist auch $i_X(X) \subset X''$ vollständig, insbesondere abgeschlossen, und nach Lemma 4.5 ist dann $i_X(X)$ reflexiv. Da $i_X : X \to i_X(X)$ ein isometrischer Isomorphismus ist, ist X selbst reflexiv. $\square$

Die linearen Funktionale $f \in X'$ sollen nun genutzt werden, um auf dem normierten Vektorraum $(X, \| \ \|)$ eine (wie sich zeigen wird von der Norm-Topologie verschiedene) Topologie einzuführen, die *schwache Topologie*. Die schwache Topologie wird gerade so konstruiert, daß jedes $f \in X'$ stetig ist, d.h. $f^{-1}(U) \subset X$ ist offen für beliebige offene Teilmengen $U \subset \mathbb{K}$. Diese Konstruktion ist wichtiges Beispiel für eine Initialtopologie:

Definition 4.7 Sei X eine Menge, $(Y_i, \mathcal{T}_i)$ eine Familie von topologischen Räumen und $f_i : X \to Y_i$ eine Familie von Abbildungen von X nach Y_i. Die *Initialtopologie* bezüglich $(Y_i, \mathcal{T}_i, f_i)$ ist die gröbste Topologie auf X, bezüglich derer alle Abbildungen f_i stetig sind.

Ist $(X, \| \ \|)$ ein normierter Vektorraum, so heißt die Initialtopologie auf X bezüglich $(\mathbb{K}, f_i)_{f_i \in X'}$ die *schwache Topologie*.

Das bedeutet (und ist äquivalent zur Definition), daß die Urbilder $f_i^{-1}(\mathcal{O}_{ij})$ aller offenen Mengen $\mathcal{O}_{ij} \in \mathcal{T}_i$ unter allen Abbildungen f_i eine Subbasis $\mathcal{S}$ der Topologie $\mathcal{T}$ bilden, d.h. jede (offene) Menge in $\mathcal{T}$ ist darstellbar als Vereinigung von endlichen Durchschnitten von Elementen aus $\mathcal{S}$.

Da jedes $f \in X'$ auch stetig in der Norm-Topologie ist, ist die schwache Topologie gröber als die Norm-Topologie. Dabei heißt eine eine Topologie $\mathcal{T}_1$ *gröber* als eine Topologie $\mathcal{T}_2$ (und $\mathcal{T}_2$ *feiner* als $\mathcal{T}_1$), wenn $\mathcal{T}_1 \subset \mathcal{T}_2$ ist, d.h. wenn jede in $\mathcal{T}_1$ offene Menge auch in $\mathcal{T}_2$ offen ist. Die schwache Topologie enthält also weniger offene Mengen als die Norm-Topologie, was umgekehrt dazu führt, daß sich die schwache Topologie hinsichtlich Kompaktheit besser verhält als die Norm-Topologie.

Satz 4.8 *Ein normierter Vektorraum X ist bezüglich der schwachen Topologie ein Hausdorff-Raum.*

Beweis. Seien $x_1 \neq x_2$ verschiedene Punkte aus X. Nach Hahn-Banach (Folgerung 3.8) gibt es ein $f \in X'$ mit $\epsilon := \frac{1}{2}|f(x_1) - f(x_2)| > 0$. Wegen der Stetigkeit von f sind $f^{-1}(U_\epsilon(f(x_1)))$ und $f^{-1}(U_\epsilon(f(x_2)))$ offen in der schwachen Topologie und per Konstruktion disjunkt. $\square$

Somit ist die schwache Konvergenz von Folgen (d.h. Konvergenz in der schwachen Topologie) erklärt, und der Grenzwert ist eindeutig.

Satz 4.9 *Sei X normierter Vektorraum. Eine Folge (x_n) in X ist genau dann schwach-konvergent gegen $x \in X$, geschrieben $\lim^w_{n\to\infty} x_n = x$, wenn für alle $f \in X'$ gilt $\lim_{n\to\infty} f(x_n) = f(x)$.*

Beweis. ($\Rightarrow$) Sei $\epsilon > 0$ gegeben, dann liegen für beliebiges $f \in X'$ fast alle x_n in $f^{-1}(U_\epsilon(f(x)))$, somit $|f(x) - f(x_n)| < \epsilon$.

($\Leftarrow$) Nach Konstruktion der schwachen Topologie aus der Subbasis enthält jede schwach-offene Umgebung von x eine Teilmenge der Form

$$\mathcal{U} := f_1^{-1}(\mathcal{O}_1) \cap f_2^{-1}(\mathcal{O}_2) \cap \dots f_m^{-1}(\mathcal{O}_m) ,$$

wobei $\mathcal{O}_i$ offene Umgebungen von $f(x) \in \mathbb{K}$ sind und $f_i \in X'$. Da nach Voraussetzung fast alle $f_i(x_n)$ in jedem $\mathcal{O}_i$ liegen, liegen auch fast alle $x_n \in f_i^{-1}(\mathcal{O}_i)$ und damit in $\mathcal{U}$. $\square$

Aus Norm-Konvergenz folgt schwache Konvergenz, die Umkehrung gilt in unendlich-dimensionalen normierten Vektorräumen aber nicht:

Beispiel 4.10 Sei $X = \ell^p$ mit $1 < p < \infty$. Wir betrachten die Folge $e_n = (\delta_{nk})_{k\in\mathbb{N}} \in \ell^p$ aus Satz 4.4. Ist $n \neq m$, so folgt $\|e_n - e_m\|_p = 2^{\frac{1}{p}}$. Damit ist die Folge der $(e_n)_{n\in\mathbb{N}}$ keine Cauchy-Folge, so daß $(e_n)_{n\in\mathbb{N}}$ in der $\| \ \|_p$-Norm nicht konvergent ist. Andererseits ist nach Satz 4.4 $(\ell^p)' = \ell^q$, d.h. zu $f \in (\ell^p)'$ gibt es $(y_n) \in \ell^q$ mit $f(x) = \sum_{k=0}^{\infty} x_k y_k$ für alle $x = (x_k) \in \ell^p$. Insbesondere gilt $f(e_n) = y_n$, und y_n ist eine Nullfolge für alle $f \in (\ell^p)'$. Somit ist $(e_n)_{n\in\mathbb{N}}$ schwach-konvergent gegen $0 \in \ell^p$. $\triangleleft$

Das Beispiel verdeutlicht insbesondere, daß Norm-abgeschlossene Mengen nicht schwach-abgeschlossen sein müssen. Für die in ℓ^p schwach-konvergente Folge $(e_n)_{n\in\mathbb{N}}$ von Punkten aus S_{ℓ^p} liegt der Grenzwert $0 \notin S_{\ell^p}$, d.h. S_{ℓ^p} ist nicht schwach-abgeschlossen! Es läßt sich zeigen, daß in unendlich-dimensionalen Banach-Räumen X allgemein gilt: Der schwache Abschluß von S_X ist B_X.

Jedoch gilt für *konvexe* Mengen:

Satz 4.11 *Sei X normierter Vektorraum und $V \subset X$ eine (Norm- oder schwach) abgeschlossene konvexe Teilmenge. Ist (v_n) eine Folge aus V, die schwach gegen ein $x \in X$ konvergent ist, so gilt $x \in V$.*

Beweis. In jedem Fall ist V Norm-abgeschlossen. Angenommen, $x \neq V$. Nach dem 2. Trennungssatz von Hahn-Banach (Folgerung 3.19) gibt es ein $f \in X'$ und ein $0 < \epsilon < \delta$ mit $\operatorname{Re} f(x) + \epsilon \leq \inf_{v\in V} \operatorname{Re} f(v)$. Damit ist $(\operatorname{Re} f)(x - v_n) > \epsilon$ für alle $n \in \mathbb{N}$, im Widerspruch zur schwachen Konvergenz gegen x nach Satz 4.9. $\square$

Auch in der schwachen Topologie gilt, daß schwach-konvergente Folgen beschränkt sind. Der Beweis beruht auf dem Baireschen Kategoriensatz und kann erst später gegeben werden. Dagegen sind schwach-konvergente Netze (auch das ist später einzuführen) nicht notwendig beschränkt.

Ebenso ist auf X' die schwache Topologie erklärt als die gröbste Topologie, so daß alle $y \in X''$ stetig sind. Fordert man nur Stetigkeit des Unterraums $i_X(X) \subset X''$, so erhält man:

Definition 4.12 Die schwach-*-Topologie ist die gröbste Topologie auf X', für die alle Abbildungen $X' \ni f \mapsto f(x) \in \mathbb{K}$ stetig sind für $x \in X$.

Ist X nicht reflexiv, so ist die schwach-*-Topologie noch gröber als die schwache Topologie auf X'. Es gilt dennoch, daß X' Hausdorff-Raum bezüglich der schwach-*-Topologie ist (Beweis wie Satz 4.8), und eine Folge (f_n) in X' ist schwach-*-konvergent gegen $f \in X'$ genau dann, wenn sie punktweise konvergent ist, $\lim_{n\to\infty} f_n(x) = f(x)$ für alle $x \in X$ (Beweis wie Satz 4.9). Deshalb heißt die schwach-*-Topologie auch Topologie der punktweisen Konvergenz. Zentrales Resultat ist:

Theorem 4.13 (Banach-Alaoglu) *Sei X normierter Vektorraum. Dann ist $B_{X'}$ kompakt in der schwach-*-Topologie.*

Beweis. Für $x \in X$ definiere kompakte Teilmengen $Y_x := \{\lambda_x \in \mathbb{K} \,:\, |\lambda_x| \leq \|x\|\}$. Nach dem Satz von Tychonoff ist dann auch $Y := \prod_{x\in X} Y_x$ kompakt.

Durch $\eta : B_{X'} \ni f \mapsto (f(x))_{x\in X}$ wird eine Abbildung $\eta : B_{X'} \to Y$ definiert (beachte $|f(x)| \leq \|x\|$). Diese ist injektiv, so daß $B_{X'} \subset Y$ identifiziert werden kann. Dabei stimmt die Relativtopologie von $B_{X'} \subset Y$, induziert aus der Produkttopologie von Y, mit der schwach-$*$-Topologie auf $B_{X'}$ überein: Die Produkt-Topologie auf $Y = \prod_{x\in X} Y_x$ ist nichts anderes als die Initialtopologie bezüglich der Projektionen $p_x : Y \to Y_x \subset \mathbb{K}$.

Es verbleibt zu zeigen, daß $B_{X'} \subset Y$ abgeschlossen ist. Dazu sei $g \in \overline{B_{X'}} \subset Y$ gegeben; wegen $|g(x)| \leq \|x\|$ ist nur die Linearität von g zu zeigen. Seien $x_1, x_2 \in X$ beliebige, aber feste Punkte. Wir betrachten die Menge

$$G := \{y \in Y \,:\, |y(x_1+x_2) - g(x_1+x_2)| < \epsilon\,,\ |y(x_i) - g(x_i)| < \epsilon\,,\ i = 1,2\}\,.$$

Jede der drei Mengen $(i_X(x))^{-1}(U_\epsilon(g(x)))$ ist offen ($x \in \{x_1, x_2, x_1+x_2\}$), somit ist G offene Umgebung von g. Wegen $g \in \overline{B_{X'}}$ existiert ein $f \in G \cap B_{X'}$. Aus der Linearität von f folgt $|g(x_1+x_2) - g(x_1) - g(x_2)| < 3\epsilon$. Die analoge Konstruktion für die skalare Multiplikation beweist die Linearität von g. □

Folgerung 4.14 *Zu jedem normierten Vektorraum $(X, \|\ \|)$ gibt es einen kompakten topologischen Raum K und eine lineare Isometrie $\phi : (X, \|\ \|) \to (\mathcal{C}(K), \|\ \|_\infty)$.*

Beweis. $K := B_{X'}$ ist kompakt in der schwach-*-Topologie. Einem $x \in X$ ordnen wir die Abbildung $\phi_x : B_{X'} \ni f \mapsto f(x) \in \mathbb{K}$ zu. Zu zeigen ist: ϕ ist linear und beschränkt mit $\|\phi_x\|_\infty := \sup_{f\in B_{X'}} |\phi_x(f)| = \|x\|$. Linearität folgt aus $\phi_{x+y}(f) = f(x+y) = f(x) + f(y) = \phi_x(f) + \phi_y(f) = (\phi_x + \phi_y)(f)$, da die Addition von Funktionen auf $K = B_{X'}$ punktweise erklärt ist. Die Isometrie ist durch die Folgerung 3.10 aus Hahn-Banach bewiesen. □

Als Dualraum von X' ist auf X'' ebenfalls die schwach-*-Topologie erklärt, in dieser ist $B_{X''}$ kompakt. Es geht nun darum, diese Kompaktheitseigenschaften auf X zu übertragen. Unser Ziel ist der Beweis des folgenden:

Theorem 4.15 *Ein Banach-Raum X ist genau dann reflexiv, wenn B_X schwach-kompakt ist.*

Folgerung 4.16 *Ein Banach-Raum X ist genau dann reflexiv, wenn auf X' die schwache Topologie und die schwach-*-Topologie übereinstimmen.*

Beweis. Stimmen beide Topologien auf X' überein, so ist $B_{X'}$ schwach-kompakt nach Theorem 4.13, somit ist X' reflexiv nach Theorem 4.15. Nach Satz 4.6 ist dann auch X reflexiv. Die Umkehrung ist klar. $\square$

Der Beweis von Theorem 4.15 beruht auf:

Satz 4.17 *Sei X normierter Vektorraum. Dann ist $i_X(B_X) \subset B_{X''}$ schwach-*-dicht.*

Es gibt einen relativ kurzen Beweis über den Trennungssatz von Hahn-Banach, falls dieser für lokal-konvexe Topologien gezeigt ist (was wir nicht getan haben). Ein Beweis innerhalb des bisher vorgestellten Rahmens ist leider etwas umfangreich. Dazu zunächst folgendes:

Lemma 4.18 *Seien $y \in B_{X''}$ und $f_1, \dots, f_n \in X'$ gegeben. Für $h(x) := \sum_{i=1}^n |y(f_i) - f_i(x)|^2$ gilt: $\inf_{x \in B_X} h(x) = 0$.*

Beweis. Sei $I := \inf_{x \in B_X} h(x)$. Dann gibt es eine Folge (a_k) in B_X mit $\lim_{k\to\infty} h(a_k) = I$. Da $(f_i(a_k))$ beschränkt ist, gibt es nach Bolzano-Weierstraß eine Teilfolge (a_{k_l}), so daß $(f_i(a_{k_l}))$ konvergent ist gegen $\alpha_i := \lim_{l\to\infty} f_i(a_{k_l})$, für alle $i = 1 \dots n$. Der Einfachheit halber sei die Teilfolge wieder mit (a_k) bezeichnet, außerdem sei $\delta_i := y(f_i) - \alpha_i$.

Für $0 \leq \lambda \leq 1$ und $x \in B_X$ ist $\lambda x + (1-\lambda)a_k \in B_X$; somit

$$\begin{aligned}
I \leq h(\lambda x + (1-\lambda)a_k) &= \sum_{i=1}^n \Big| y(f_i) - \lambda f_i(x) - (1-\lambda) f_i(a_k) \Big|^2 \\
&= \sum_{i=1}^n |y(f_i) - f_i(a_k)|^2 - 2\lambda \mathrm{Re}\Big(\sum_{i=1}^n \overline{(y(f_i) - f_i(a_k))}(f_i(x) - f_i(a_k)) \Big) \\
&\qquad + \lambda^2 \sum_{i=1}^n |f_i(x) - f_i(a_k)|^2 \\
&\overset{k\to\infty}{\longrightarrow} I - 2\lambda \mathrm{Re} \sum_{i=1}^n \overline{\delta_i}(f_i(x) - \alpha_i) + \lambda^2 \sum_{i=1}^n |f_i(x) - \alpha_i|^2 \,.
\end{aligned}$$

Für $\lambda \to 0$ erhalten wir $\mathrm{Re} \sum_{i=1}^n \overline{\delta_i}(f_i(x) - \alpha_i) \leq 0$. Somit erfüllt $f := \sum_{i=1}^n f_i \overline{\delta_i}$ für alle $x \in B_X$ die Abschätzung $\mathrm{Re}\, f(x) \leq \mathrm{Re} \sum_{i=1}^n \overline{\delta_i}\alpha_i$, aus der sich nach Variation über $e^{i\beta}x \in B_X$ ergibt: $\|f\| \leq \mathrm{Re} \sum_{i=1}^n \overline{\delta_i}\alpha_i$. Andererseits ist

$$f(a_k) = \sum_{i=1}^n f_i(a_k)\overline{\delta_i} \overset{k\to\infty}{\longrightarrow} \sum_{i=1}^n \overline{\delta_i}\alpha_i \,, \qquad \Rightarrow \qquad \|f\| \geq \Big| \sum_{i=1}^n \overline{\delta_i}\alpha_i \Big|$$

und insgesamt $\|f\| = \sum_{i=1}^n \overline{\delta_i}\alpha_i \geq 0$. Das ergibt

$$I = \lim_{k\to\infty} \sum_{i=1}^n |y(f_i) - f_i(a_k)|^2 = \sum_{i=1}^n \overline{\delta_i}\big(y(f_i) - \alpha_i\big) = y(f) - \|f\| < 0$$

wegen $|y(f)| \leq \|y\|\|f\| \leq \|f\|$. Somit $I = 0$. $\square$

Beweis von Satz 4.17. Sei $y \in B_{X''}$. Nach Definition der schwach-*-Topologie enthält jede schwach-*-Umgebung von y eine offene Teilmenge der Form

$$\{b \in B_{X''} \ : \ |y(f_i) - b(f_i)| < \epsilon\ , \quad \epsilon > 0\ ,\ f_i \in X'\ ,\ i = 1, \ldots n\}$$

Nach Lemma 4.18 enthält diese Menge ein $b = i_X(x)$ für $x \in B_X$. $\square$

Beweis von Theorem 4.15. Die schwach-*-Topologie auf X'' ist definiert als jene Topologie, für die alle Abbildungen $X'' \ni y \mapsto y(f) \in \mathbb{K}$ für beliebige $f \in X'$ stetig sind. Für $y = i_X(x)$ ist $y(f) = f(x)$. Die Stetigkeit der Abbildung $X' \ni x \mapsto f(x) \in \mathbb{K}$ definiert gerade die schwache Topologie auf X, d.h. die schwache Topologie auf X ist die Relativtoplogie der schwach-*-Topologie von $i_X(X) \subset X''$.

Sei X reflexiv, dann werden mittels i_X beide Topologien mitenander identifiziert. Da $B_{X''} = i_X(B_X)$ schwach-*-kompakt nach Theorem 4.13 ist, ist B_X schwach-kompakt.

Sei umgekehrt B_X schwach-kompakt, so ist $i_X(B_X)$ schwach-*-kompakt, insbesondere schwach-*-abgeschlossen. Da nach Satz 4.17 $i_X(B_X)$ schwach-*-dicht in $B_{X''}$ ist, gilt bereits $i_X(B_X) = B_{X''}$ und dann $i_X(X) = X''$. $\square$

Wir widmen uns nun Fragen der *Folgenkompaktheit* bezüglich der schwachen Topologien. Dazu sind zunächst wichtige topologische Begriffe einzuführen.

Definition 4.19 Sei $(X, \mathcal{T})$ ein topologischer Raum.

i) X heißt *separabel*, wenn es eine abzählbare dichte Teilmenge $A \subset X$ gibt.

ii) X erfüllt das *erste Abzählbarkeitsaxiom*, falls es zu jedem $x \in X$ eine abzählbare Umgebungsbasis gibt, d.h. es gibt abzählbar viele Umgebungen $(U_i)_{i\in\mathbb{N}}$ von x, so daß es zu jeder Umgebung V von x ein $U_i \subset V$ gibt.

iii) X erfüllt das *zweite Abzählbarkeitsaxiom*, falls $\mathcal{T}$ eine abzählbare (Sub-)Basis hat, d.h. es gibt abzählbar viele offene Teilmengen $(W_i)_{i\in\mathbb{N}}$ in X, so daß für jede offene Teilmenge $U \subset X$ gilt: $U = \bigcup_{W_j \subset U} W_j$.

Jeder metrische Raum erfüllt das erste Abzählbarkeitsaxiom, jeder kompakte metrische Raum das zweite Abzählbarkeitsaxiom.

Lemma 4.20 *Für einen normierten Raum X sind äquivalent:*

i) *X ist separabel.*

ii) *Es gibt eine abzählbare Teilmenge $A \subset X$ mit $X = \overline{\mathrm{span}(A)}$.*

Beweis. i)$\Rightarrow$ii) ist klar; für die Umkehrung betrachte die abzählbare Menge $B := \mathrm{span}_{\mathbb{Q}}(A)$ bzw. $B := \mathrm{span}_{\mathbb{Q}+i\mathbb{Q}}(A)$. Sei $x \in X$. Zu jedem $\epsilon > 0$ gibt es nach Voraussetzung endlich viele $x_1, \ldots, x_m \in A$ und $\lambda_1, \ldots, \lambda_m \in \mathbb{K}$, so daß $\|x - a\| <$

ϵ für $a := \sum_{k=1}^m \lambda_k x_k$. Weiter gibt es $\rho_i \in \mathbb{Q}$ bzw. $\rho_i \in \mathbb{Q} + \mathrm{i}\mathbb{Q}$ mit $|\lambda_i - \rho_i| \leq \big(\sum_{k=1}^m \|x_k\|\big)^{-1}\epsilon$. Für $b := \sum_{k=1}^m \rho_k x_k$ gilt dann

$$\|x - b\| \leq \|x - a\| + \|a - b\| < \epsilon + \sum_{k=1}^m |\lambda_k - \rho_k| \|x_k\| < 2\epsilon\,. \qquad \square$$

Satz 4.21 $(\mathcal{C}([a,b]), \|\ \|_\infty)$ *ist separabel.*

Beweisidee: (Grundlagen werden später bereitgestellt) Betrachte die Polynomfunktionen mit rationalen Koeffizienten. Diese sind dicht in $(\mathcal{C}([a,b]), \|\ \|_\infty)$ nach dem Weierstraßschen Approximationssatz. $\square$

Man kann zeigen, daß $(\mathcal{C}(K), \|\ \|_\infty)$ separabel ist für beliebige kompakte metrische Räume K. Unter gleichen Voraussetzungen ist auch $L^p(K,\mu)$ mit $1 \leq p < \infty$ separabel für reguläre Borelmaße μ.

Satz 4.22 *Sei $(X, \|\ \|)$ ein normierter Vektorraum, und X' sei separabel. Dann ist auch X separabel.*

Beweis. Mit X' ist auch $S_{X'}$ separabel, d.h. es gibt eine abzählbare dichte Teilmenge $A = (x_n')_{n\in\mathbb{N}} \subset S_{X'}$. Nach Definition der Norm gibt es $x_n \in S_X$ mit $\frac{1}{2} \leq |x_n'(x_n)| \leq 1$. Setze $U := \mathrm{span}(x_n\ :\ n \in \mathbb{N})$ und betrachte $x' \in X'$ mit $x'\big|_U = 0$. Angenommen, $x' \neq 0$. Durch Skalieren kann $x' \in S_{X'}$ angenommen werden, so daß es ein $x_k' \in A$ gibt mit $\|x' - x_k'\| < \frac{1}{4}$. Somit gilt

$$\frac{1}{2} \leq |x_k'(x_k)| = |x_k'(x_k) - x'(x_k)| \leq \|x_k' - x_k\| \|x_k\| \leq \frac{1}{4}\,,$$

Widerspruch. Somit ist $x' = 0$ identisch auf X, und nach Folgerung 3.12 aus Hahn-Banach ist U dicht in X. Nach Lemma 4.20 ist X separabel. $\square$

Folgerung 4.23 *Ein reflexiver Raum X ist genau dann separabel, wenn X' separabel ist.* $\square$

Satz 4.24 *Ein normierter Vektorraum $(X, \|\ \|)$ sei separabel in der Normtopologie. Dann erfüllt $B_{X'}$ (und damit jede beschränkte Teilmenge von X') das erste Abzählbarkeitsaxiom bezüglich der schwach-*-Topologie.*

Beweis. Zu $f \in X'$ betrachten wir folgende Umgebungen in der schwach-*-Topologie, parametrisiert durch $\epsilon > 0$ und $x_1, \dots, x_r \in X$:

$$U(f; x_1, \dots, x_r, \epsilon) := \{g \in X'\ :\ |g(x_i) - f(x_i)| < \epsilon\ \forall i = 1, \dots, r\}\,.$$

Nach Konstruktion der schwach-*-Topologie aus der Subbasis enthält jede schwach-*-Umgebung von f eine Umgebung der Form $U(f; x_1, \dots, x_r, \epsilon)$. Ist X

separabel in der Norm-Topologie mit Norm-dichter Teilmenge $A = \{a_n : n \in \mathbb{N}\}$, dann betrachten wir die abzählbare Familie $U(f; a_{n_1}, \ldots, a_{n_r}, \frac{1}{m})$ und zeigen: Für beliebige $\epsilon > 0$ und $x_1, \ldots, x_r \in X$ gibt es $m \in \mathbb{N}$ und $a_{n_1}, \ldots, a_{n_r} \in A$ mit $U(f; a_{n_1}, \ldots, a_{n_r}, \frac{1}{m}) \subset U(f; x_1, \ldots, x_r, \epsilon)$. Wähle $m > \frac{3}{\epsilon}$ und $a_{n_i} \in A$ mit $\|a_{n_i} - x_i\| < \frac{1}{m}$. Dann gilt für $f, g \in B_{X'}$ mit $g \in U(f; a_{n_1}, \ldots, a_{n_r}, \frac{1}{m})$ die Abschätzung

$$\begin{aligned} |g(x_k) - f(x_k)| &\leq |g(a_k) - f(a_k)| + |(g - f)(x_k - a_k)| \\ &< \frac{1}{m} + \|g - f\| \|x_k - a_k\| < \frac{1}{m} + 2 \cdot \frac{1}{m} < \epsilon \,. \end{aligned}$$ □

Bemerkung 4.25 In topologischen Räumen, in denen das erste Abzählbarkeitsaxiom *nicht* erfüllt ist, kann die Topologie nicht allein durch die Betrachtung von Folgen erfaßt werden. Man muß dann zu Netzen übergehen. Der vorige Satz zeigt, daß in *separablen* normierten Räumen die Betrachtung von Folgen genügt.

Satz 4.26 *Es sei X ein kompakter topologischer Raum, der das erste Abzählbarkeitsaxiom erfüllt. Dann ist X folgenkompakt, d.h. jede Folge in X besitzt eine konvergente Teilfolge.*

Beweis. Für die Folge $(x_n)_{n\in\mathbb{N}}$ betrachte die Mengen $A_n := \overline{\{x_k : k \geq n\}} \subset X$, mit Abschluß in der Topologie von X. Diese haben die *endliche Durchschnittseigenschaft*, $A_{i_1} \cap \cdots \cap A_{i_k} \neq \varnothing$. Da X kompakt ist, gibt es ein $a \in \bigcap_{n\in\mathbb{N}} A_n \neq \varnothing$. Insbesondere ist $a \in A_0$, d.h. jede Umgebung von a enthält einen Punkt x_k der Folge. Nach dem ersten Abzählbarkeitsaxiom gibt es eine Folge $(U_i)_{i\in\mathbb{N}}$ von Umgebungen von a, so daß es zu jeder Umgebung V von a ein $U_i \subset V$ gibt. Betrachte die Folge $(V_j)_{j\in\mathbb{N}}$ der Umgebungen $V_j = \bigcap_{i=0}^{j} U_i$. Dann gibt es zu jedem V_j ein $x_{n_j} \in V_j$. Somit konvergiert die Teilfolge $(x_{n_j})_{j\in\mathbb{N}}$ gegen a. □

Folgerung 4.27 *Sei X ein normierter Vektorraum, der separabel in der Norm-Topologie ist. Dann ist $B_{X'}$ schwach-*-folgenkompakt.* □

In metrischen Räumen sind Kompaktheit und Folgenkompaktheit äquivalent, in allgemeinen topologischen Räumen aber nicht. Ein einfaches Beispiel für (kompakt $\not\Rightarrow$ folgenkompakt) ist:

Beispiel 4.28 Wir betrachten $X = \ell^\infty$. Nach Banach-Alaoglu ist $B_{X'}$ kompakt in der schwach-*-Topologie. Aber $B_{X'}$ ist nicht schwach-*-folgenkompakt. Dazu sei $\delta_n \in X'$ definiert durch $\delta_n(x) := x_n$ für $x = (x_n) \in \ell^\infty$, dann gilt $\|\delta_n\| \leq 1$, so daß $(\delta_n)_{n\in\mathbb{N}}$ eine (beschränkte) Folge in $B_{X'}$ ist. Angenommen, eine Teilfolge $(\delta_{n_k})_{k\in\mathbb{N}}$ wäre schwach-*-konvergent gegen $\delta \in B_{X'}$, dann gilt nach der Charakterisierung der schwach-*-Topologie als Topologie der punktweisen Konvergenz

$\lim_{k\to\infty} \delta_{n_k}(x) = \delta(x)$ für alle $x \in \ell^\infty$, d.h. die *gleiche* Auswahl von Folgegliedern $(x_{n_k})_{k\in\mathbb{N}}$ wäre für jede Folge $x = (x_n) \in \ell^\infty$ konvergent, Widerspruch. Insbesondere gilt: ℓ^∞ ist nicht separabel in der Norm-Topologie.

Ebenso lassen sich Beispiele für (folgenkompakt $\not\Rightarrow$ kompakt) finden, selbst in topologischen Räumen, die das erste Abzählbarkeitsaxiom erfüllen.

Satz 4.29 *Ist X reflexiv, so ist B_X schwach-folgenkompakt.*

In Kombination von Theorem 4.15 mit Satz 4.26 ist zu zeigen, daß X das erste Abzählbarkeitsaxiom erfüllt. Sei zunächst X separabel, dann auch X' und X'' nach Folgerung 4.23. Nach Satz 4.24 erfüllt $B_{X''}$ das erste Abzählbarkeitsaxiom bezüglich der schwach-*-Topologie. Entsprechend der Diskussion im Beweis von Theorem 4.15 erfüllt dann B_X das erste Abzählbarkeitsaxiom bezüglich der schwachen Topologie.

Im allgemeinen Fall betrachte zu gegebener Folge (x_n) in B_X den Norm-abgeschlossenen Untervektorraum $U = \overline{\mathrm{span}(x_n \; : \; n \in \mathbb{N})} \subset X$. Nach Konstruktion ist U separabel, außerdem reflexiv nach Lemma 4.5. Nach dem 1. Teil besitzt (x_n) eine konvergente Teilfolge. □

In Kombination mit Satz 4.6, Folgerung 4.16 und Beispiel 4.28 folgt, daß ℓ^∞ nicht reflexiv sein kann.

Bemerkung 4.30 Es gilt auch die Umkehrung: *Ist B_X schwach-folgenkompakt, so ist X reflexiv (beruht auf Satz von Eberlein-Šmulian).* Dieser ist aber schwieriger zu beweisen.

5 Gleichmäßig konvexe Räume

In diesem Abschnitt werden wir beweisen, daß gleichmäßig konvexe Räume immer reflexiv sind, und daß die Räume $L^p(X, \mu)$ gleichmäßig konvex, damit reflexiv, sind. Gleichmäßge Konvexität ist sehr nützlich für Approximationen.

Definition 5.1 Ein normierter Vektorraum $(X, \| \, \|)$ heißt *gleichmäßig konvex*, wenn eine der folgenden äquivalenten Bedingungen erfüllt ist:

i) Für alle $\epsilon > 0$ gibt es ein $\delta > 0$, so daß für $x, y \in B_X$ (insbesondere $x, y \in S_X$) gilt:

$$\begin{aligned} & \|x - y\| \geq \epsilon \quad \Rightarrow \quad \|\tfrac{1}{2}(x + y)\| \leq 1 - \delta \\ \text{bzw.} \quad & \|\tfrac{1}{2}(x + y)\| > 1 - \delta \quad \Rightarrow \quad \|x - y\| < \epsilon \, . \end{aligned}$$

ii) Für beliebige Folgen (x_n), (y_n) in S_X (bzw. B_X) gilt

$$\lim_{n\to\infty} \|\tfrac{1}{2}(x_n + y_n)\| = 1 \quad \Rightarrow \quad \lim_{n\to\infty} (x_n - y_n) = 0 \, .$$

Beweis der Äquivalenzen. i)⇒ii) ist klar für Folgen aus S_X. Sind die Folgen aus B_X, so ist $\limsup_n \|x_n\| \leq 1$ und tatsächlich $\lim_n \|x_n\| = \lim_n \|y_n\| = 1$ wegen $\lim_n \|\frac{1}{2}(x_n + y_n)\| = 1$. Damit ist $x_n, y_n \neq 0$ für alle $n \geq N$, und man zeigt für die normierten Folgen $\tilde{x}_n := \frac{x_{N+n}}{\|x_{N+n}\|}$ und $\tilde{y}_n := \frac{y_{N+n}}{\|y_{N+n}\|}$, daß $\lim_{n\to\infty}(\tilde{x}_n - \tilde{y}_n) = 0$. Daraus ergibt sich $\lim_{n\to\infty}(x_n - y_n) = 0$.

ii)⇒i) Angenommen, es gibt ein $\epsilon > 0$, so daß für alle $\delta_n = \frac{1}{n}$ gilt: Es gibt $x_n, y_n \in B_X$ mit $\|x_n - y_n\| > \epsilon$ aber $\|\frac{1}{2}(x_n + y_n)\| > 1 - \delta$. Nach ii) ist das ausgeschlossen. □

Satz 5.2 *Jeder unitäre Vektorraum* $(X, \langle\ ,\ \rangle)$ *ist gleichmäßig konvex.*

Beweis. In X gilt die Parallelogrammgleichung $\|x+y\|^2 + \|x-y\|^2 = 2(\|x\| + \|y\|)$. Sei $0 < \delta < 1$. Sind $x, y \in B_X$ mit $\frac{1}{2}\|x + y\| > 1 - \delta$, so folgt

$$\|x - y\|^2 = 2(\|x\| + \|y\|) - \|x + y\|^2 \leq 4 - 4(1 - \delta)^2 = 8\delta - 4\delta^2 < 8\delta\ .$$

Wähle $0 < \delta < \frac{1}{8}\epsilon^2$, dann $\|x - y\| < \epsilon$ für alle $x, y \in B_X$ mit $\frac{1}{2}\|x + y\| > 1 - \delta$. □

Beispiel 5.3 ℓ^1 ist nicht gleichmäßig konvex: Wähle $x = e_1 = (1, 0, 0, \dots)$ und $y = e_2 = (0, 1, 0, \dots)$, dann ist $\|\frac{1}{2}(x + y)\| = 1$ aber $\|x - y\|_1 = 2$. ◁

Wir werden dagegen sehen, daß ℓ^p gleichmäßig konvex ist für alle $1 < p < \infty$.

Theorem 5.4 (Milman) *Jeder gleichmäßig konvexe Banach-Raum* $(X, \|\ \|)$ *ist reflexiv.*

Beweis. Zu zeigen ist, daß es zu $y \in X''$ ein $x \in X$ gibt mit $y(f) = f(x)$ für alle $f \in X'$. Wir können $y \neq 0$ annehmen (sonst $x = 0$) und wegen der Linearität von f außerdem $\|y\| = 1$. Somit gibt es eine Folge $(f_n)_{n\geq 1}$ in $B_{X'}$ mit $1 - \frac{1}{n} < |y(f_n)| \leq 1$. Durch Multiplikation mit geeigneten $e^{\mathrm{i}\alpha_n}$ kann $1 - \frac{1}{n} < y(f_n) \leq 1$ erreicht werden. Nach Lemma 4.18 gibt es zu jedem $n \geq 1$ ein $x_n \in B_X$ mit

$$\begin{aligned} & |y(f_i) - f_i(x_n)| < \tfrac{1}{2n} \quad \text{für alle } i = 1, \dots, n\ , && (*) \\ \Rightarrow\quad & 1 - \tfrac{3}{2n} < \operatorname{Re} f_i(x_n) \quad \text{für alle } i = 1, \dots, n\ . \end{aligned}$$

Somit gilt für $m \geq n$ unter Verwendung von $f_i \in B_{X'}$ die Abschätzung

$$\big(1 - \tfrac{3}{2n}\big) + \big(1 - \tfrac{3}{2m}\big) < \operatorname{Re}(f_n(x_n) + f_n(x_m)) \leq |f_n(x_n + x_m)| \leq \|x_n + x_m\|\ . \quad (**)$$

Da X gleichmäßig konvex ist, gibt es zu $\epsilon > 0$ ein $\frac{3}{2\delta} \leq N \in \mathbb{N}$, so daß für alle $m \geq n \geq N$ gilt $\|x_n - x_m\| < \epsilon$. Somit ist (x_n) eine Cauchy-Folge, die wegen der Vollständigkeit von X einen Norm-Grenzwert $x \in X$ hat. Aus (**) folgt für $m = n \to \infty$ sogar $\|x\| = 1$, und die Norm-Stetigkeit von f_i in (*) liefert

$$y(f_i) = f_i(x) \qquad \text{für alle } i \in \mathbb{N} \setminus \{0\}\ . \qquad (***)$$

Der Grenzwert x ist durch die Folge (f_i) sowie (***) und $\|x\| = 1$ eindeutig bestimmt. Denn gäbe es einen weitere Lösung $\tilde{x} \in S_X$ von (***), so erfüllt auch die Folge $\tilde{x}_n$ mit $\tilde{x}_{2k} = x$ und $\tilde{x}_{2k+1} = \tilde{x}$ alle früheren Bedingungen, ist somit Cauchy-Folge mit $x = \tilde{x}$.

Wir behaupten, daß sogar $y(f_0) = f_0(x)$ gilt für beliebiges $f_0 \in X'$. Wir können $f_0 \neq 0$ annehmen. Dann finden wir nach Lemma 4.18 eine u.U. andere Folge $(\tilde{x}_n)_{n\geq 1}$, für die statt (*) gilt $|y(f_i) - f_i(\tilde{x}_n)| < \frac{1}{2n}$ für alle $i = 0, 1, \ldots, n$. Da diese Folge $(\tilde{x}_n)$ auch alle früheren Bedingungen erfüllt, konvergiert sie wegen der Eindeutigkeit des Grenzwertes ebenfalls gegen x. Dann folgt für $i = 0, n \to \infty$ die Behauptung. $\square$

Folgerung 5.5 *Jeder Hilbert-Raum $(H, \langle\ ,\ \rangle)$ ist reflexiv.* $\square$

Satz 5.6 (Clarkson) *Sei $1 < p < \infty$. Dann ist $L^p(X, \mu)$ gleichmäßig konvex, damit reflexiv, für beliebige Maßräume (X, μ).*

Wir geben nur den Beweis für $2 \leq p < \infty p \geq 2$. Der Fall $1 < p < 2$ ist wesentlich komplizierter und z.B. in [Hirzebruch-Scharlau] zu finden. Wir benötigen folgende Abwandlung der Parallelogrammgleichung für $p = 2$:

Lemma 5.7 *Für alle $f, g \in L^p(X, \mu)$, $2 \leq p < \infty$, gilt $\|f + g\|_p^p + \|f - g\|_p^p \leq 2^{p-1}(\|f\|_p^p + \|g\|_p^p)$.*

Beweis. i) Wir zeigen: Für alle $a, b \in \mathbb{K}$ und $p \geq 2$ gilt $(|a|^p + |b|^p)^{\frac{1}{p}} \leq (|a|^2 + |b|^2)^{\frac{1}{2}}$. Wir können $a, b \neq 0$ annehmen, und dann $|a|^2 + |b|^2 = 1$. Dann ist $|a|, |b| \leq 1$, somit $|a|^p \leq |a|^2$ und $|b|^p \leq |b|^2$, und schließlich $(|a|^p + |b|^p)^{\frac{1}{p}} \leq (|a|^2 + |b|^2)^{\frac{1}{p}} = 1 = (|a|^2 + |b|^2)^{\frac{1}{2}}$.

ii) Wir zeigen: $|a + b|^p + |a - b|^p \leq 2^{p-1}(|a|^p + |b|^p)$ für alle $a, b \in \mathbb{K}$ und $2 < p < \infty$ (klar für $p = 2$). Sei $p' := \frac{p}{2} > 1$ und $q' := \frac{p}{p-2} > 1$. Es gilt $\frac{1}{p'} + \frac{1}{q'} = 1$. Wir nutzen Hölder für $(\mathbb{K}^2, \|\ \|_{p'})$ und $(\mathbb{K}^2, \|\ \|_{q'})$:

$$\begin{aligned} |a|^2 + |b|^2 = |a|^2 \cdot 1 + |b|^2 \cdot 1 &\leq \|(|a|^2, |b|^2)\|_{p'} \|(1, 1)\|_{q'} \qquad (*) \\ &= (|a|^{2p'} + |b|^{2p'})^{\frac{1}{p'}} (1 + 1)^{\frac{1}{q'}} = (|a|^p + |b|^p)^{\frac{2}{p}} (2)^{\frac{p-2}{p}} . \end{aligned}$$

Es folgt:

$$(|a+b|^p + |a-b|^p)^{\frac{1}{p}} \overset{\text{i)}}{\leq} (|a+b|^2 + |a-b|^2)^{\frac{1}{2}} = (2|a|^2 + 2|b|^2)^{\frac{1}{2}} \overset{(*)}{\leq} 2^{\frac{1}{2}} (|a|^p + |b|^p)^{\frac{1}{p}} 2^{\frac{p-2}{2p}} ,$$

die p-te Potenz liefert die Behauptung.

iii) Zu $f, g \in L^p(X, \mu)$ wählen wir beliebige Repräsentanten in $\tilde{f}, \tilde{g} \in \mathcal{L}^p(X, \mu)$. Dann ist $\tilde{f}(x), \tilde{g}(x) \in \mathbb{K}$ für $x \in X \setminus N$, wobei N eine μ-Nullmenge ist. Somit folgt $|\tilde{f}(x) + \tilde{g}(x)|^p + |\tilde{f}(x) - \tilde{g}(x)|^p \leq 2^{p-1}(|\tilde{f}(x)|^p + |\tilde{g}(x)|^p)$ für $x \in X \setminus N$

nach ii). Das Integral erhält die Monotonie und ist unabhängig von der Wahl der Repräsentanten sowie der Nullmenge N. $\square$

Beweis von Satz 5.6 für $p \geq 2$. Zu $0 < \epsilon < 2$ wähle $\delta := 1-(1-(\frac{\epsilon}{2})^p)^{\frac{1}{p}} > 0$. Seien $f, g \in L^p(X,\mu)$ mit $\|f\|_p, \|g\|_p \leq 1$ sowie $\|f-g\|_p \geq \epsilon$. Dann gilt nach Lemma 5.7

$$\frac{1}{2^p}\|f+g\|_p^p \leq \frac{1}{2}(\|f\|_p^p + \|g\|_p^p) - \frac{1}{2^p}\|f-g\|_p^p \leq 1 - \frac{\epsilon^p}{2^p} = (1-\delta)^p \,.$$

Die p-te Wurzel liefert die Behauptung nach Definition 5.1.i). $\square$

Bemerkung 5.8 Für $1 < p < 2$ gilt statt Lemma 5.7 die Ungleichung $\|f+g\|_p^q + \|f-g\|_p^q \leq 2(\|f\|_p^p + \|g\|_p^p)^{q-1}$ mit $\frac{1}{p}+\frac{1}{q} = 1$, deren Beweis wesentlich schwieriger ist. Der Beweis des Satzes von Clarkson ist dann analog.

Wir können nun die Bemerkung 2.21 vollenden:

Satz 5.9 *Für $1 < p < \infty$ ist die lineare Isometrie $S : L^q(X,\mu) \to (L^p(X,\mu))'$ aus Satz 2.20 surjektiv.*

Beweis. Sei $P := L^p(X,\mu)$ und $Q := S(L^q(X,\mu))$, dann ist Q vollständig, insbesondere abgeschlossener Untervektorraum von P'. Angenommen, es gibt ein $f \in P' \setminus Q$. Nach Hahn-Banach (Folgerung 3.11) gibt es ein $y \in P''$ mit $y\big|_Q = 0$ und $y(f) \neq 0$. Da P reflexiv ist, gibt es ein $x \in P$ mit $y(f) = f(x) \neq 0$, andererseits gilt $0 = y(S_g) = S_g(x)$ für alle $g \in L^q(X,\mu)$. Nun ist S_g isometrisch, d.h. es gilt $x = 0$ and damit auch $y(f) = 0$, Widerspruch. $\square$

Abschließend widmen wir uns Approximationsfragen.

Definition 5.10 Ein reeller normierter Vektorraum $(X, \|\ \|)$ heißt *strikt normiert*, falls für alle $x, y \in X \setminus \{0\}$ gilt:

$$\|x+y\| = \|x\| + \|y\| \quad \Rightarrow \quad x = \lambda y \text{ für ein } \lambda > 0 \,.$$

Lemma 5.11 *Jeder gleichmäßig konvexe reelle normierte Vektorraum ist strikt normiert.*

Beweis. In der Situation von Definition 5.10 können wir $\|x\| = 1$ annehmen. Sei $\tilde{x} := \frac{x+y}{\|x+y\|} = \frac{x+y}{1+\|y\|}$, dann gilt $\|x+y-\tilde{x}\| = \|y\|$ und dann $\mu := \|\frac{x-\tilde{x}'}{2} + y\| \leq \|y\|$. Es folgt $1 + \|y\| = \|x+y\| \leq \mu + \|\frac{1}{2}(x+\tilde{x})| \leq \|y\| + 1$, somit $\mu = \|y\|$ und $\|\frac{1}{2}(x+\tilde{x})\| = 1$. Aus der gleichmäßigen Konvexität folgt $x = \tilde{x}$, also $y = \|y\|x$.$\square$

Satz 5.12 (Approximationssatz) *Sei X ein strikt normierter reeller Vektorraum, $x \in X$ und $V \subset X$ eine abgeschlossene konvexe Teilmenge. Dann gilt:*

i) *Es gibt höchstens ein $w \in V$ mit $\|x-w\| = d(x,V) = \inf_{v \in V} \|x-v\|$.*

ii) *Ist X gleichmäßig konvexer Banach-Raum, so gibt es genau ein $w \in V$ mit $\|x - w\| = d(x, V)$.*

Beweis. Für $x \in V$ ist nichts zu zeigen. Für $x \notin V$ ist $d := \inf_{v \in V} \|x - v\| > 0$ wegen der Abgeschlossenheit von V.

i) Seien $w_1, w_2 \in V$ zwei Minima, dann ist wegen der Konvexität $\frac{1}{2}(w_1 + w_2) \in V$. Es folgt $d \leq \|x - \frac{1}{2}(w_1 + w_2)\| \leq \frac{1}{2}\|x - w_1\| + \frac{1}{2}\|x - w_2\| = d$, somit Gleichheit. Aus der strikten Normiertheit folgt $x - w_1 = \lambda(x - w_2)$ mit $\lambda > 0$, aus der Gleichheit der Normen dann $w_1 = w_2$.

ii) Es gibt eine Folge (v_n) in V mit $\lim_{n\to\infty} \|x - v_n\| = d$. Aus der Konvexität folgt $d \leq \|x - \frac{1}{2}(v_n + v_m)\|$. Wir zeigen: (v_n) ist Cauchy-Folge, wegen der Vollständigkeit von X und Abgeschlossenheit von V konvergent gegen ein $w \in V$; dieses ist nach i) das einzige Minimum.

Zu jedem $\epsilon > 0$ gibt es ein $0 < \delta < 1$ nach Definition 5.1. Wegen der Konvergenz von $(\|x - v_n\|)$ gegen d gibt es ein $N \in \mathbb{N}$, so daß für alle $n \geq N$ gilt $x_n := (1 - \frac{\delta}{2})\frac{1}{d}(x - v_n) \in B_X$. Wäre $\|x_n - x_m\| \geq \epsilon$ für alle $n, m \geq N$, so gilt nach gleichmäßiger Konvexität $\|\frac{1}{2}(x_n + x_m)\| \leq 1 - \delta$, also

$$d \leq \|x - \tfrac{1}{2}(v_n + v_m)\| \leq d\frac{1 - \delta}{1 - \frac{\delta}{2}} < d\ ,$$

Widerspruch. Also ist (x_n) Cauchy-Folge und damit auch (v_n). □

Der Satz wird vor allem genutzt in der Situation, daß V abgeschlossener Untervektorraum eines gleichmäßig konvexen Banach-Raums X ist. Dann gibt es zu $x \in X$ ein eindeutiges $w \in V$ mit $d(x, V) = d(x, w)$. Als Beispiel wähle $X = L^2([a, b], \lambda)$ und $V = P_n$ als Vektorraum der Polynome vom Grad $\leq n$. Dann gibt es zu jeder Äquivalenzklasse von Funktionen $[f] \in L^2([a, b], \lambda)$ ein eindeutiges Polynom $p_n(x) = \sum_{i=0}^{n} a_i x^i$, für das der quadratische Fehler $\int_{[a,b]} d\lambda\ |f(x) - p_n(x)|^2$ minimal wird.

6 Der Satz von Baire und seine Konsequenzen

In vollständigen metrischen Räumen X bleibt der Durchschnitt abzählbar vieler offener und dichter Teilmengen immer noch dicht in X. Daraus ergeben sich eine Reihe der wichtigsten Theoreme der Funktionalanalysis, die auf den ersten Blick unerwartet sind.

Satz 6.1 (Baire) *Sei (X, d) ein vollständiger metrischer Raum und $(\mathcal{O}_n)$ eine Folge offener und dichter Teilmengen. Dann ist $Y := \bigcap_{n \in \mathbb{N}} \mathcal{O}_n$ dicht in X.*

Beweis. Wir zeigen induktiv, daß jede ϵ-Umgebung $U_\epsilon(x)$ ein $y \in Y$ enthält. Da $\mathcal{O}_1$ dicht ist, gibt es ein $x_1 \in \mathcal{O}_1 \cap U_\epsilon(x)$, und da $\mathcal{O}_1 \cap U_\epsilon(x)$ offen, gibt es ein $\epsilon_1 > 0$ mit $U_{\epsilon_1}(x_1) \subset \mathcal{O}_1 \cap U_\epsilon(x)$. Durch Verkleinerung von ϵ_1 läßt sich

$\overline{U_{\epsilon_1}(x_1)} \subset \mathcal{O}_1 \cap U_\epsilon(x)$ erreichen, und es kann $\epsilon_1 < \frac{1}{2}\epsilon$ gewählt werden. Induktiv konstruiert man eine Folge (x_n) in X und eine Folge (ϵ_n) positiver Zahlen mit

i) $\overline{U_{\epsilon_n}(x_n)} \subset \mathcal{O}_n \cap U_{\epsilon_{n-1}}(x) \subset \mathcal{O}_1 \cap \cdots \cap \mathcal{O}_n \cap U_\epsilon(x)$,

ii) $\epsilon_n < \frac{1}{2}\epsilon_{n-1} < \frac{1}{2^n}\epsilon$.

Nach Konstruktion ist (x_n) eine Cauchy-Folge, die wegen der Vollständigkeit von X gegen ein $y \in X$ konvergiert. Dabei gilt $y \in \overline{U_{\epsilon_n}(x_n)} \subset \mathcal{O}_n \cap U_\epsilon(x)$ für alle $n \in \mathbb{N}$ und deshalb $y \in \big(\bigcap_{n\in\mathbb{N}} \mathcal{O}_n\big) \cap U_\epsilon(x)$, insbesondere $y \in Y$. □

Es wird nicht behauptet, daß $\bigcap_{n\in\mathbb{N}} \mathcal{O}_n$ offen ist, was i.a. auch nicht gilt. Der Satz läßt sich ebenfalls zeigen in lokal-kompakten Hausdorff-Räumen. Der ursprüngliche Satz von Baire formuliert die komplementäre Aussage und benötigt:

Definition 6.2 Eine Teilmenge A eines metrischen Raumes X heißt *nirgends dicht*, falls $\bar{A}$ keinen inneren Punkt besitzt.

Dabei heißt $\mathrm{int}(A) := \bigcup\{U \subset A \text{ offen}\}$ das *Innere* von A, und Punkte $y \in \mathrm{int}(A)$ heißen innere Punkte.

Lemma 6.3 *In einem metrischen Raum X gilt:* $\mathrm{Int}(A) = \varnothing \quad \Leftrightarrow \quad X \setminus A$ *dicht.*

Beweis. ($\Rightarrow$) Keine ϵ-Umgebung $U_\epsilon(x)$ kann vollständig in A liegen. Somit existiert ein $y \in (X \setminus A) \cap U_\epsilon(x)$, d.h. $X \setminus A$ ist dicht.

($\Leftarrow$) Ist $X \setminus A$ dicht in X, dann existiert für alle $U_\epsilon(x)$ ein $y \in (X \setminus A) \cap U_\epsilon(x)$. Folglich ist A nirgends dicht. □

Beispiel 6.4 i) Jeder Punkt $x \in X$ ist nirgends dicht.

ii) $S^1 \subset \mathbb{R}^2$ ist nirgends dicht.

iii) $\mathbb{Q} \subset \mathbb{R}$ ist abzählbare Vereinigung nirgends dichter Teilmengen, und dicht in $\mathbb{R}$.

Satz 6.5 (Baire, 2. Version) *Sei (X, d) vollständiger metrischer Raum und (A_n) eine Folge nirgends dichter abgeschlossener Teilmengen $A_n \subset X$. Dann enthält auch $\bigcup_{n\in\mathbb{N}} A_n$ keine inneren Punkte.*

Achtung: Es wird nicht behauptet, daß $\bigcup_{n\in\mathbb{N}} A_n$ nirgends dicht ist, siehe $\mathbb{Q} \subset \mathbb{R}$!

Beweis. Sei $\mathcal{O}_n := X \setminus A_n$. Dann ist $\bigcap_{n\in\mathbb{N}} \mathcal{O}_n$ dicht, also enthält die folgende Menge keine inneren Punkte:

$$X \setminus \Big(\bigcap_{n\in\mathbb{N}} (X \setminus A_n)\Big) = X \setminus \Big(X \setminus \Big(\bigcup_{n\in\mathbb{N}} A_n\Big)\Big) = \bigcup_{n\in\mathbb{N}} A_n\,. \qquad \square$$

Der Satz von Baire hat viele wichtige Anwendungen in der Funktionalanalysis. Eine davon ist:

Theorem 6.6 (Banach-Steinhaus: Prinzip der gleichmäßigen Beschränktheit) *Seien X ein Banach-Raum, Y ein normierter Vektorraum und I eine Indexmenge. Zu jedem $i \in I$ sei ein $T_i \in \mathcal{L}(X,Y)$ gegeben. Falls $\sup_{i\in I} \|T_i(x)\| < \infty$ für alle $x \in X$, so gilt sogar $\sup_{i\in I} \|T_i\| < \infty$.*

In Worten: Jede punktweise beschränkte Familie linearer stetiger Abbildungen ist auch gleichmäßig beschränkt (unter den gegebenen Voraussetzungen). Zur Bedeutung des Theorems erinnere man sich daran, daß für eine Folge stetiger Funktionen (f_n), die punktweise gegen eine Grenzfunktion f konvergiert, diese nicht notwendig stetig ist.

Beweis. Für $n \in \mathbb{N}$ setze $A_n := \{x \in X \ : \ \|T_i(x)\| \leq n \text{ für alle } i \in I\}$. Nach Voraussetzung ist $X = \bigcup A_n$. Die Abbildung $X \ni x \mapsto p_i(x) := \|T_i(x)\| \in \mathbb{R}_+$ ist stetig, so daß $A_n = \bigcap_{i\in I} p_i^{-1}([0,n])$ abgeschlossen ist. Da X innere Punkte hat, können nach dem Satz 6.5 von Baire nicht alle A_n nirgends dicht sein, d.h. es gibt $N \in \mathbb{N}$ und $U_\epsilon(x) \subset X$ mit $U_\epsilon(x) \subset A_N$. Mit $y \in A_N$ ist auch $-y \in A_N$, somit $x+u, x-u, u-x \in A_N$ für alle $u \in X$ mit $\|u\| < \epsilon$. Weiter ist A_N konvex: Sind $x, y \in A_N$ und $0 \leq \lambda \leq 1$, so gilt $\|T_i(\lambda x+(1-\lambda)y)\| \leq \lambda\|T_i(x)\|+(1-\lambda)\|T_i(y)\| \leq N$. Somit folgt

$$\begin{aligned}
\|u\| < \epsilon \qquad &\Rightarrow \qquad u = \frac{1}{2}(x+u) + \frac{1}{2}(u-x) \in A_N \\
&\Rightarrow \qquad \|T_i\| = \sup_{\frac{u}{\epsilon}\in B_X} \Big\|T_i\Big(\frac{u}{\epsilon}\Big)\Big\| \leq \frac{N}{\epsilon} < \infty \quad \text{für alle } i \in I\ .
\end{aligned}$$

Somit auch $\sup_{i\in I} \|T_i\| < \infty$. □

Der Satz von Banach-Steinhaus ist eine reine Existenzaussage. Es kann keine Aussage über den Wert von $\sup_{i\in I} \|T_i\| < \infty$ erhalten werden. Die Vollständigkeit von X ist unverzichtbar.

Folgerung 6.7 *Sei $(X, \|\ \|)$ normierter Vektorraum. Für $M \subset X$ sind äquivalent:*

i) *M ist beschränkt.*

ii) *$f(M)$ ist beschränkt für alle $f \in X'$.*

Beweis. i)⇒ii) ist klar. Umgekehrt genügt es wegen der Isometrie von i_X zu zeigen, daß $i_X(M) \subset X''$ beschränkt ist. Nach Voraussetzung gibt es für alle $f \in X'$ ein $c_f \geq 0$ mit $|f(m)| = |(i_X(m))(f)| \leq c_f$ für alle $m \in M$. Da X' Banach-Raum ist, folgt mit Banach-Steinhaus $\|i_X(m)\| = \|m\| < \infty$. □

Die analoge Aussage für X' erfordert die Vollständigkeit: Nur dann gilt: $M \subset X'$ ist genau dann beschränkt, wenn $i_X(M)$ beschränkt ist.

Folgerung 6.8 *Jede schwach-konvergente Folge in einem normierten Vektorraum $(X, \| \ \|)$ ist beschränkt.*

Beweis. (x_n) ist genau dann schwach-konvergent, wenn $f(x_n)$ konvergent in $\mathbb{K}$ für alle $f \in X'$. Insbesondere ist $(f(x_n))_{n\in\mathbb{N}}$ beschränkt. Die Behauptung folgt aus Folgerung 6.7 für $M = \{x_n \ : \ n \in \mathbb{N}\}$. □

Die Aussage gilt nicht für Netze (später einzuführen), da im Gegensatz zu Folgen ein in $\mathbb{K}$ konvergentes Netz nicht beschränkt sein muß.

Folgerung 6.9 *Sei X Banach-Raum, Y normierter Vektorraum und (T_n) eine Folge in $\mathcal{L}(X, Y)$, so daß $(T_n(x))_{n\in\mathbb{N}}$ in Y Norm-konvergent ist für alle $x \in X$. Sei $T : X \to Y$ definiert durch $Tx := \lim_{n\to\infty} T_n x$. Dann gilt $T \in \mathcal{L}(X, Y)$.*

Beweis. Linearität ergibt sich wie in Satz/Definition 1.14. Nach Voraussetzung ist $(T_n(x))_{n\in\mathbb{N}}$ konvergent und damit beschränkt, $\sup_n \|T_n(x)\| < \infty$ für alle $x \in X$. Nach dem Satz von Banach-Steinhaus gilt dann $\|T_n\| \leq C < \infty$ für alle $n \in \mathbb{N}$, somit

$$\|T\| = \sup_{x\in B_X} \|T(x)\| = \sup_{x\in B_X} \Big(\lim_{n\to\infty} \|T_n(x)\| \Big) \leq C \ . \qquad \square$$

Unter diesen Bedingungen heißt $(T_n)_{n\to\infty}$ *stark konvergent gegen T.*

Das Prinzip der gleichmäßigen Beschränktheit hat eine Reihe überraschender Konsequenzen, die in den Übungen behandelt werden:

- Die Menge der stetigen, aber nirgends differenzierbaren Funktionen auf $[a, b]$ ist dicht in $\mathcal{C}([a, b])$.
- Die Fourier-Reihe einer stetigen periodischen Funktion konvergiert nicht in jedem Punkt.

Eine weitere sehr wichtige Anwendung des Satzes von Baire ist:

Theorem 6.10 (Satz von der offenen Abbildung) *Seien X, Y Banach-Räume. Für $T \in \mathcal{L}(X, Y)$ sind äquivalent:*

i) *T ist surjektiv.*

ii) *T ist offen, d.h. $T(U)$ ist offen in Y für alle offenen $U \subset X$.*

Zum Beweis benötigen wir:

Lemma 6.11 *Seien X, Y normierte Vektorräume und $T \in \mathcal{L}(X, Y)$. Dann gilt*

$$T \textit{ offen} \quad \Leftrightarrow \quad \exists\, \delta > 0 \textit{ mit } U_\delta(0) \subset T(U_1(0)) \ .$$

Beweis. ($\Rightarrow$) ist die Definition der Offenheit.

($\Leftarrow$) Sei $\mathcal{O} \subset X$ offen, $x \in \mathcal{O}$. Dann gibt es ein $\epsilon > 0$ mit $U_\epsilon(x) \subset \mathcal{O}$. Aus der Voraussetzung und der Linearität von T folgt:

$$U_{\delta\epsilon}(T(x)) = T(x) + \epsilon U_\delta(0) \subset T(x) + \epsilon T(U_1(0)) = T(x + \epsilon U_1(0)) = T(U_\epsilon(x)) \ .$$

Damit ist T offen. □

Beweis von Theorem 6.10. Wir bezeichnen mit $V_\epsilon(y)$ die ϵ-Umgebung von $y \in Y$ in Y, mit $U_\epsilon(x)$ die ϵ-Umgebung von $x \in X$ in X.

ii)$\Rightarrow$i) Sei $0 \neq y \in Y$ beliebig und $\delta > 0$ wie in Lemma 6.11. Dann ist $\frac{\delta}{2\|y\|} y \in V_\delta(0)$, somit gibt es ein $\tilde{x} \in U_1(0)$ mit $T(\tilde{x}) = \frac{\delta}{2\|y\|} y$, folglich $T(\frac{2\|y\|}{\delta}\tilde{x}) = y$.

i)$\Rightarrow$ii) Es gilt $X = \bigcup_{n=1}^\infty U_n(0)$. Da T surjektiv ist, folgt $Y = \bigcup_{n=1}^\infty T(U_n(0)) \subset \bigcup_{n=1}^\infty \overline{T(U_n(0))}$. Sei $A_n := \overline{T(U_n(0))}$. Nach dem Satz 6.5 von Baire (hier wird die Vollständigkeit von Y gebraucht!) können nicht alle A_n nirgends dicht sein, d.h. es gibt ein $N \in \mathbb{N}$ und $V_{\delta_0}(y)$ mit $V_{\delta_0}(y) \subset A_N$. Mit $y \in A_N$ ist auch $-y \in A_N$ (klar für Punkte aus $T(U_n(0))$, überträgt sich auf Abschluß). Sei $v \in V_{\delta_0}(0)$, dann sind $y+v, y-v \in V_{\delta_0}(y)$, somit $v+y, v-y \in A_N$. Nun ist A_N konvex als Abschluß der konvexen Menge $T(U_N(0))$, also $v = \frac{1}{2}(v+y) + \frac{1}{2}(v-y) \in A_N$, d.h. $V_{\delta_0}(0) \subset A_N$. Aus der Linearität von T folgt: Zu jedem $\epsilon > 0$ gibt es ein $0 < \delta < \frac{\epsilon\delta_0}{N}$ mit $V_\delta(0) \subset \overline{T(U_\epsilon(0))}$.

Wir zeigen, daß für X vollständig sogar $V_\delta(0) \subset T(U_1(0))$ gilt für ein $\delta > 0$. Als Ergebnis des 1. Teils gibt es eine Nullfolge (δ_n) positiver Zahlen mit $V_{\delta_n}(0) \subset \overline{T(U_{\frac{1}{2^n}}(0))}$. Sei $y_1 \in V_{\delta_1}(0)$ beliebig, aber fest gewählt. Wir konstruieren per Induktion Folgen $(y_n)_{n\geq 1}$ in Y und $(x_n)_{n\geq 1}$ in X mit $y_n \in V_{\delta_n}(0)$, $x_n \in U_{\frac{1}{2^n}}(0)$ und $T(x_n) = y_n - y_{n+1}$.

- $n = 1$: Nach Definition des Abschlusses enthält jede Umgebung von y_1 einen Punkt aus $T(U_{\frac{1}{2^1}}(0))$. Somit gibt es ein $x_1 \in U_{\frac{1}{2^1}}(0)$ mit $T(x_1) \in V_{\delta_1}(0) \cap V_{\delta_2}(y_1)$. Setze $y_2 := y_1 - T(x_1)$, dann ist $\|y_2\| < \delta_2$, also $y_2 \in V_{\delta_2}(0)$.
- $n \to n+1$: Es sind $\{y_1, \dots, y_n\}$ sowie $\{x_1, \dots, x_{n-1}\}$ konstruiert, mit $y_n \in V_{\delta_n}(0)$. Nach Definition des Abschlusses enthält jede Umgebung von y_n einen Punkt aus $T(U_{\frac{1}{2^n}}(0))$. Somit gibt es ein $x_n \in U_{\frac{1}{2^n}}(0)$ mit $T(x_n) \in V_{\delta_n}(0) \cap V_{\delta_{n+1}}(y_n)$. Setze $y_{n+1} := y_n - T(x_n)$, dann ist $\|y_{n+1}\| < \delta_{n+1}$, also $y_{n+1} \in V_{\delta_{n+1}}(0)$.

Nach Konstruktion gilt $y_{n+1} = y_1 - T(x_1 + \dots + x_n)$ für $n \geq 1$. Wegen $\|x_n\| < \frac{1}{2^n}$ gilt $\sum_{n=1}^\infty \|x_n\| < 1 < \infty$. Da X vollständig ist, ist nach Lemma 1.12 die folgende Reihe in X konvergent: $x := \sum_{n=1}^\infty x_n$. Aus der Stetigkeit von T folgt:

$$T(x) = T\Big(\lim_{N\to\infty} \sum_{n=1}^N x_n\Big) = \lim_{N\to\infty} T\Big(\sum_{n=1}^N x_n\Big) = \lim_{N\to\infty} \big(y_1 - y_{N+1}\big) = y_1$$

wegen $\|y_{N+1}\| < \delta_{N+1} \to 0$. Somit gilt $y_1 \in T(U_1(0))$, und da $y_1 \in V_{\delta_1}(0)$ beliebig war, folgt $V_{\delta_1}(0) \subset T(U_1(0))$. Nach Lemma 6.11 ist T offen. $\square$

Folgerung 6.12 (Satz vom inversen Operator) *Seien X, Y Banach-Räume. Ist $T \in \mathcal{L}(X,Y)$ bijektiv, so ist $T^{-1} \in \mathcal{L}(Y,X)$ stetig.* $\square$

Auch diese Aussage ist bemerkenswert. Für allgemeine Homöomorphismen $T : X \to Y$ zwischen topologischen Räumen sind Stetigkeit von T, Bijektivität von T *und* Stetigkeit von T^{-1} zu fordern. Für lineare stetige Abbildungen zwischen Banach-Räumen reicht die Bijektivität!

Folgerung 6.13 (Homomorphiesatz) *Seien X, Y Banach-Räume und $T \in \mathcal{L}(X,Y)$ sei surjektiv. Definieren wir $\tilde{T} : X/\ker T \to Y$ durch $\tilde{T}([x]) := T(x)$, mit $[x] = x + \ker T$, so ist $\tilde{T}$ ein Homöomorphismus. Es gilt $\|\tilde{T}\| = \|T\|$.*

Beweis. Da T stetig, ist $\ker T \subset X$ abgeschlossener Untervektorraum. Nach Satz 1.16.iv) ist $X/\ker T$ ein Banach-Raum, nach Konstruktion ist $\tilde{T}$ linear und bijektiv. Unter Verwendung von Folgerung 6.12 bleibt zu zeigen, daß $\tilde{T}$ beschränkt ist. Sei $\epsilon > 0$ gegeben und $\|[x]\| = 1$. Dann gibt es nach Definition 1.15 ein $y \in \ker T$ mit $\|x - y\| \leq 1 + \epsilon$. Es folgt

$$\|\tilde{T}([x])\| = \|T(x)\| = \|T(x-y)\| \leq \|T\|(1+\epsilon)\ .$$

Da $\epsilon > 0$ beliebig war, gilt $\|\tilde{T}\| \leq \|T\|$.

Umgekehrt gilt wegen $T = \tilde{T} \circ \iota$ mit $\iota(x) = [x]$ und $\|\iota\| \leq 1$ nach Satz 1.16.iii) die Ungleichung $\|T\| \leq \|\tilde{T}\|$, d.h. $\|T\| = \|\tilde{T}\|$. $\square$

Folgerung 6.14 *Sei X ein Vektorraum, versehen mit zwei Normen $\|\ \|_1$ und $\|\ \|_2$, so daß X in beiden Normen vollständig ist. Sei $\mathcal{T}_i$, $i = 1, 2$ die durch $\|\ \|_i$ definierte Topologie auf X. Gilt $\mathcal{T}_1 \supset \mathcal{T}_2$, so gilt sogar $\mathcal{T}_1 = \mathcal{T}_2$.*

Beweis. Betrachte die bijektive stetige Abbildung $T = \mathrm{id} : (X, \mathcal{T}_1) \to (X, \mathcal{T}_2)$. Aus T^{-1} stetig nach Folgerung 6.12 folgt $\mathcal{T}_2 \supset \mathcal{T}_1$. $\square$

Wir kommen nun zu einer weiteren wichtigen Anwendung des Satzes von Baire.

Definition 6.15 Seien X, Y Mengen und $T : X \to Y$ eine Abbildung, dann heißt $G(T) := \{(x, T(x))\ :\ x \in X\} \subset X \times Y$ der *Graph* von T.

Sind X, Y Vektorräume und T linear, so ist $G(T)$ Untervektorraum von $X \times Y$. Sind X, Y normierte Vektorräume, dann wird durch $\|(x,y)\| := \|x\| + \|y\|$ auf dem kartesischen Produkt $X \times Y$ eine Norm erklärt. Sind X, Y vollständig, so auch $X \times Y$ in dieser Norm.

Theorem 6.16 (Satz vom abgeschlossenen Graphen) *Seien X, Y Banach-Räume und $T : X \to Y$ linear. Dann gilt:*

$$T \text{ stetig} \quad \Leftrightarrow \quad G(T) \text{ ist abgeschlossen in } X \times Y\ .$$

Beweis. ($\Rightarrow$) Sei $(x_n, T(x_n))_{n\in\mathbb{N}}$ eine beliebige Folge aus $G(T)$, die in $X \times Y$ konvergent ist. Insbesondere konvergiert (x_n) gegen ein $x \in X$. Da T stetig ist, konvergiert $(T(x_n))$ in Y gegen $T(x)$. Somit konvergiert $(x_n, T(x_n))_{n\in\mathbb{N}}$ gegen $(x, T(x)) \in G(T)$, d.h. $G(T)$ ist abgeschlossen.

($\Leftarrow$) Sei $G(T)$ abgeschlossener Untervektorraum von $X \times Y$. Nach Satz 1.16.vi) ist $G(T)$ Banach-Raum. Die Projektion auf die erste Komponente $\pi : G(T) \ni (x, T(x)) \mapsto x \in X$ ist linear, bijektiv und stetig. Nach dem Satz vom inversen Operator ist auch $\pi^{-1} : X \ni x \mapsto (x, T(x)) \in G(T)$ stetig. Ebenso ist die Projektion $P_Y : X \times Y \ni (x, y) \mapsto y \in Y$ stetig, folglich ist auch die Komposition $P_Y \circ \pi^{-1} : X \ni x \mapsto T(x) \in Y$ stetig. □

Der Satz vom abgeschlossenen Graphen ist die Grundlage für die Funktionalanalysis *unbeschränkter Operatoren*. Seien X, Y wieder Banach-Räume, dann betrachten wir lineare Abbildungen $T : \mathrm{dom}(T) \to Y$, wobei der *Definitionsbereich* $\mathrm{dom}(T) \subset X$ ein Untervektorraum ist. Für solche Operatoren heißt

$$G(T) := \{(x, T(x))\ :\ x \in \mathrm{dom}(T)\}$$

der Graph von T. Dieser ist ein Untervektorraum von $X \times Y$, und T heißt *abgeschlossen*, falls $G(T) \subset X \times Y$ abgeschlossen ist. Eine lineare Abbildung $S : \mathrm{dom}(S) \to Y$ heißt *Erweiterung* von T, falls $G(T) \subset G(S)$. In diesem Fall gilt $S\big|_{\mathrm{dom}(T)} = T$. Oft ist T nicht selbst abgeschlossen, es gibt aber eine abgeschlossene Erweiterung. Viele Resultate über lineare stetige Operatoren lassen sich auf abgeschlossene Operatoren $T : \mathrm{dom}(T) \to Y$ übertragen, deren Definitionsbereich $\mathrm{dom}(T)$ dicht in X ist. Zu beachten ist, daß für einen abgeschlossenen Operator T der Definitionsbereich $\mathrm{dom}(T)$ selbst nicht abgeschlossen sein muß. Ist $\mathrm{dom}(T) \subset X$ aber abgeschlossen, damit Banach-Raum, so liefert der Satz vom abgeschlossenen Graphen: $T : \mathrm{dom}(T) \to Y$ ist genau dann stetig, wenn $G(T)$ abgeschlossen ist.

Beispiel 6.17 Wir betrachten $X = Y = \mathcal{C}([-1,1])$ versehen mit der Supremumsnorm $\|x\|_\infty = \sup_{t\in[-1,1]} |x(t)|$. Die Ableitung $T : x \mapsto \dot{x}$ ist aufzufassen als Operator $T : \mathrm{dom}(T) \to Y$ mit $\mathrm{dom}(T) = \mathcal{C}^1(]-1,1[) \cap \mathcal{C}([-1,1])$. Konvergenz von $(x_n, \dot{x}_n)$ gegen $(x, y) \in X \times Y$ heißt gleichmäßige Konvergenz. Einer der Vertauschungssätze der Analysis liefert unter diesen Voraussetzungen die Differenzierbarkeit von $x = \lim_{n\to\infty} x_n$ sowie die Identität $y = \lim_{n\to\infty} \dot{x}_n = \dot{x}$. Damit sind $G(T)$ und T selbst abgeschlossen. ◁

Beispiel 6.18 Sei jetzt $X, Y = L^2([-1,1], \lambda)$ und $T : \mathrm{dom}(T) \to Y$ wie oben mit $\mathrm{dom}(T) = \mathcal{C}^1(]-1,1[) \cap \mathcal{C}([-1,1])$. Dann ist T nicht abgeschlossen: Betrachte die Folge $x_n(t) = \sqrt{t^2 + \frac{1}{n+1}}$ in $\mathrm{dom}(T)$, die gegen $|t| \in X$ konvergiert. Dabei konvergiert $\dot{x}_n(t) = \frac{t}{\sqrt{t^2+\frac{1}{n+1}}}$ gegen $\mathrm{sign}(t) \in L^2([-1,1],\lambda)$, aber $(|t|, \mathrm{sign}(t)) \notin G(T)$. Diese Folge ist nicht zulässig in Beispiel 6.17, da $(\dot{x}_n)$ nicht *gleichmäßig* konvergent ist.

Das Problem läßt sich beheben durch Übergang zu schwachen Ableitungen. Mit $\mathcal{C}_0^\infty([-1,1])$ werde der Raum der Testfunktionen bezeichnet, also der beliebig oft differenzierbaren Funktionen ϕ, die am Rand verschwinden: $\phi(-1) = \phi(1) = 0$. Eine Funktion $\dot{x}$ heißt *verallgemeinerte Ableitung* von x, falls

$$\int_{[-1,1]} d\lambda\; \phi(t)\dot{x}(t) = -\int_{[-1,1]} d\lambda\; \dot{\phi}(t)x(t)$$

für alle $\phi \in \mathcal{C}_0^\infty([-1,1])$. Betrachten wir $\mathrm{dom}(T) := \{x \in X \;:\; \dot{x} \in Y\}$, wobei $\dot{x}$ die verallgemeinerte Ableitung ist, dann ist die verallgemeinerte Ableitung $T : x \mapsto \dot{x}$ ein abgeschlossener Operator (ohne Beweis). $\triangleleft$

Teil II
Lokalkonvexe Vektorräume

7 Netze

Da es in dieser Allgemeinheit notwendig ist, führen wir zunächst die Netze ein:

Definition 7.1 Eine nichtleere Menge Λ heißt *gerichtet*, falls es eine Relation $\leq$ auf Λ gibt mit

i) $\lambda_1 \leq \lambda_2$ und $\lambda_2 \leq \lambda_3 \quad \Rightarrow \quad \lambda_1 \leq \lambda_3$.

ii) Für alle $\lambda_1, \lambda_2 \in \Lambda$ gibt es ein $\lambda_3 \in \Lambda$ mit $\lambda_1 \leq \lambda_3$ und $\lambda_2 \leq \lambda_3$.

Sei $(X, \mathcal{T})$ ein topologischer Raum und Λ eine gerichtete Menge. Ist $x_\lambda \in X$ für alle $\lambda \in \Lambda$, so heißt $(x_\lambda)_{\lambda\in\Lambda}$ ein *Netz* in X. Ein Netz $(x_\lambda)_{\lambda\in\Lambda}$ heißt *konvergent* gegen ein $x \in X$, falls es zu jeder Umgebung V von x ein $\mu \in \Lambda$ gibt mit $x_\lambda \in V$ für alle $\lambda \geq \mu$. Wir schreiben dann $x_\lambda \to x$.

Beispiel 7.2 i) Jede Folge $(x_n)_{n\in\mathbb{N}}$ ist ein Netz bezüglich der gerichteten Menge $\mathbb{N}$ mit der üblichen $\leq$-Relation, und die Konvergenzbegriffe stimmen überein.

ii) $\mathbb{R}$ ist gerichtete Menge mit der üblichen $\leq$-Relation. Ist $x_t \in X$ für alle $t \in \mathbb{R}$, so ist $(x_t)_{t\in\mathbb{R}}$ ein Netz in X. $\triangleleft$

In topologischen Räumen, in denen das erste Abzählbarkeitsaxiom (siehe Def. 4.19) gilt, kann man sich wegen der Abzählbarkeit der Umgebungsbasis bei Konvergenzbetrachtungen immer auf Folgen einschränken, ansonsten aber nicht!

Satz 7.3 *Sei $(X, \mathcal{T})$ ein topologischer Raum. Dann sind äquivalent:*

i) *X ist Hausdorffsch.*

ii) *Jedes in X konvergente Netz besitzt genau einen Grenzwert.*

Beweis. i)$\Rightarrow$ii) ist klar. Sei X nicht Hausdorffsch. Dann gibt es $x, y \in X$, $x \neq y$, so daß für alle Umgebungen U von x und V von y gilt $U \cap V \neq \varnothing$. Seien $\mathcal{U}_x$ bzw. $\mathcal{V}_y$ Umgebungsbasen von x bzw. y. Wir setzen

$$\Lambda := \mathcal{U}_x \times \mathcal{V}_y \,, \qquad (U_1, V_1) \leq (U_2, V_2) \quad \Leftrightarrow \quad U_1 \supset U_2 \text{ und } V_1 \supset V_2 \,.$$

Dann ist Λ gerichtet (Def. 7.1.i) ist klar, für Def. 7.1.ii) wähle eine Basismene aus dem Durchschnitt), und für jedes $\lambda = (U, V) \in \Lambda$ gibt es ein $x_\lambda \in U \cap V$. Wir zeigen: Das zugehörige Netz $(x_\lambda)_{\lambda\in\Lambda}$ konvergiert sowohl gegen x als auch gegen y, was den Widerspruch liefert. Seien U, V beliebige Umgebungen von x, y, dann gibt es $U_\mu \subset \mathcal{U}_x$ und $V_\mu \in \mathcal{V}_x$ mit $U_\mu \subset U$ und $V_\mu \subset V$. Setze $\mu := (U_\mu, V_\mu)$, dann gilt $U_\lambda \subset U_\mu$ und $V_\lambda \subset V_\mu$ für alle $\lambda \geq \mu$, somit $x_\lambda \in U_\mu \subset U$ und $x_\lambda \in V_\mu \subset V$ für alle $\lambda \geq \mu$. □

Wir können den Beweis genau dann auf Folgen reduzieren, falls die Umgebungsbasen abzählbar sind.

Satz 7.4 *Sei $(X, \mathcal{T})$ ein topologischer Raum. Für $A \subset X$ sind äquivalent:*

i) *A ist abgeschlossen.*

ii) *Sei $(x_\lambda)_{\lambda \in \Lambda}$ ein Netz in A, das gegen ein $x \in X$ konvergiert, so gilt $x \in A$.*

Ist $B \subset X$ eine beliebige Teilmenge, so gilt

iii) $x \in \bar{B}$ genau dann, wenn es ein Netz $(x_\lambda)_{\lambda \in \Lambda}$ in B gibt, das gegen x konvergiert.

Beweis. i)⇒ii) Angenommen, $x \notin A$. Da $X \setminus A$ offen ist, ist $X \setminus A$ Umgebung von x. Diese entält aber kein x_λ, Widerspruch.

ii)⇒i) Angenommen, $X \setminus A$ ist nicht offen. Dann gibt es ein $x \in X \setminus A$, so daß für alle Umgebungen U von x, insbesondere jene aus der Umgebungsbasis gilt: $U \cap A \neq \varnothing$. Also: zu jedem $U_\lambda \in \mathcal{U}_x$ gibt es ein $x_\lambda \in U_\lambda$ mit $x_\lambda \in A$. Wie in Satz 7.3 ist $(x_\lambda)_{\lambda \in \Lambda}$ mit $\Lambda = \mathcal{U}_x$ ein Netz in A, das gegen x konvergiert. Nach Voraussetzung ist $x \in A$, Widerspruch.

iii) klar für $x \in B$.

(⇒) Sei $x \notin B$ Häufungspunkt. Dann enthält jede Umgebung U von x, insbesondere jene aus der Umgebungsbasis $\mathcal{U}_x$, einen Punkt aus B. Wie zuvor wird dadurch ein gegen x konvergentes Netz erklärt.

(⇐) nach Definition des Grenzwertes enthält jede Umgebung ein $x_\lambda \in B$. Damit ist x Häufungspunkt von B. □

Satz 7.5 *Seien $(X, \mathcal{T})$ und $(Y, \mathcal{S})$ topologische Räume und $x \in X$. Für eine Abbildung $f : X \to Y$ sind äquivalent:*

i) *f ist stetig in x.*

ii) *Für jedes Netz $(x_\lambda)_{\lambda \in \Lambda}$ mit $x_\lambda \to x$ gilt $f(x_\lambda) \to f(x)$.*

Beweis. i)⇒ii) Sei V Umgebung von $f(x)$. Da f stetig ist, gibt es eine Umgebung $U \subset X$ von x mit $f(U) \subset V$. Ist $x_\lambda \to x$, so gibt es ein $\mu \in \Lambda$ mit $x_\lambda \in U$ für alle $\lambda \geq \mu$. Damit ist $f(x_\lambda) \in V$ für alle $\lambda \geq \mu$.

ii)⇒i) Angenommen, f ist nicht stetig. Dann gibt es eine Umgebung V von $f(x)$, so daß für alle Umgebungen U von x, insbesondere jene aus der Umgebungsbasis $\mathcal{U}_x$ gilt $f(U) \not\subset V$. Also: zu jedem $U_\lambda \in \mathcal{U}_x$ gibt es ein $x_\lambda \in U_\lambda$ mit $f(x_\lambda) \notin V$. Wie in Satz 7.3 ist $(x_\lambda)_{\lambda \in \Lambda}$ mit $\Lambda = \mathcal{U}_x$ ein Netz, das gegen x konvergiert. Aber $f(x_\lambda) \notin V$ für alle λ, Widerspruch. □

Somit kann man in allgemeinen topologischen Räumen wie in metrischen arbeiten, wenn man Folgen durch Netze ersetzt. Es gibt aber ein paar Fallen. Im Gegensatz zu Folgen muß ein konvergentes Netz nicht beschränkt sein: Betrachte $\Lambda = \mathbb{R} \setminus \{0\}$ und das Netz $(x_t)_{t\in\Lambda}$ in $\mathbb{R}$ mit $x_t = \frac{1}{t}$. Dann konvergiert $(x_t)_{t\in\Lambda}$ gegen 0, ist aber unbeschränkt.

8 Systeme von Halbnormen

In normierten Vektorräumen wurde die Topologie über die Norm induziert. Wir werden in diesem Abschnitt zeigen, daß auch eine Halbnorm oder ein System von Halbnormen auf einem Vektorraum diesen stets zu einem topologischen Vektorraum machen. Die entscheidende Frage ist dann, ob diese Topologie fein genug ist, um die Punkte zu trennen, d.h. ob der so konstruierte topologische Vektorraum Hausdorffsch ist. Nur so wird ein brauchbarer Konvergenzbegriff erhalten.

Beispiel 8.1 Sei X ein topologischer Raum. Auf dem Vektorraum der stetigen Funktionen $\mathcal{C}(X)$ betrachten wir die Familie $\mathcal{P} := \{p_x \ : \ x \in X\}$ der Halbnormen $p_x(f) = |f(x)|$. Es gilt $f = 0 \quad \Leftrightarrow \quad p_x(f) = 0 \ \forall x \in X$, so daß das System $\mathcal{P}$ zu einem sinnvollen Konvergenzbegriff führt: Eine Folge (f_n) in $\mathcal{C}(X)$ heißt bezüglich $\mathcal{P}$ konvergent gegen $f \in \mathcal{C}(X)$, falls $p_x(f_n - f) \to 0$ für alle $x \in X$. Offenbar ist das die punktweise Konvergenz der Folge (f_n). $\triangleleft$

Beispiel 8.2 Stetige Funktionen auf ganz $\mathbb{R}$ müssen nicht beschränkt sein. Durch $\{p_n\}_{n=1}^{\infty}$ mit $p_n(f) := \sup_{x\in[-n,n]} |f(x)|$ wird ein System von Halbnormen auf $\mathcal{C}(\mathbb{R})$ erklärt. Wieder gilt $f = 0 \quad \Leftrightarrow \quad p_n(f) = 0 \ \forall n \in \mathbb{N} \setminus \{0\}$, so daß ein sinnvoller Konvergenzbegriff erhalten wird (gleichmäßige Konvergenz auf Kompakta). $\triangleleft$

Definition 8.3 Sei X ein Vektorraum über $\mathbb{K}$ und $\mathcal{P} = \{p_i \ : \ i \in I\}$ ein System von Halbnormen $p_i : X \to \mathbb{R}$ auf X. Für $x \in X$, $p_1, \ldots, p_n \in \mathcal{P}$ und $\epsilon > 0$ setze

$$U(x; p_1, \ldots, p_n, \epsilon) := \{y \in X \ : \ p_k(x - y) < \epsilon \ \forall k = 1, \ldots, n\} .$$

Eine (nichtleere) Teilmenge $\mathcal{O} \subset X$ heißt *offen bezüglich* $\mathcal{P}$ (oder $\mathcal{P}$*-offen*), falls es zu jedem $x \in \mathcal{O}$ ein $\epsilon > 0$ und $p_1, \ldots, p_n \in \mathcal{P}$ gibt mit $U(x; p_1, \ldots, p_n, \epsilon) \subset \mathcal{O}$.

Ist $\mathcal{P} = \{\| \ \|\}$ für eine Norm $\| \ \|$ auf X, so ist $U(x; \| \ \|, \epsilon) = U_\epsilon(x)$ genau die ϵ-Umgebung von X.

Satz 8.4 *Sei $\mathcal{P}$ ein System von Halbnormen auf einem $\mathbb{K}$-Vektorraum X und $\mathcal{T} := \{\mathcal{O} \subset X$ offen bzgl. $\mathcal{P}\}$. Dann ist $(X, \mathcal{T})$ ein topologischer Vektorraum. Ferner gilt:*

i) $\mathcal{U}_x := \{U(x; p_1, \ldots, p_n, \epsilon) \ : \ \epsilon > 0\,,\ n \in \mathbb{N}\,,\ p_1, \ldots, p_n \in \mathcal{P}\}$ *ist Umgebungsbasis von $x \in X$.*

ii) *$\mathcal{T}$ ist die gröbste Topologie, in der alle $p \in \mathcal{P}$ stetig sind.*

iii) *Ein Netz* $(x_\lambda)_{\lambda\in\Lambda}$ *in* X *konvergiert genau dann in* $\mathcal{T}$ *gegen* $x \in X$*, wenn* $p(x_\lambda - x) \to 0$ *für alle* $p \in \mathcal{P}$.

Beweis. Zu zeigen sind zunächst die Topologie-Axiome. Seien $\mathcal{O}_1, \mathcal{O}_2 \in \mathcal{T}$ und $x \in \mathcal{O}_1 \cap \mathcal{O}_2 \neq \varnothing$. Dann gibt es $\epsilon_1, \epsilon_2 \geq 0$, $n, m \in \mathbb{N}$ und $p_1, \ldots, p_n, \tilde{p}_1, \ldots, \tilde{p}_m$ mit $U(x; p_1, \ldots, p_n, \epsilon_1) \subset \mathcal{O}_1$ und $U(x; \tilde{p}_1, \ldots, \tilde{p}_m, \epsilon_2) \subset \mathcal{O}_2$. Sei $\epsilon := \min(\epsilon_1, \epsilon_2)$, so folgt $U(x; p_1, \ldots, p_n, \tilde{p}_1, \ldots, \tilde{p}_m, \epsilon) \subset \mathcal{O}_1 \cap \mathcal{O}_2$. Die restlichen Axiome sind klar.

i) Die $U(x; p_1, \ldots, p_n, \epsilon)$ sind selbst offen: Für $y \in U(x; p_1, \ldots, p_n, \epsilon)$ wähle ein $0 < \delta < \epsilon - \max_{i=1\ldots,n}(p_i(x-y))$, dann ist nach Dreiecksungleichung $U(y; p_1, \ldots, p_n, \delta) \subset U(x; p_1, \ldots, p_n, \epsilon)$. Für $\mathcal{O} \in \mathcal{T}$ gilt $\mathcal{O} = \bigcup_{x\in\mathcal{O}} U(x; p_1^x, \ldots, p_{n_x}^x, \epsilon_x)$ mit geeigneten $\epsilon_x > 0$, $n_x \in \mathbb{N}$ und $p_1^x, \ldots, p_{n_x}^x \in \mathcal{P}$. Somit ist $\{\mathcal{U}_x : x \in X\}$ Basis von $\mathcal{T}$, insbesondere $\mathcal{U}_x$ Umgebungsbasis von x.

ii) folgt sofort aus der Definition der Initialtopologie, soll der Vollständigkeit halber aber direkt bewiesen werden. Zunächst überprüfen wir die Stetigkeit der p_i: Sei $(x_\lambda)_{\lambda\in\Lambda}$ ein beliebiges gegen $x \in X$ konvergentes Netz in X und $\epsilon > 0$ beliebig. Dann ist $U(x; p, \epsilon)$ offene Umgebung von x. Somit gibt es ein $\mu \in \Lambda$, so daß für alle $\lambda \geq \mu$ gilt $x_\lambda \in U(x; p, \epsilon)$. Damit ist $p(x_\lambda - x) < \epsilon$, und wegen $|p(x) - p(y)| \leq p(x-y)$ konvergiert $p(x_\lambda)$ gegen $p(x)$, d.h. p ist stetig.

Sei $\tilde{\mathcal{T}}$ eine beliebige Topologie auf X, so daß alle $p \in \mathcal{P}$ stetig sind. Dann gilt $U(x, p_1, \ldots, p_n, \epsilon) = \bigcap_{i=1}^n p_i^{-1}(]-\infty, \epsilon[)$, und wegen der Stetigkeit der p_i ist $U(x, p_1, \ldots, p_n, \epsilon)$ offen in $\tilde{\mathcal{T}}$, d.h. $\mathcal{T} \subset \tilde{\mathcal{T}}$, d.h. $\mathcal{T}$ ist gröber als $\tilde{\mathcal{T}}$.

iii) Die Richtung $x_\lambda \to x \quad \Rightarrow \quad p(x_\lambda - x) \to 0$ ist bereits in ii) gezeigt. Sei umgekehrt $\mathcal{O} \subset X$ eine beliebige Umgebung von x. Dann gibt es ein $U(x; p_1, \ldots, p_n, \epsilon) \subset \mathcal{O}$. Wegen $p(x_\lambda - x) \to 0$ für alle $p \in \mathcal{P}$ gibt es $\mu_1, \ldots, \mu_n \in \Lambda$ mit $p_i(x_{\lambda_i} - x) < \epsilon$ für alle $\lambda_i \geq \mu_i$ und alle $i = 1\ldots, n$. Wähle $\mu = \max(\mu_1, \ldots, \mu_n)$, dann ist $x_\lambda \in U(x; p_1, \ldots, p_n, \epsilon) \subset \mathcal{O}$ für alle $\lambda \geq \mu$, somit $x_\lambda \to x$.

iv) Stetigkeit der Addition: Jede Umgebung U von $x + y$ enthält ein $U(x + y; p_1, \ldots, p_n, \epsilon)$. Dann ist $U_1 := U(x; p_1, \ldots, p_n, \frac{\epsilon}{2})$ Umgebung von x und $U_2 := U(y; p_1, \ldots, p_n, \frac{\epsilon}{2})$ Umgebung von y, und nach Dreiecksungleichung gilt $U_1 + U_2 \subset U(x + y; p_1, \ldots, p_n, \epsilon) \subset U$.

v) Stetigkeit der skalaren Multiplikation: Jede Umgebung U von λx, $\lambda \in \mathbb{K}$, enthält ein $U(\lambda x; p_1, \ldots, p_n, \epsilon)$. Sei $a := \max(p_1(x), \ldots, p_n(x), 1)$ und $b := |\lambda| + \frac{\epsilon}{2a} > 0$, dann gilt für $\mu \in U_{\frac{\epsilon}{2a}}(\lambda) \subset \mathbb{K}$ und $y \in U(x; p_1, \ldots, p_n, \frac{\epsilon}{2b})$:

$$\begin{aligned} p_i(\lambda x - \mu y) &\leq p_i((\lambda - \mu)x) + p_i(\mu(x - y) = |\lambda - \mu| p_i(x) + |\mu| p_i(x - y) \\ &\leq |\lambda - \mu| p_i(x) + (|\mu - \lambda| + |\lambda|) p_i(x - y) \\ &< \frac{\epsilon}{2a} \cdot a + (\frac{\epsilon}{2a} + |\lambda|) \cdot \frac{\epsilon}{2b} \leq \epsilon\,, \end{aligned}$$

somit $U_{\frac{\epsilon}{2a}}(\lambda) \cdot U(x; p_1, \ldots, p_n, \frac{\epsilon}{2b}) \subset U(\lambda x; p_1, \ldots, p_n, \epsilon) \subset U$. □

Aus ii) folgt, daß zwei Halbnormensysteme $\mathcal{P}$, $\mathcal{Q}$ genau dann die gleiche Topologie erzeugen, wenn alle $p \in \mathcal{P}$ bezüglich $\mathcal{Q}$ stetig sind und alle $q \in \mathcal{Q}$ bezüglich $\mathcal{P}$ stetig sind.

Definition 8.5 Sei X ein Vektorraum, $\mathcal{P}$ ein Halbnormensystem auf X und $\mathcal{T}$ die durch $\mathcal{P}$ erzeugte Topologie. Ist $\mathcal{T}$ eine Hausdorff-Topologie auf X, so heißt $(X, \mathcal{T})$ *lokalkonvexer (topologischer) Vektorraum* und $\mathcal{T}$ *lokalkonvexe Topologie auf* X.

Beispiel 8.6

i) Sei X normierter Vektorraum, dann ist die schwache Topologie auf X die lokalkonvexe Topologie erzeugt durch das Halbnormensystem $\mathcal{P} = \{p_{x'} \; : \; x' \in X'\}$, mit $p_{x'}(x) = |x'(x)|$. Die Hausdorff-Eigenschaft ist in Satz 4.8 gezeigt.

ii) Sei X normierter Vektorraum, dann ist die schwach-*-Topologie auf dem Dualraum X' die lokalkonvexe Topologie erzeugt durch das Halbnormensystem $\mathcal{P} = \{p_x \; : \; x \in X\}$, mit $p_x(x') = |x'(x)|$. Die Hausdorff-Eigenschaft folgt in Analogie zu Satz 4.8. $\triangleleft$

Zur Vorbereitung eines Satzes über Kriterien zur Hausdorff-Eigenschaft zeigen wir:

Lemma 8.7 *Sei X Vektorraum, $\mathcal{P}$ ein Halbnormensystem auf X und $\mathcal{T}$ die durch $\mathcal{P}$ induzierte Topologie. Dann gilt $\overline{\{0\}} = \bigcap_{p \in \mathcal{P}} p^{-1}(\{0\})$, und $\mathcal{N}_{\mathcal{P}} := \overline{\{0\}}$ ist abgeschlossener Untervektorraum von X.*

Beweis. Nach der Bemerkung vor Lemma 1.6 gilt $\bar{A} = \bigcap_{V \in \mathcal{U}_0}(A+V)$ für beliebige Teilmengen $A \subset X$, insbesondere

$$\begin{aligned}\overline{\{0\}} &= \bigcap \{U(0; p_1, \ldots, p_n, \epsilon) \; : \; \epsilon > 0, n \in \mathbb{N}, p_1, \ldots, p_n \in \mathcal{P}\} \\ &= \bigcap \{U(0; p, \epsilon) \; : \; \epsilon > 0, p \in \mathcal{P}\} \\ &= \bigcap_{p \in \mathcal{P}} \Big(\bigcap_{\epsilon > 0} \{x \in X \; : \; p(x) < \epsilon\} \Big) \\ &= \bigcap_{p \in \mathcal{P}} p^{-1}(\{0\}) \, .\end{aligned}$$

Sind $x, y \in \mathcal{N}_{\mathcal{P}}$ und $\lambda, \mu \in \mathbb{K}$, so ist $0 \leq p(\lambda x + \mu y) \leq |\lambda| p(x) + |\mu| p(y) = 0$ für alle $p \in \mathcal{P}$. $\square$

Satz 8.8 *Sei X ein Vektorraum, $\mathcal{P}$ ein Halbnormensystem auf X. Für die durch $\mathcal{P}$ auf X erzeugte Topologie $\mathcal{T}$ sind äquivalent:*

i) *$(X, \mathcal{T})$ ist Hausdorff-Raum.*

ii) *$p(x) = 0$ für alle $p \in \mathcal{P} \quad \Leftrightarrow \quad x = 0$.*

iii) *$\{0\}$ ist abgeschlossen.*

Beweis. i⇒ii) Angenommen, es gibt ein $0 \neq x \in X$ mit $p(x) = 0$ für alle $p \in \mathcal{P}$. Dann ist $x \in U(0, p_1, \ldots, p_n, \epsilon)$ für alle $\epsilon > 0$, $n \in \mathbb{N}$ und $p_1, \ldots, p_n \in \mathcal{P}$, also liegt x in jeder Umgebung von 0 und kann von 0 nicht durch offene Mengen getrennt werden, Widerspruch.

ii)⇒i) Angenommen, X ist nicht Hausdorffsch. Nach Satz 7.3 gibt es $x, y \in X$, $x \neq y$ sowie ein Netz $(x_\lambda)_{\lambda \in \Lambda}$ in X mit $x_\lambda \to x$ und $x_\lambda \to y$. Dann folgt $p(x_\lambda - x) \to 0$ und $p(x_\lambda - y) \to 0$, somit $p(x - y) < \epsilon$ für alle $\epsilon > 0$. Nach Voraussetzung ist dann $x = y$, Widerspruch.

ii)⇔iii) Zunächst gilt ii) $\Leftrightarrow$ $\{0\} = \bigcap_{p \in \mathcal{P}} p^{-1}(\{0\})$. Aus Lemma 8.7 folgt ii)⇔ $(\{0\} = \overline{\{0\}})$. □

Beispiel 8.9 i) Sind X, Y normierte Vektorräume, dann heißt die auf $\mathcal{L}(X, Y)$ durch das Halbnormensystem $\mathcal{P} = \{p_x : x \in X\}$ mit $p_x(T) := \|T(x)\|$ erzeugte lokalkonvexe Topologie die *starke Operatortopologie*. Ist $\|T(x)\| = 0$, so $T(x) = 0$ für alle $x \in X$, somit $T = 0$, und $\mathcal{T}$ trennt die Punkte. Ein Netz $(T_\lambda)_{\lambda \in \Lambda}$ in $\mathcal{L}(X, Y)$ konvergiert somit genau dann gegen $T \in \mathcal{L}(X, Y)$, wenn $T_\lambda(x) \to T(x)$ für alle $x \in X$. Die starke Operatortopologie ist gröber als die (Operator-)Norm-Topologie.

ii) Sind X, Y normierte Vektorräume, dann heißt die auf $\mathcal{L}(X, Y)$ durch das Halbnormensystem $\mathcal{P} = \{p_{x,y'} : x \in X\,,\ y' \in Y'\}$ mit $p_{x,y'}(T) := |y'(T(x))|$ erzeugte lokalkonvexe Topologie die *schwache Operatortopologie*. Da Y' die Punkte aus Y trennt, ist $p_{x,y'}(T) = 0$ für alle $p_{x,y'} \in \mathcal{P}$ genau dann, wenn $T = 0$, somit ist die schwache Operatortopologie Hausdorffsch. Ein Netz $(T_\lambda)_{\lambda \in \Lambda}$ in $\mathcal{L}(X, Y)$ konvergiert somit genau dann gegen $T \in \mathcal{L}(X, Y)$, wenn $y'(T_\lambda(x)) \to y'(T(x))$ für alle $x \in X$ und alle $y' \in Y'$. Die schwache Operatortopologie ist gröber als starke Operatortopologie und noch gröber als die (Operator-)Norm-Topologie.

iii) Die Topologie der punktweisen Konvergenz auf $\mathcal{C}(X)$ nach Beispiel 8.1 ist eine lokalkonvexe Topologie.

iv) Die Topologie der gleichmäßigen Konvergenz auf Kompakta auf $\mathcal{C}(\mathbb{R})$ nach Beispiel 8.2 ist eine lokalkonvexe Topologie.

v) Auf $\mathcal{C}^\infty(\mathbb{R}^n)$ definiert $p_{K,\alpha}(x) := \sup_{t \in K} |(\partial^\alpha x)(t)|$ für $K \subset \mathbb{R}^n$ und $\alpha \in \mathbb{N}^n$ (Multi-Index) eine Halbnorm. Dann definiert $\mathcal{P} = \{p_{K,\alpha} : K \subset \mathbb{R}^n \text{ kompakt}\,,\ \alpha \in \mathbb{N}^n\}$ ein Halbnormensystem, welches $\mathcal{C}^\infty(\mathbb{R}^n)$ zu einem lokalkonvexen Vektorraum macht.

vi) Der Schwartz-Raum $\mathcal{S}(\mathbb{R}^n)$ ist der Vektorraum der beliebig oft differenzierbaren Funktionen, welche im Unendlichen schneller als jede Potenz verschwinden. Durch das Halbnormensystem $\mathcal{P} = \{p_{m,\alpha} : m \in \mathbb{N}\,,\ \alpha \in \mathbb{N}^n\}$ mit $p_{m,\alpha}(x) := \sup_{t \in \mathbb{R}^n} |(1 + \|t\|^m)(\partial^\alpha x)(t)|$ wird $\mathcal{S}(\mathbb{R}^n)$ zu einem lokalkonvexen Vektorraum. ◁

Bemerkung 8.10 Ähnlich wie beim Schritt von $\mathcal{L}^p(X,\mu)$ zu $L^p(X,\mu)$ läßt sich eine durch ein Halbnormsystem induzierte Topologie "Hausdorffisieren" durch Übergang zu $\tilde{X} := X/\mathcal{N}_\mathcal{P}$ topologisiert durch das Halbnormensystem $\tilde{P} = \{\tilde{p} \,:\, p \in \mathcal{P}\}$ mit $\tilde{p}([x]) := p(x)$. Denn sei $n \in \mathcal{N}_\mathcal{P}$ beliebig, so gilt $p(x+n) \leq p(x) + p(n) = p(x)$ und $p(x) = p(x+n-n) \leq p(x+n) + p(-n) = p(x+n)$, so daß $\tilde{p}$ wohldefiniert ist. Nach Konstruktion gilt $\bigcap_{\tilde{p}\in\tilde{P}} \tilde{p}^{-1}(\{0\}) = \mathcal{N}_\mathcal{P} = \{[0]\}$, damit ist $\tilde{X}$ mit der aus $\tilde{P}$ induzierten Topologie ein Hausdorff-Raum.

9 Konvexität

Definition 9.1 Sei X Vektorraum über $\mathbb{K}$. Eine (nichtleere) Teilmenge $A \subset X$ heißt

i) *absolutkonvex*, falls für alle $x_1, \ldots, x_n \in A$ und $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ mit $\sum_{i=1}^n |\lambda_i| \leq 1$ gilt: $\sum_{i=1}^n \lambda_i x_i \in A$.

ii) *kreisförmig*, falls für $x \in A$ und $\lambda \in \mathbb{K}$ mit $|\lambda| \leq 1$ gilt $\lambda x \in A$.

Jede absolutkonvexe Menge ist konvex, nicht aber umgekehrt. Jede absolutkonvexe Menge enthält $0 \in X$.

Lemma 9.2 *Eine Teilmenge $A \subset X$ ist genau dann absolutkonvex, wenn sie konvex und kreisförmig ist.*

Beweis. ($\Rightarrow$) ist klar, für die Umkehrung schreibe (es kann $\lambda_i \neq 0$ angenommen werden) $\sum_{i=1}^n \lambda_i x_i = \sum_{i=1}^n |\lambda_i| \frac{\lambda_i}{|\lambda_i|} x_i$. □

In $\mathbb{C}$ sind die offenen oder abgeschlossenen Kreischeiben mit Mittelpunkt 0 die einzigen absolutkonvexen Mengen.

Lemma 9.3 *Ist $\mathcal{P}$ ein Halbnormensystem auf X, so ist jede Nullumgebung $U(0; p_1, \ldots, p_n, \epsilon)$ absolutkonvex.*

Beweis. Sind $x_1, \ldots, x_k \in U(0; p_1, \ldots, p_n, \epsilon)$ und $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ mit $\sum_{j=1}^k |\lambda_j| \leq 1$, so folgt nach Dreiecksungleichung $p_i(\sum_{j=1}^k \lambda_j x_j) \leq \sum_{j=1}^k |\lambda_j| p_i(x_j) < \epsilon$ für alle $i = 1, \ldots, n$. □

Ist A absolutkonvex, so ist $x + A$ konvex. Somit besitzt jeder Punkt $x \in X$ des von einem Halbnormensystem $\mathcal{P}$ auf X erzeugten topologischen Vektorraums eine Umgebungsbasis aus konvexen Mengen (daher der Name "lokalkonvexer Vektorraum"). Es gilt sogar die Umkehrung:

Satz 9.4 *Sei $(X, \mathcal{T})$ ein topologischer Vektorraum. Dann sind äquivalent:*

i) *Jedes $x \in X$ besitzt eine Umgebungsbasis aus konvexen Mengen.*

ii) *$0 \in X$ besitzt Umgebungsbasis aus absolutkonvexen Mengen.*

iii) *Es gibt ein Halbnormensystem $\mathcal{P}$ auf X, welches $\mathcal{T}$ erzeugt.*

Beweis. iii)⇒i) ist Lemma 9.3.

i)⇒ii) Sei V konvexe Umgebung von 0. Wegen der Stetigkeit der skalaren Multiplikation gibt es $\epsilon > 0$ und eine Umgebung U_1 von 0 mit $U_2 := \overline{U_\epsilon(0)} \cdot U_1 = \{\lambda u \ : \ |\lambda| \leq \epsilon\,,\ u \in U_1\} \subset V$. Da V konvex ist, liegt auch die konvexe Hülle $\mathrm{conv}(U_2)$ von U_2 in V, wobei

$$\mathrm{conv}(B) := \Big\{ \sum_{i=1}^{l} \lambda_i x_i \ : \ l \in \mathbb{N}\ ,\ x_1, \dots, x_l \in B\ ,\ \lambda_1, \dots, \lambda_l \in [0,1]\ ,\ \sum_{i=1}^{l} \lambda_i = 1\}$$

für eine Teilmenge $B \subset X$. Da U_2 kreisförmig ist, ist auch $\mathrm{conv}(U_2)$ kreisförmig, damit nach Lemma 9.2 absolutkonvexe Umgebung von 0.

ii)⇒iii) Sei U absolutkonvexe Umgebung von 0. Durch Übergang zu $\mathrm{int}(U)$ kann U als offen angenommen werden (man zeigt, daß mit U absolutkonvex auch $\mathrm{int}(U)$ absolutkonvex ist) nehmen. Nach Lemma 3.16 ist das zugehörige Minkowski-Funktional $p_U : X \to \mathbb{R}_+$ mit $p_U(x) = \inf\{\lambda > 0 \ : \ \frac{x}{\lambda} \in U\}$ sublinear, für absolutkonvexes U sogar Halbnorm: Denn für $\mu \in \mathbb{K} \setminus 0$ gilt

$$p_U(\mu x) = \inf\{\lambda > 0 \ : \ \tfrac{\mu x}{\lambda} \in U\} = \inf\{\lambda > 0 \ : \ \tfrac{|\mu| x}{\lambda} \in \tfrac{|\mu|}{\mu} U\} = |\mu| p_U(x)\ .$$

Ist $\mathcal{U}_0$ eine Umgebungsbasis von 0 aus absolutkonvexen offenen Mengen, so setzen wir $\mathcal{P} = \{p_U \ : U \in \mathcal{U}_0\}$.

Sei $\tilde{\mathcal{T}}$ die von $\mathcal{P}$ erzeugte Topologie auf X. Wir zeigen $\tilde{\mathcal{T}} = \mathcal{T}$, genauer, daß $\mathrm{id} : (X, \mathcal{T}) \to (X, \tilde{\mathcal{T}})$ und das Inverse stetig in 0 sind (was nach Satz 1.7 genügt). Sei V Umgebung von 0 bezüglich $\tilde{\mathcal{T}}$. Dann gibt es $\epsilon > 0$, $n \in \mathbb{N}$ und $U_1, \dots, U_n \in \mathcal{U}_0$ mit $U(0, p_{U_1}, \dots, p_{U_n}, \epsilon) \subset V$. Nach Lemma 3.16.ii) gilt $U_i = p_{U_i}^{-1}([0,1[)$, damit $\epsilon U_i = p_{U_i}^{-1}([0,\epsilon[)$ und $U(0, p_{U_1}, \dots, p_{U_n}, \epsilon) = \bigcap_{i=1}^n (\epsilon U_i)$. Wegen der $\mathcal{T}$-Stetigkeit der skalaren Multiplikation ist auch ϵU offen, somit ist $U(0, p_{U_1}, \dots, p_{U_n}, \epsilon) \in \mathcal{T}$. Damit ist die Stetigkeit von $\mathrm{id} : (X, \mathcal{T}) \to (X, \tilde{\mathcal{T}})$ in 0 gezeigt. Umgekehrt sei $U \in \mathcal{U}_0$, dann ist $U = U(0, p_U, 1)$ offen in $\tilde{\mathcal{T}}$, somit $\mathrm{id} : (X, \tilde{\mathcal{T}}) \to (X, \mathcal{T})$ stetig in 0. □

Zur Vorbereitung der Stetigkeit von Abbildungen zwischen lokalkonvexen Vektorräumen benötigen wir:

Lemma 9.5 *Sei $\mathcal{P}$ ein Halbnormensystem auf X und q eine weitere Halbnorm auf X. Dann sind äquivalent:*

i) *q ist stetig bzgl. $\mathcal{P}$.*

ii) *Es gibt $c \geq 0$ und $p_1, \dots, p_n \in \mathcal{P}$ mit $q(x) \leq c \max(p_1(x), \dots, p_n(x))$ für alle $x \in X$.*

iii) *Es gibt $\lambda_1, \dots, \lambda_n \geq 0$ und $p_1, \dots, p_n \in \mathcal{P}$ mit $q(x) \leq \sum_{i=1}^n \lambda_i p_i(x)$ für alle $x \in X$.*

iv) *Es gibt eine bezüglich $\mathcal{P}$ stetige Halbnorm $\tilde{p}$ mit $q(x) \leq \tilde{p}(x)$ für alle $x \in X$.*

Beweis. i)⇒ii) $q^{-1}(]-\infty, 1[)$ ist $\mathcal{P}$-offen und enthält $0 \in X$, somit gibt es $U(0; p_1, \ldots, p_n, \epsilon) \subset q^{-1}(]-\infty, 1[)$. Ist $\max(p_i(x), \ldots, p_n(x)) = 0$, so ist $\lambda x \in q^{-1}(]-\infty, 1[)$ für alle $\lambda \in \mathbb{K}$, was nur für $q(x) = 0$ möglich ist, und ii) gibt keine Einschränkung an c. Sei also $m_x := \max(p_i(x), \ldots, p_n(x)) > 0$. Dann ist $p_i(\frac{\epsilon}{2m_x}x) < \epsilon$ für alle $i = 1, \ldots, m$, somit $q(\frac{\epsilon}{2m_x}x) < 1$, und dann $0 \leq q(x) < \frac{2}{\epsilon}m_x$. Damit ist $c := \frac{2}{\epsilon}$ unabhängig von x.

ii)⇒iii) $\max(p_i(x), \ldots, p_n(x)) \leq \sum_{i=1}^n p_i(x)$

iii)⇒iv) $\tilde{p} := \sum_{i=1}^n \lambda_i p_i(x)$ ist $\mathcal{P}$-stetige Halbnorm.

iv)⇒i) Sei $(x_\lambda)_{\lambda \in \Lambda}$ ein gegen $x \in X$ konvergentes Netz in x. Dann konvergiert $x_\lambda - x$ gegen 0, somit $\tilde{p}(x_\lambda - x) \to 0$. Schließlich gilt $|q(x_\lambda) - q(x)| \leq q(x_\lambda - x) \leq \tilde{p}(x_\lambda - x) \to 0$. Damit ist q stetig. □

Satz 9.6 *Seien X bzw. Y lokalkonvexe Vektorräume, deren Topologien durch Halbnormensysteme $\mathcal{P}$ bzw. $\mathcal{Q}$ induziert seien. Für eine lineare Abbildung $T : X \to Y$ sind äquivalent:*

i) *T ist stetig.*

ii) *Für alle $q \in \mathcal{Q}$ ist $q \circ T : X \to \mathbb{R}_+$ stetig.*

iii) *Für jedes $q \in \mathcal{Q}$ gibt es ein $c \geq 0$ und $p_1, \ldots, p_n \in \mathcal{P}$ mit $q(T(x)) \leq c \max(p_1(x), \ldots, p_n(x)$ für alle $x \in X$.*

Beweis. i)⇒ii) T und q sind stetig, somit auch $q \circ T$.

ii)⇔iii) Offenbar ist $q \circ T$ Halbnorm auf X. Dann folgt die Äquivalenz aus Lemma 9.5.

ii)⇒i) Sei $(x_\lambda)_{\lambda \in \Lambda}$ ein gegen x konvergentes Netz in X. Dann konvergiert $q(T(x_\lambda - x))$ gegen 0 für alle $q \in \mathcal{Q}$. Nach Satz 8.4.iii) konvergiert $T(x_\lambda - x)$ gegen 0, d.h. $T(x_\lambda)$ konvergiert gegen $T(x)$. Damit ist T stetig. □

Folgerung/Definition 9.7 *Sei X lokalkonvexer Vektorraum bezüglich $\mathcal{P}$. Eine lineare Abbildung $f : X \to \mathbb{K}$ ist genau dann stetig, wenn es $c \geq 0$ und $p_1, \ldots, p_n \in \mathcal{P}$ gibt mit $|f(x)| \leq c \max(p_1(x), \ldots, p_n(x)$ für alle $x \in X$.*

Der Dualraum X' von X ist der Vektorraum aller bezüglich $\mathcal{P}$ stetigen linearen Abbildungen $f : X \to \mathbb{K}$.

Mit $\mathcal{L}(X, Y)$ werde der Vektorraum aller linearen stetigen Abbildungen zwischen lokalkonvexen Vektorräumen X, Y bezeichnet.

Beispiel 9.8 Sei $(X, \|\ \|)$ normierter Vektorraum, $(X, \|\ \|)'$ sein Dualraum bezüglich der Norm-Topologie. Dieser definiert einen lokalkonvexen Vektorraum $(X, \mathcal{P} = \{p_{x'}\})$, von dem wieder der Dualraum $(X, \mathcal{P})'$ bezüglich der schwachen Topologie erklärt ist. Wir zeigen: $(X, \|\ \|)' = (X, \mathcal{P})'$. Jedes $f \in (X, \mathcal{P})$ ist erst recht norm-stetig, also $(X,\mathcal{P})' \subset (X,\|\ \|)'$. Sei umgekehrt $x' \in (X,\|\ \|)'$, dann ist $|x'(x)| = p_{x'}(x)$ durch eine Halbnorm beschränkt und damit $x' : (X,\mathcal{P}) \to \mathbb{K}$ stetig nach Satz 9.6. Also $(X,\|\ \|)' \subset (X,\mathcal{P})'$. ◁

Beispiel 9.9 Der Dualraum $(\mathcal{S}(\mathbb{R}^n))'$ des Schwartz-Raumes, siehe Beispiel 8.9.vi), heißt Vektorraum der *temperierten Distributionen*. Diese spielen eine große Rolle in der Theorie partieller Differentialgleichungen. Es sind sozusagen die langsam wachsenden Distributionen. Ein wichtiges Beispiel ist die Diracsche δ-Distribution $\delta_t(f) := f(t)$ für $f \in \mathcal{S}(\mathbb{R}^n)$. Diese ist linear, und $|\delta_t(f)| \leq p_{0,0}(f) = \sup_{t\in\mathbb{R}^n} |f(t)|$, d.h. δ_t ist stetig. ◁

Satz 9.10 (Hahn-Banach VI: Fortsetzungssatz für lokalkonvexe VR)
Sei X lokalkonvexer Vektorraum und $U \subset X$ ein Untervektorraum. Dann gilt: Zu jedem linearen stetigen Funktional $f : U \to \mathbb{K}$ gibt es eine Fortsetzung zu einem linearen stetigen Funktional $F : X \to \mathbb{K}$ mit $F\big|_U = f$.

Beweis. Der Untervektorraum U trägt die Relativtopologie gegeben durch Einschränkung der Halbnormen $p \in \mathcal{P}$ auf U. Sei zunächst $\mathbb{K} = \mathbb{R}$. Da f stetig ist, gibt es nach Satz 9.6 ein $c \geq 0$ und Halbnormen $p_1,\dots,p_n \in \mathcal{P}$ mit $f(u) \leq |f(u)| \leq c\max(p_1(u),\dots,p_n(u))$ für alle $u \in U$. Dann definiert $p(x) := c\max(p_1(x),\dots,p_n(x))$ eine sublineare Abbildung $p : X \to \mathbb{R}$, so daß es nach dem Satz 3.5 von Hahn-Banach für reelle Vektorräume eine lineare Fortsetzung $F : X \to \mathbb{R}$ gibt mit $F\big|_U = f$ und $F(x) \leq p(x)$ für alle $x \in X$. Wegen $p(x) = p(-x)$ gilt $-F(x) = F(-x) \leq p(x)$, also $|F(x)| \leq c\max(p_1(x),\dots,p_n(x))$ für alle $x \in X$. Damit ist F stetig. Die übliche Diskussion beweist dann den Fall $\mathbb{K} = \mathbb{C}$. □

Der Trennungssatz (Theorem 3.17) von Hahn-Banach überträgt sich wörtlich auf lokalkonvexe Vektorräume, da nur die Eigenschaften des Minkowski-Funktionals p_U eingehen. Die Stetigkeit der so konstruierten p_U ergibt sich wie folgt: Da U offene Umgebung von 0 ist, gibt es $U(0;p_1,\dots,p_n,\epsilon) \subset U$. Sei $x \in X$ beliebig. Ist $\max(p_1(x),\dots,p_n(x)) > 0$, so ist $\frac{x}{\frac{2}{\epsilon}\max(p_1(x),\dots,p_n(x))} \in U(0;p_1,\dots,p_n,\epsilon) \subset U$ und damit $p_U(x) \leq \frac{2}{\epsilon}\max(p_1(x),\dots,p_n(x))$ stetig. Für $\max(p_1(x),\dots,p_n(x)) = 0$ ist $\lambda x \in U(0;p_1,\dots,p_n,\epsilon) \subset U$ für alle $\lambda > 0$ und dann $0 \leq p_U(\lambda x) = \lambda p_U(x) < 1$, somit auch $p_U(x) = 0$. Damit ist P_U stetig. Wir können dann zeigen:

Theorem 9.11 (Hahn-Banach VII: Trennungssatz für lokalkonvexe VR)
Sei $(X,\mathcal{P})$ lokalkonvexer Vektorraum, $V \subset X$ konvexe abgeschlossene Teilmenge und $x \in X \setminus V$. Dann gibt es ein $x' \in (X,\mathcal{P})'$ und ein $\delta > 0$ mit $\operatorname{Re} x'(x) + \delta \leq \operatorname{Re} x'(v)$ für alle $v \in V$.

Beweis. Da $X \setminus V$ offene Umgebung von x, gibt es $\epsilon > 0$ und $p_1,\dots,p_n \in \mathcal{P}$ mit $U(x;p_1,\dots,p_n,\epsilon) \cap V = \varnothing$. In den Bezeichnungen von Theorem 3.17 ist $V_1 := U(x;p_1,\dots,p_n,\epsilon)$ und $V_2 := V$. Sei $v_0 \in V$ beliebig, dann ist $U := V_1 - V_2 -$

$(x - v_0) = U(0; p_1, \dots, p_n, \epsilon) - (V - v_0)$ eine offene konvexe Umgebung von 0 mit $U(0; p_1, \dots, p_n, \epsilon) \subset U$. Das (nach der lokalkonvexen Version von Theorem 3.17 existierende) lineare stetige Funktional $\operatorname{Re} x'$ mit

$$\operatorname{Re} x'(x) + \operatorname{Re} x'(u) < \operatorname{Re} x'(v) \qquad \text{für alle } u \in U(0; p_1, \dots, p_n, \epsilon) \text{ und } v \in V$$

ist nach Konstruktion eine Fortsetzung des Funktionals y' auf $\mathbb{R}(x - v_0)$ gegeben durch $y'(\lambda(v_0 - x)) = \lambda p_U(v_0 - x)$. Wegen $p_U(v_0 - x) \geq 1$ ist $x'(\lambda(v_0 - x)) \neq 0$ für alle $\lambda \neq 0$. Dann ist auch $p_i(v_0 - x) \neq 0$ für mindestens ein $i \in \{1, \dots, n\}$, denn ansonsten wäre $v_0 - x \in U(0; p_1, \dots, p_n, \epsilon)$ im Widerspruch zu $V \cap U(x; p_1, \dots, p_n, \epsilon) = \varnothing$. Somit ist $u := \frac{(v_0 - x)}{\frac{2}{\epsilon} \max(p_1(v_0 - x), \dots, p_n(v_0 - x))} \in U(0; p_1, \dots, p_n, \epsilon)$ mit $x'(u) = \frac{\epsilon}{2} \frac{p_U(v_0 - x)}{\max(p_1(v_0 - x), \dots, p_n(v_0 - x))} \neq 0$. Da $U(0; p_1, \dots, p_n, \epsilon)$ kreisförmig ist, ist $\delta := \sup_{u \in U(0; p_1, \dots, p_n, \epsilon)} \operatorname{Re} x'(u) > 0$, woraus die Behauptung folgt. □

Folgerung 9.12 *Der Dualraum $(X, \mathcal{P})'$ eines lokalkonvexen Vektorraums $(X, \mathcal{P})$ trennt die Punkte, d.h. zu $x, y \in X$ mit $x \neq y$ gibt es ein $x' \in (X, \mathcal{P})'$ mit $x'(x) \neq x'(y)$.*

Beweis. Ein lokalkonvexer Vektorraum ist nach Definition Hausdorffsch, somit ist $V := \{y\}$ nach Satz 8.8 abgeschlossen und trivialerweise konvex. Die Behauptung folgt dann aus Theorem 9.11. □

10 Der Satz von Stone-Weierstraß

Der Approximationssatz von Stone-Weierstraß ist unverzichtbares Hilfsmittel bei vielen Betrachtungen über stetige Funktionen. Für den Beweis benötigen wir ein Theorem von eigenem Interesse: den Satz von Krein-Milman über Extrempunkte in kompakten konvexen Mengen. Wir geben dann einen operatoralgebraischen Beweis des Satzes von Stone-Weierstraß. Es gibt einen alternativen maßtheoretischen Beweis über den Rieszschen Darstellungssatz, siehe [Werner]. Dieser besagt, daß der Dualraum von $C(K)$, K konpakter Hausdorff-Raum, gegeben ist durch den Raum der signierten bzw. komplexen regulären Borel-Maße.

Definition 10.1 Sei X ein Vektorrraum und $K \subset X$ konvex. Eine Teilmenge $F \subset K$ heißt *Seite*, falls F konvex ist und aus $\frac{x_1 + x_2}{2} \in F$ für $x_1, x_2 \in K$ folgt $x_1, x_2 \in F$. Eine Seite von K, die nur aus einem Punkt besteht, heißt *Extrempunkt* von K. Mit $\mathrm{ex}(K)$ werde die Menge der Extrempunkte von K bezeichnet.

Beispiel 10.2
i) In normierten Vektorräumen gilt $\mathrm{ex}(B_X) \subset S_X$, denn für $x \notin \{S_X, 0\}$ ist $x = \frac{1}{2} \frac{x}{\|x\|} + \frac{1}{2}(2x - \frac{x}{\|x\|})$ Konvexkombination von Punkten $B_X \ni \frac{x}{\|x\|}, (2x - \frac{x}{\|x\|}) \neq x$, ebenso $0 = \frac{1}{2}x + \frac{1}{2}(-x)$ für beliebige $0 \neq x \in B_X$.

ii) Ist $(X, \| \, \|)$ gleichmäßig konvex, so gilt $\mathrm{ex}(B_X) = S_X$, denn aus $x, y \in B_X$ mit $\|\frac{1}{2}(x + y)\| = 1$ folgt $x = y \in S_X$.

iii) Für $X = (c_0, \| \ \|_\infty)$ (Vektorraum der Nullfolgen) besitzt B_X keine Extrempunkte. Denn für einen solchen Extrempunkt $x = (x_n) \in c_0$ gäbe es ein $k \in \mathbb{N}$ mit $|x_k| < \frac{1}{2}$, so daß $x \pm \frac{1}{2}e_k \in B_{c_0}$ für $(e_k)_n := \delta_{nk}$. Dann folgt $x = \frac{1}{2}(x + \frac{1}{2}e_k) + \frac{1}{2}(x - \frac{1}{2}e_k)$.

Die Beispiele zeigen, daß Extrempunkte selten sein können. Kompakte konvexe Teilmengen K besitzen jedoch immer Extrempunkte, und ihre Konvexkombinationen spannen sogar K auf. Erinnert sei an die Bezeichnung $\mathrm{conv}(A) := \Big\{ \sum_{i=1}^l \lambda_i x_i \;:\; l \in \mathbb{N}\,,\; x_1, \dots, x_l \in A\,,\; \lambda_1, \dots, \lambda_l \in [0,1]\,,\; \sum_{i=1}^l \lambda_i = 1\}$.

Theorem 10.3 (Krein-Milman) *Sei $(X, \mathcal{P})$ lokalkonvexer Vektorraum, $K \subset X$ nichtleer, konvex und kompakt (in der lokalkonvexen Topologie). Dann gilt:*

i) $\mathrm{ex}(K) \neq \varnothing$.

ii) $\overline{\mathrm{conv}(\mathrm{ex}(K))} = K$ *(Abschluß in der lokalkonvexen Topologie).*

Beweis. i) Sei $\mathcal{F}$ die Menge aller abgeschlossenen Seiten $F \subset K$. Insbesondere ist K selbst eine abgeschlossene Seite (kompakte Teilmengen in Hausdorff-Räumen sind immer abgeschlossen), somit $\mathcal{F} \neq \varnothing$. Eine Seite $F_1 \subset F$ einer Seite $F \subset K$ ist auch Seite von K. Somit definiert $\subset$ eine Ordnung (siehe Definition 3.1) auf $\mathcal{F}$: Wir sagen $F_1 \leq F_2$ falls $F_2 \subset F_1$. Sei $\{F_i \;:\; i \in I\}$ eine beliebige Kette in $\mathcal{F}$. Dann ist der Durchschnitt endlich vieler F_i nichtleer. Nach der endlichen Durchschnittseigenschaft von K ist dann auch $\bigcap_{i \in I} F_i \neq \varnothing$, somit gibt es eine obere Schranke $f \in \bigcap_{i \in I} F_i$ mit $f_i \leq f$ für alle $i \in I$. Nach dem Lemma 3.2 von Zorn gibt es ein maximales Element F_m in $\mathcal{F}$.

Wir zeigen, daß F_m Extrempunkt ist. Angenommen, es gibt $x, y \in F_m$ mit $x \neq y$. Nach Folgerung 9.12 aus dem Trennungssatz von Hahn-Banach gibt es ein $x' \in (X, \mathcal{P})'$ mit $\mathrm{Re}\, x'(x) < \mathrm{Re}\, x'(y)$. Sei $s := \sup_{y \in F_m}(\mathrm{Re}\, x'(y))$. Da F_m kompakt ist als abgeschlossene Teilmenge von K, nimmt die stetige reellwertige Funktion $\mathrm{Re}\, x'$ das Supremum an, d.h. $F_s := \{y \in F_m \;:\; \mathrm{Re}\, x'(y) = s\} \neq \varnothing$. Als Urbild von $\{s\}$ unter $\mathrm{Re}\, x'$ ist F_s abgeschlossen, und aus $\mathrm{Re}\, x'(\frac{y_1+y_2}{2}) = s$ für $y_1, y_2 \in F_m$ folgt $y_1, y_2 \in F_s$. Damit ist $F_s \subset F_m$ abgeschlossene Seite, d.h. $F_m \leq F_s$. Aber es gibt $x \in F_m \setminus F_s$, d.h. $F_s \neq F_m$ im Widerspruch zur Maximalität von F_m. Somit ist $F_m = \{y_m\} \in \mathrm{ex}(K)$ Extrempunkt.

ii) Sei $K_1 := \overline{\mathrm{conv}(\mathrm{ex}(K))}$, also abgeschlossene konvexe Teilmenge von K, somit kompakt und nach i) nichtleer. Angenommen, $K_1 \neq K$. Dann gibt es ein $x \in K \setminus K_1$ und nach dem Trennungssatz (Theorem 9.11) von Hahn-Banach ein $(-x') \in (X, \mathcal{P})'$ und $\delta > 0$ mit $\mathrm{Re}\,(-x'(x)) + \delta \leq \mathrm{Re}\,(-x'(y))$ für alle $y \in K_1$, also $\mathrm{Re}\, x'(y) \leq \mathrm{Re}\, x'(x) - \delta$. Sei $s := \sup_{y \in K} \mathrm{Re}\, x'(y)$, dann ist $F_s := \{y \in K \;:\; \mathrm{Re}\, x'(y) = s\}$ wieder abgeschlossene, nichtleere Seite von K, insbesondere konvex und kompakt, so daß F nach i) einen Extrempunkt y_s besitzt. Dann ist y_s auch Extrempunkt von K, somit $y_s \in K_1$. Einerseits war $\mathrm{Re}\, x'(y_s) = s$ wegen $y_s \in F_s$, andererseits

$$s = \mathrm{Re}\, x'(y_s) \leq \mathrm{Re}\, x'(x) - \delta \leq s - \delta\,,$$

Widerspruch. Somit $K_1 = K$. $\square$

Sei nun K ein kompakter Hausdoff-Raum. Wir erarbeiten uns grundlegende Eigenschaften des Banach-Raums $(\mathcal{C}(K), \| \ \|)$ unter Verwendung von "besonderen" Elementen seines Dualraums $(\mathcal{C}(K), \| \ \|)'$.

Eine *Algebra* $\mathcal{A}$ über $\mathbb{K}$ ist gleichzeitig ein Ring und ein $\mathbb{K}$-Vektorraum, so daß beide Strukturen verträglich sind. Insbesondere ist $\mathcal{C}(K)$ eine Algebra.

Definition 10.4 Sei $\mathcal{A}$ eine Algebra. Eine Untervektorraum $\mathcal{J} \subset \mathcal{A}$ heißt *Linksideal* bzw. *Rechtsideal*, falls $ab \subset \mathcal{J}$ bzw. $ba \subset \mathcal{J}$ für alle $a \in \mathcal{A}$ und $b \in \mathcal{J}$. Ist $\mathcal{A}$ kommutativ (wie z.B. $\mathcal{C}(K)$), so fallen Links- und Rechtsideale zusammen, wir sprechen dann kurz von *Idealen*.

Ein Ideal $\mathcal{J}$ heißt *echtes Ideal*, falls $\mathcal{J} \neq \mathcal{A}$ und $\mathcal{J} \neq \{0\}$. Ein Ideal $\mathcal{M}$ von $\mathcal{A}$ heißt *maximales Ideal*, falls für jedes echte Ideal $\mathcal{J}$ von $\mathcal{A}$ mit $\mathcal{M} \subset \mathcal{J}$ gilt $\mathcal{M} = \mathcal{J}$.

Definition 10.5 Eine Funktion $f \in \mathcal{C}(K)$ heißt *positiv*, geschrieben $f \geq 0$, falls $f(t) \geq 0$ für alle $t \in K$. Mit $\mathcal{C}_+(K)$ werde die Menge der positiven stetigen Funktionen bezeichnet. Dann wird auf $\mathcal{C}(K)$ eine Ordnung $\leq$ erklärt durch $f \leq g$ falls $g - f \in \mathcal{C}_+(K)$.

Ein lineares stetiges Funktional $\rho \in (\mathcal{C}(K))'$ heißt

i) *hermitesch*, falls $\rho(\bar{f}) = \overline{\rho(f)}$ für alle $f \in \mathcal{C}(K)$.

ii) *positiv*, falls $\rho(f) \geq 0$ für alle $f \geq 0$; wir schreiben dann $\rho \geq 0$.

iii) *Zustand*, falls $\rho \geq 0$ und $\rho(1) = 1$. Dabei ist 1 die konstante Funktion $1(t) = 1$.

iv) *multiplikativ*, falls $\rho(fg) = \rho(f)\rho(g)$ für alle $f, g \in \mathcal{C}(K)$.

Ist ρ hermitesch, f reell, so ist auch $\rho(f)$ reell.

Lemma 10.6 i) *Jedes positive lineare stetige Funktional ist hermitesch.*

ii) *Jedes positive lineare stetige Funktional ρ auf $\mathcal{C}(K)$ erfüllt die Ungleichung von Cauchy-Schwarz*

$$\big|\rho(\bar{f}g)\big|^2 \leq \rho(|f|^2)\rho(|g|^2) \qquad \text{für alle } f, g \in \mathcal{C}(X) \ .$$

iii) *Für jeden Zustand ω gilt $\|\omega\| = 1$.*

iv) *Für jedes positive lineare stetige Funktional ρ gilt $\|\rho\| = \rho(1)$.*

Beweis. i) Sei ρ positiv, dann $\rho(|f + \lambda 1|^2) = \rho(|f|^2) + |\lambda|^2\rho(1) + \lambda\rho(\bar{f}) + \bar{\lambda}\rho(f)$ reell für alle $\lambda \in \mathbb{K}$. Somit $\lambda\rho(\bar{f}) + \bar{\lambda}\rho(f) = \lambda\big(\rho(\bar{f}) - \overline{\rho(f)}\big) + 2\mathrm{Re}\big(\bar{\lambda}\rho(f)\big)$ reell für alle λ und deshalb $\rho(\bar{f}) - \overline{\rho(f)}$.

ii) Für beliebige $\lambda \in \mathbb{K}$ gilt unter Verwendung von i)

$$0 \leq \rho(|f + \lambda g|^2) = \rho(|f|^2) + |\lambda|^2\rho(|g|^2) + 2\mathrm{Re}\big(\lambda\rho(\bar{f}g)\big) \ .$$

Ist $\rho(|g|^2) = 0$, so muß auch $\rho(\bar{f}g) = 0$ gelten, da ansonsten $\lambda := -\frac{\rho(|f|^2)}{\rho(\bar{f}g)}$ zum Widerspruch führt. Ist $\rho(|g|^2) \neq 0$, so setze $\lambda := -\frac{\overline{\rho(\bar{f}g)}}{\rho(|g|^2)}$, dann folgt

$$0 \leq \rho(|f|^2) + \frac{\left|\rho(\bar{f}g)\right|^2}{\rho(|g|^2)} - 2\frac{\left|\rho(\bar{f}g)\right|^2}{\rho(|g|^2)} .$$

iii) Sei $f \in \mathcal{C}(K)$ reellwertig, dann gilt $-\|f\|_\infty 1 \leq f \leq \|f\|_\infty 1$. Anwendung von ω erhält die Ordnung, also $-\|f\|_\infty \omega(1) \leq \omega(f) \leq \|f\|_\infty \omega(1)$, also $|\omega(f)| \leq \|f\|_\infty$, und dann $\|\omega\| \leq 1$. Umgekehrt ist $\omega(1) = 1$, deshalb $\|\omega\| = 1$. Im komplexen Fall gilt $|\omega(f1)|^2 \leq \omega(|1|^2)\omega(|f|^2) \leq \||f|^2\|_\infty \leq \|f\|_\infty^2$ nach ii).

iv) Ist ρ positiv mit $\rho(1) = 0$, dann ist $\rho = 0$ auf allen positiven Funktionen f wegen $0 \leq f \leq \|f\|_\infty 1$. Jede stetige Funktion ist Linearkombination von 4 positiven stetigen Funktionen, somit $\rho = 0$. Ansonsten ist $\omega := \frac{\rho}{\rho(1)}$ Zustand und damit $\|\rho\| = \rho(1)$. □

Lemma 10.7 *Jedes hermitesche lineare stetige Funktional ρ besitzt eine eindeutige Zerlegung $\rho = \rho_+ - \rho_-$ in positive lineare stetige Funktionale $\rho_+, \rho_- \geq 0$ mit $\|\rho\| = \|\rho_+\| + \|\rho_-\|$.*

Beweis. i) Wir nehmen die Existenz solcher Funktionale an und beweisen die Eindeutigkeit.

Nach Definition des Supremums gibt es zu jedem $n \in \mathbb{N}\setminus\{0\}$ eine reellwertige Funktion $g \in B_{\mathcal{C}(K)}$ mit

$$\|\rho\| - \frac{1}{n} \leq \rho(g_n) = \rho_+(g_n) - \rho_-(g_n) \leq \|\rho\| ,$$

Wegen $\|\rho\| = \rho_+(1) + \rho_-(1) = \|\rho_+\| + \|\rho_-\|$ folgt $\rho_+(1 - g_n) \leq \frac{1}{n} - \rho_-(1 + g_n)$. Wegen $1 \pm g \geq 0$ folgt $\lim_{n\to\infty} \rho_+(g_n) = \rho_+(1) = \|\rho_+\|$ und $\lim_{n\to\infty} \rho_-(g_n) = -\rho_-(1) = -\|\rho_-\|$.

Wir zerlegen $g_n = g_{n+} - g_{n-}$ mit $g_{n+} = \max(0, g_+) \geq 0$ und $g_{n-} = \max(-g_n, 0) \geq 0$. Insbesondere sind g_{n+}, g_{n-} stetig. Dann ist

$$\rho_+(1 - g_{n+} + g_{n-}) = \rho_+(1 - g_{n+}) + \rho_+(g_{n-}) \stackrel{n\to\infty}{\longrightarrow} 0 ,$$

woraus wegen der Positivität $\rho_+(g_{n+}) \stackrel{n\to\infty}{\longrightarrow} \|\rho_+\|$ und $\rho_+(g_-) \stackrel{n\to\infty}{\longrightarrow} 0$ folgt. Analog ergibt sich $\rho_-(g_{n+}) \stackrel{n\to\infty}{\longrightarrow} 0$. Ist $f \in \mathcal{C}_+(X)$, so folgt mit $1 - g_{n+} \geq 0$ und Cauchy-Schwarz

$$\begin{aligned} 0 \leq |\rho_+((1 - g_{n+})f)|^2 &= \left|\rho_+(\sqrt{1 - g_{n+}} \cdot (\sqrt{1 - g_{n+}} f))\right|^2 \\ &\leq \rho_+(|\sqrt{1 - g_{n+}}|^2) \cdot \rho_+(|\sqrt{1 - g_{n+}} f|^2) \stackrel{n\to\infty}{\longrightarrow} 0 , \end{aligned}$$

also $\rho_+(g_+ f) \stackrel{n\to\infty}{\longrightarrow} \rho_+(f)$. Analog folgt $\rho_-(g_{n+} f) \stackrel{n\to\infty}{\longrightarrow} 0$ und deshalb

$$\rho_+(f) = \lim_{n\to\infty} \rho(g_{n+} f) \leq \sup\{\rho(f_1) \ : \ 0 \leq f_1 \leq f \ , \ f_1 \in \mathcal{C}(K)\} . \tag{*}$$

Ist umgekehrt $f_1 \in \mathcal{C}_+(X)$ mit $0 \le f_1 \le f$, so gilt $\rho(f_1) \le \rho_+(f_1) \le \rho_+(f)$ und damit $\sup\{\rho(f_1) \,:\, 0 \le f_1 \le f\} \le \rho_+(f)$, so daß mit $(*)$ folgt:

$$\begin{aligned}\rho_+(f) &= \sup\{\rho(f_1) \,:\, 0 \le f_1 \le f \,,\ f_1 \in \mathcal{C}(K)\} \,, \\ \rho_-(f) &= \sup\{-\rho(f_2) \,:\, 0 \le f_2 \le f \,,\ f_2 \in \mathcal{C}(K)\} \,.\end{aligned} \tag{**}$$

Damit sind ρ_+, ρ_- auf $\mathcal{C}_+(K)$ eindeutig bestimmt. Sie werden für die (eindeutige) Zerlegung einer Funktion $f = (\operatorname{Re} f)_+ - (\operatorname{Re} f)_- + \mathrm{i}(\operatorname{Im} f)_+ - \mathrm{i}(\operatorname{Im} f)_- \in \mathcal{C}(K)$ eindeutig fortgesetzt zu $\rho_+(f) := \rho_+((\operatorname{Re} f)_+) - \rho_+((\operatorname{Re} f)_-) + \mathrm{i}\rho_+((\operatorname{Im} f)_+) - \mathrm{i}\rho_+((\operatorname{Im} f)_-)$, analog für ρ_-.

ii) Existenz: Zu zeigen ist, daß die nach (**) *definierten* ρ_+, ρ_- wirklich positive lineare stetige Funktionale sind. Zunächst zur Linearität. Hier genügt es, die Additivität auf positiven Funktionen $0 \le f, g$ zu zeigen. Seien $0 \le f_1 \le f$ und $0 \le g_1 \le g$ Approximationen mit $|\rho_+(f) - \rho(f_1)| < \epsilon$ und $|\rho_+(g) - \rho(g_1)| < \epsilon$, dann ist

$$\rho(f_1) + \rho(g_1) = \rho(f_1 + g_1) \le \Big(\sup_{0 \le h_1 \le f+g} \rho(h_1) \Big) = \rho_+(f+g) \,,$$

und für $\epsilon \to 0$ entsteht

$$\rho_+(f) + \rho(g) \le \rho_+(f+g) \,. \tag{†}$$

Sei nun eine stetige positive Funktion $0 \le h \le f + g$ gegeben mit $|\rho(h) - \rho_+(f+g)| < \epsilon$. Dann läßt sich h zerlegen[1] in $h = h_1 + h_2$ mit $0 \le h_1 \le f$ und $0 \le h_2 \le g$. Dann ist

$$\rho(h) = \rho(h_1) + \rho(h_2) \le \Big(\sup_{0 \le f_1 \le f} \rho(f_1) + \sup_{0 \le g_1 \le g} \rho(g_1) \Big) = \rho_+(f) + \rho_+(g) \,.$$

Für $\epsilon \to 0$ folgt

$$\rho_+(f+g) \le \rho_+(f) + \rho_+(g) \,. \tag{‡}$$

Damit ist ρ_+, und analog ρ_-, linear.

Die Relation $\rho = \rho_+ - \rho_-$ ergibt sich aus der Definition (**):

$$\begin{aligned}-\rho_-(f) &= \inf\{\rho(f_2) \,:\, 0 \le f_2 \le f \,,\ f_2 \in \mathcal{C}(K)\} \\ &= \inf\{\rho(f - f_1) \,:\, 0 \le f_1 \le f \,,\ f_1 \in \mathcal{C}(K)\} = \rho(f) - \rho_+(f) \,.\end{aligned}$$

Die Relation $\|\rho\| \le \|\rho_+\| + \|\rho_-\|$ ist klar. Umgekehrt ist

$$\begin{aligned}\|\rho_+\| + \|\rho_-\| &= \rho_+(1) + \rho_-(1) \\ &= \sup_{0 \le f_1 \le 1} \rho(f_1) - \inf_{0 \le f_2 \le 1} \rho(f_2) \,,\end{aligned}$$

[1] Setze $h_1 = \min(h, f) \ge 0$ und $h_2 = h - h_1 \ge 0$. Dann sind h_1, h_2 stetig. Angenommen, es gilt *nicht*: $h_2 \le g$. Dann gibt es ein $t \in K$ mit $0 \le g(p) < h_2(p) = h(p) - \min(h(p), f(p))$, also $\min(h(p), f(p)) < h(p)$ und damit $\min(h(p), f(p)) = f(p)$. Es folgt $f(p) + g(p) < h(p)$, Widerspruch. Also $h_2 \le g$.

d.h. für beliebige $\epsilon > 0$ gibt es $0 \leq f_1, f_2 \leq 1$ mit $\rho_+(1) + \rho_-(1) \leq \rho(f_1 - f_2) + \epsilon$. Aber $\rho(f_1 - f_2) \leq \|\rho\| \|f_1 - f_2\|_\infty \leq \|\rho\|$ und damit $\rho_+(1) + \rho_-(1) \leq \|\rho\|$. Insbesondere sind ρ_+, ρ_- beschränkt. $\square$

Satz 10.8 *Sei K kompakter Hausdorff-Raum (der nicht nur aus einem Punkt besteht) und $\rho \in (\mathcal{C}(K))'$ multiplikativ und nicht identisch 0. Dann gibt es einen Punkt $t \in K$ mit $\rho(f) = f(t)$ für alle $f \in \mathcal{C}(K)$.*

Beweis. Es sei $\mathcal{M} := \{f \in \mathcal{C}(K) \ : \ \rho(f) = 0\}$. Als Urbild von $\{0\}$ unter ρ ist $\mathcal{M}$ abgeschlossen. Für $f \in \mathcal{M}$ ist $\rho(|f|^2) = \rho(f)\rho(\bar{f}) = 0$ und dann $\rho(fg) = 0$ für alle $g \in \mathcal{C}(K)$ nach Cauchy-Schwarz. Somit ist $\mathcal{M}$ ein Ideal von $\mathcal{C}(K)$. Wegen $\rho \neq 0$ ist es von $\mathcal{C}(X)$ verschieden. Damit ist $1 \notin \mathcal{M}$, so daß $0 \neq \rho(1) = \rho(1^2) = \rho(1)\rho(1)$ und damit $\rho(1) = 1$. Außerdem ist $\mathcal{M} \neq \{0\}$, denn ansonsten gäbe es eine nichtkonstante stetige Funktion $f \in \mathcal{C}(K)$ mit $\rho(f) \neq 0$, und damit folgt $0 \neq 1 - \frac{f}{\rho(f)} \in \mathcal{M}$. Somit ist $\mathcal{M}$ echtes Ideal von $\mathcal{C}(K)$.

Wir zeigen, daß $\mathcal{M}$ maximales Ideal ist. Angenommen, $\mathcal{M} \subset \mathcal{J}$ für ein echtes Ideal $\mathcal{J}$, und es gibt ein $g \in \mathcal{J} \setminus \mathcal{M}$. Dann ist $\rho(g) \neq 0$, und für $h := \frac{1}{\rho(g)} \in \mathcal{C}(K)$ gilt $\rho(g)\rho(h) = 1 = \rho(\frac{g}{\rho(g)})$. Aber $\frac{g}{\rho(g)} \in \mathcal{J}$ und $1 - \frac{g}{\rho(g)} \in \mathcal{M} \subset \mathcal{J}$, also auch $1 = 1 - \frac{g}{\rho(g)} + \frac{g}{\rho(g)} \in \mathcal{J}$ und dann $\mathcal{J} = \mathcal{C}(K)$, d.h. $\mathcal{J}$ wäre kein echtes Ideal.

Die Menge aller Funktionen, die in einem Punkt $t \in K$ verschwinden, bilden ein maximales Ideal $\mathcal{M}_t$: Denn setze $\rho_t(g) := g(t)$, dann ist ρ_t multiplikativ, nicht identisch 0 wegen $\rho_t(1) \neq 0$ und $\mathcal{M}_t := \{f \in \mathcal{C}(K) \ : \ \rho_t(f) = 0\}$ nach obiger Diskussion maximales Ideal.

Wir zeigen, daß jedes maximale Ideal $\mathcal{M}$ ein solches $\mathcal{M}_t$ ist. Angenommen, $\mathcal{M}$ ist verschieden von allen $\mathcal{M}_t$, insbesondere in keinem $\mathcal{M}_t$ enthalten. Dann gibt es zu jedem $t \in K$ eine stetige Funktion $f_t \in \mathcal{M}$ mit $f_t(t) \neq 0$. Wegen der Stetigkeit von f_t gibt zu jedem t sogar eine offene Umgebung $U_t \subset K$ mit $t \in U_t$ und $f_t(y_t) \neq 0$ für alle $y_t \in U_t$. Da K kompakt ist (die Kompaktheit geht genau hier entscheidend ein!), wird K durch endlich viele dieser U_t überdeckt, d.h. es gibt Punkte $t_1, \ldots, t_m$ mit

$$f := \sum_{i=1}^{m} f_{t_i} \cdot \overline{f_{t_i}} \in \mathcal{M} \ .$$

Nach Konstruktion ist $f(t) \neq 0$ für alle $t \in K$. Somit ist die Funktion $\frac{1}{f}$ stetig, und es folgt $1 = f \cdot \frac{1}{f} \in \mathcal{M}$, Widerspruch. $\square$

Theorem 10.9 (Stone-Weierstraß) *Es sei K ein kompakter Hausdorff-Raum und $\mathcal{A}$ eine Unteralgebra von $\mathcal{C}(K)$ mit*

i) *$1 \in \mathcal{A}$.*

ii) *$\mathcal{A}$ trennt die Punkte von K, d.h. für $t, s \in K$ mit $t \neq s$ gibt es ein $g \in \mathcal{A}$ mit $g(t) \neq g(s)$.*

iii) *Mit $g \in \mathcal{A}$ ist auch $\bar{g} \in \mathcal{A}$.*

Dann ist $\mathcal{A}$ dicht in $\mathcal{C}(K)$ bezüglich der Supremums-Norm $\| \ \|$.

Bemerkung: In ii) geht ein, daß K Hausdorff-Raum sein muß.

Beweis. Angenommen, $\bar{\mathcal{A}} \neq \mathcal{C}(K)$ (Abschluß in $\| \ \|_\infty$). Nach der Folgerung 3.11 aus Hahn-Banach gibt es ein $0 \neq \rho \in (\mathcal{C}(K))'$ mit $\rho(g) = 0$ für alle $g \in \bar{\mathcal{A}}$. Sei $\rho^*(f) := \overline{\rho(\bar{f})}$, mit $f \in \mathcal{C}(K)$, dann sind $\rho_+ := \frac{1}{2}(\rho + \rho^*)$ und $\rho_- := \frac{1}{2\mathrm{i}}(\rho - \rho^*)$ hermitesch und verschwinden nicht beide, so daß wir $\rho(\bar{f}) = \overline{\rho(f)}$ annehmen können (hier geht iii) ein!). Sei

$$B := \{\rho \in (\mathcal{C}(K))' \ : \ \|\rho\| \leq 2 \ , \ \rho\big|_{\bar{\mathcal{A}}} = 0 \ , \ \rho(\bar{f}) = \overline{\rho(f)}\} \ .$$

Dann ist B konvex (klar) und abgeschlossen in der schwach-*-Topologie auf $(\mathcal{C}(K))'$ als $B = \bigcap_{p_f} p_f^{-1}(\{0\})$, mit $p_g(\rho) = |\rho(g)|$ für $g \in \bar{\mathcal{A}}$. Außerdem ist B enthalten im Ball in $(\mathcal{C}(K))'$ vom Radius 2, welcher nach dem Satz von Banach-Alaoglu (Theorem 4.13) schwach-*-kompakt ist. Somit ist auch B schwach-*-kompakt, und nach dem Satz von Krein-Milman (Theorem 10.3) gibt es einen Extrempunkt $\eta \in B$. Insbesondere ist $\|\eta\| = 2$.

Nach Lemma 10.7 gibt es eine eindeutige Zerlegung $\eta = \eta_+ - \eta_-$ mit $\eta_+, \eta_- \geq 0$ und $2 = \|\eta\| = \|\eta_+\| + \|\eta_-\|$. Wegen $1 \in \mathcal{A}$ nach i) ist

$$0 = \eta(1) = \eta_+(1) - \eta_-(1) = \|\eta_+\| - \|\eta_-\| \ ,$$

(siehe Lemma 10.6), also $\eta_+(1) = 1 = \eta_-(1)$, d.h. η_+, η_- sind Zustände von $\mathcal{C}(\mathcal{K})$. Wir werden zeigen, daß es η_+, η_- multiplikativ sind. Dann gibt es nach Satz 10.8 Punkte $t_+, t_- \in K$ mit $\eta_\pm(f) = f(t_\pm)$ für alle $f \in \mathcal{C}(K)$. Es folgt $t_+ \neq t_-$, denn ansonsten wäre $(\eta_+ - \eta_-)(f) = \eta(f) = 0$ für alle $f \in \mathcal{C}(K)$ im Widerspruch zu $\|\eta\| = 2$. Dann ist aber $\eta(g) = (\eta_+ - \eta_-)(g) = g(t_+) - g(t_-) = 0$ für alle $g \in \mathcal{A}$, im Widerspruch zu ii). Somit folgt $\bar{A} = \mathcal{C}(K)$.

Wir haben also zu zeigen: $\eta_+(fh) = \eta_+(f)\eta_+(h)$ für alle $f, h \in \mathcal{C}(K)$, und analog für η_-. Wir zeigen zunächst, daß die Einschränkung $\eta_+\big|_{\bar{\mathcal{A}}}$ multiplikativ ist. Wegen der Linearität von $g_1 \mapsto \eta_+(gg_1)$ genügt der Nachweis für reelle Elemente von $\bar{\mathcal{A}}$. Durch Verschiebung und Skalierung $h := \frac{(1+\|g_1\|_\infty)1+g_1}{1+2\|g_1\|_\infty} \in \bar{\mathcal{A}}$ genügt es zu zeigen: $\eta_+(gh) = \eta_+(g)\eta_+(h)$ für alle $g, h \in \bar{\mathcal{A}}$ mit $0 \leq h \leq 1$. Wir diskutieren drei Fälle:

- Sei $0 = \eta_+(h) = \eta_+(|\sqrt{h}|^2)$. Dann folgt $\eta_+(gh) = \eta_+(g\sqrt{h} \cdot \sqrt{h}) = 0$ für alle $g \in \bar{\mathcal{A}}$ nach Cauchy-Schwarz, d.h. $\eta_+(gh) = \eta_+(g)\eta_+(h)$.
- Sei $1 = \eta_+(h)$, dann ist $\eta_+(|\sqrt{1-h}|^2) = 0$, somit $\eta_+(g(1-h)) = \eta_+(g\sqrt{1-h} \cdot \sqrt{1-h}) = 0$ für alle $g \in \bar{\mathcal{A}}$ nach Cauchy-Schwarz, d.h. $\eta_+(gh) = \eta_+(g)\eta_+(h)$.

- Verbleibt $0 < \eta_+(h) := \lambda < 1$. Dann ist auch $\eta_-(h) = \lambda$. Für $g \in \bar{\mathcal{A}}$ setze

$$\eta_{1+}(g) := \frac{\eta_+(gh)}{\lambda}\,, \qquad \eta_{1-}(g) := \frac{\eta_-(gh)}{\lambda}\,, \qquad \eta_1 := \eta_{1+} - \eta_{1-}\,,$$
$$\eta_{2+}(g) := \frac{\eta_+((1-h)g)}{1-\lambda}\,, \quad \eta_{2-}(g) := \frac{\eta_-((1-h)g)}{1-\lambda}\,, \quad \eta_2 := \eta_{2+} - \eta_{2-}\,.$$

  Nach Konstruktion sind $\eta_{1+}, \eta_{1-}, \eta_{2+}, \eta_{2-}$ positive lineare Funktionale, sogar Zustände. Da $\bar{\mathcal{A}}$ Unteralgebra ist, verschwinden η_1, η_2 identisch auf $\bar{\mathcal{A}}$. Weiter gilt $\|\eta_1\|, \|\eta_2\| \leq 2$, insgesamt also $\eta_1, \eta_2 \in B$. Andererseits gilt $\eta = \lambda\eta_1 + (1-\lambda)\eta_2$, da aber η Extrempunkt war, folgt $\eta = \eta_1 = \eta_2$. Dann ist aber $\eta = \eta_{1+} - \eta_{1-}$ und $2 = \|\eta\| = \|\eta_{1+}\| + \|\eta_{1-}\|$, und aus der Eindeutigkeit der Zerlegung folgt $\eta_+ = \eta_{1+}$, also $\eta_+(h) = \frac{\eta_+(fh)}{\eta_+(f)}$, d.h. η_+, und analog η_-, ist multiplikativ auf $\bar{\mathcal{A}}$.

Sei $\mathcal{J} := \{h \in \mathcal{C}(K) \ : \ \eta_+(|h|^2) = 0\} \subset \mathcal{C}(K)$. Für $h \in \mathcal{J}$ und $f \in \mathcal{C}(K)$ ist $\eta_+(|fh|^2) = \eta_+(h \cdot \bar{h}|f|^2) = 0$ nach Cauchy-Schwarz, somit $\mathcal{J}$ ein Ideal, wie in Satz 10.8 echtes Ideal von $\mathcal{C}(K)$. Analog ist $\mathcal{J} \cap \bar{\mathcal{A}}$ ein echtes Ideal von $\bar{\mathcal{A}}$. Da η_+ auf $\bar{\mathcal{A}}$ multiplikativ ist, folgt wie in Satz 10.8, daß daß $\mathcal{J} \cap \bar{\mathcal{A}}$ maximales Ideal von $\bar{\mathcal{A}}$ ist. Nach dem Lemma von Zorn ist $\mathcal{J}$ in einem maximalen Ideal von $\mathcal{C}(K)$ enthalten, und nach Satz 10.8 ist es ein $\mathcal{M}_t \supset \mathcal{J}$ für $t \in K$. Dann ist $\mathcal{M}_t \cap \bar{\mathcal{A}}$ ein Ideal von $\bar{\mathcal{A}}$, welches $\mathcal{J} \cap \bar{\mathcal{A}}$ enthält. Da letzteres maximal ist, gilt $\mathcal{M} \cap \bar{\mathcal{A}} = \mathcal{J} \cap \bar{\mathcal{A}}$ und somit $\eta_+(g) = g(t)$ für alle $g \in \mathcal{A}$. Da $\mathcal{A}$ die Punkte trennt, kann es nur ein $t = t_+ \in K$ mit dieser Eigenschaft geben.

Bezeichnet ρ_t das Auswertungsfunktional $\rho_t(f) = f(t_+)$, so haben wir gezeigt, daß $(\eta_+ - \rho_t)\big|_{\bar{\mathcal{A}}} = 0$ ist. Damit folgt aber noch nicht $\eta_+ = \rho_t$ auf ganz $\mathcal{C}(K)$. Dieses Ziel erreichen wir erst im nächsten Schritt.

Bezeichne $\mathcal{S}$ die Menge aller Zustände auf $\mathcal{C}(K)$ und

$$W := \{\omega \in \mathcal{S} \ : \ (\omega - \eta_+)\big|_{\bar{\mathcal{A}}} = 0\}$$

die Teilmenge jeder Zustände, die auf $\bar{\mathcal{A}}$ mit η_+ übereinstimmen. Wie zuvor B ist W konvex und schwach-*-abgeschlossen in $B_{C(K)'}$, damit schwach-*-kompakt. Nach dem Satz von Krein-Milman hat W einen Extrempunkt $\tau \in W \subset \mathcal{S}$. Wir zeigen nun, daß τ sogar Extrempunkt von $\mathcal{S}$ ist. Angenommen, es gibt Zustände $\omega_1, \omega_2 \in \mathcal{S}$ mit $\tau = \frac{1}{2}(\omega_1 + \omega_2)$. Dann ist $\frac{1}{2}(\omega_1 + \omega_2)(g) = \eta_+(g)$ für alle $g \in \bar{\mathcal{A}}$. Aus der Positivität folgt $0 \leq \frac{1}{2}\omega_1\big|_{\bar{\mathcal{A}}}, \frac{1}{2}\omega_2\big|_{\bar{\mathcal{A}}} \leq \eta_+\big|_{\bar{\mathcal{A}}}$, damit $\omega_1(|h|^2) = 0$ und $\omega_2(|h|^2) = 0$ für alle $h \in \mathcal{J} \cap \bar{\mathcal{A}}$, und dann $\omega_1(h) = \omega_2(h) = 0$ nach Cauchy-Schwarz. Das bedeutet aber $\omega_1, \omega_2 \in W$, denn $g - \eta_+(g)1 \in \mathcal{J} \cap \bar{\mathcal{A}}$ (da η_+ multiplikativ) für alle $g \in \bar{\mathcal{A}}$ und folglich $\omega_1(g) = \eta_+(g) = \omega_2(g)$. Da aber τ Extrempunkt von W war, folgt $\tau = \omega_1 = \omega_2 \in \mathcal{S}$. d.h. τ ist sogar Extrempunkt von $\mathcal{S}$. Nach dem folgenden Satz 10.11 ist dann τ multiplikativ, somit gegeben als Auswertung im selben Punkt $t_+ \in K$, so daß $\tau = \rho_t$ der *einzige* Extrempunkt von W ist. Nach dem Satz von Krein-Milman (Theorem 10.3.ii)) ist W die konvexe

Hülle der Extrempunkte, somit $W = \overline{\{\rho_t\}} = \{\rho_t\}$, und wegen $\eta_+ \in W$ folgt $\eta_+ = \rho_t$. $\square$

Definition 10.10 Ein Extrempunkt der Menge $\mathcal{S}$ der Zustände auf $\mathcal{C}(K)$ heißt *reiner Zustand.*

Satz 10.11 *Sei K kompakter Hausdorff-Raum. Ein positives lineares stetiges Funktional ω auf $\mathcal{C}(K)$ ist genau dann reiner Zustand, wenn es multiplikativ ist.*

Beweis. Sei ω reiner Zustand und $f, h \in \mathcal{C}(K)$ mit $0 \leq f \leq 1$. Definiere ein positives lineares stetiges Funktional ρ durch $\rho(h) := \omega(fh)$. Somit $0 \leq \rho(1) \leq 1$. Wir zeigen $\rho = \lambda\omega$ für ein $\lambda \in [0, 1]$:

i) Wäre $\rho(1) = 0$, dann $\rho = 0$ nach Lemma 10.6.iv).

ii) Wäre $\rho(1) = 1 = \omega(1)$, dann folgt $\rho = \omega$ nach Lemma 10.6.iv).

iii) Sei $0 < \lambda := \rho(1) < 1$, dann definiere $\rho_1 := \frac{\rho}{\lambda}$ und $\rho_2 := \frac{\omega-\rho}{1-\lambda}$. Dann sind ρ_1, ρ_2 Zustände mit $\rho = \lambda\rho_1 + (1-\lambda)\rho_2$. Da ω reiner Zustand ist, folgt $\omega = \rho_1$ und $\rho = \lambda\omega$.

Sei nun $\omega(h) = 0$, dann gilt für beliebige $f \in \mathcal{C}(K)$ mit $0 \leq f \leq 1$ die Relation $\omega(fh) = \rho(h) = \lambda\omega(h) = 0$. Da jede stetige Funktion Linearkoombination von 4 stetigen Funktionen $0 \leq f_i \leq 1$ ist, ist $\mathcal{M} = \ker\omega = \{h \in \mathcal{C}(K) \,:\, \omega(h) = 0\}$ ein Ideal von $\mathcal{C}(K)$. Wie in Satz 10.8 ergibt sich, daß es maximal ist, somit ist $\omega = \rho_t$ das Auswertungsfunktional (wieder nach Satz 10.8), und dieses ist multiplikativ.

Ist umgekehrt $\omega \neq 0$ multiplikativ, dann Auswertungsfunktional nach Satz 10.8 und damit automatisch Zustand. Angenommen, $\omega = \frac{1}{2}\rho_1 + \frac{1}{2}\rho_2$, für Zustände ρ_1, ρ_2. Dann ist $0 \leq \frac{1}{2}\rho_1, \frac{1}{2}\rho_2 \leq \omega$. ist $h \in \ker\omega$, dann (wie zuvor) auch $\rho_1(h) = 0 = \rho_2(h)$, somit $0 = \rho_1(f - \omega(f)1) = \rho_1(f) - \omega(f) = \rho_2(f) - \omega(f)$. Damit ist ω reiner Zustand. $\square$

Folgerung 10.12 *Ist K kompakter metrischer Raum, dann ist $\mathcal{C}(K)$ Norm-separabel.*

Beweis. Unter diesen Voraussetzungen gibt es eine abzählbare dichte Menge von Punkten $\{t_n\}$ in K. Setze $f_n(t) := d(t, t_k)$. Dann ist f_n stetig und trennt die Punkte. Sei $\mathcal{A}$ die Algebra erzeugt von $\{f_n\}$, d.h. die Menge der Polynome in f_n. dann erfüllt $\mathcal{A}$ die Voraussetzungen des Satzes von Stone-Weierstraß, ist also dicht in $\mathcal{C}(K)$. $\square$

11 Metrisierbarkeit und Vollständigkeit

Wir klären nun die Frage, wann eine lokalkonvexe Topologie durch eine Metrik induziert wird. Ist das der Fall, so ist naheliegend zu fragen, ob der lokalkonvexe Vektorraum vollständig ist.

Satz 11.1 *Sei $(X, \mathcal{T})$ lokalkonvexer Vektorraum. Dann sind äquivalent:*

i) *$(X, \mathcal{T})$ ist metrisierbar, d.h. es gibt eine Metrik d auf X, welche $\mathcal{T}$ erzeugt.*

ii) *$0 \in X$ besitzt eine abzählbare Umgebungsbasis.*

iii) *Es gibt ein abzählbares Halbnormensystem $\mathcal{P} = \{p_n \ : \ n \in \mathbb{N}\}$, welches $\mathcal{T}$ erzeugt.*

Beweis. i)⇒ii) ist klar: Wähle die $(\epsilon = \frac{1}{n})$-Umgebungen.

ii)⇒iii) Sei $\{V_n\}_{n\in\mathbb{N}}$ abzählbare Umgebungsbasis von 0. Jedes V_n enthält eine absolutkonvexe Umgebung $U_n \subset V_n$ von 0. Wähle für $\mathcal{P}$ die zugehörige Menge der Minkowski-Funktionale.

iii)⇒i) Setze $d(x,y) := \displaystyle\sum_{n=0}^{\infty} \frac{1}{2^n} \frac{p_n(x-y)}{1+p_n(x-y)}$. Die Reihe konvergiert, und d ist eine Metrik: Ist $d(x,y) = 0$, dann $p_n(x-y) = 0$ für alle n, und dann $x = y$ aus der Hausdorff-Eigenschaft. Sind $a, b, c \geq 0$ mit $c \leq a+b$, so folgt $\frac{c}{1+c} \leq \frac{a}{1+a} + \frac{b}{1+b}$. Dann liefert $a := p_n(x-y)$, $b = p_n(y-z)$, $c = p_n(x-z)$ die Dreiecksungleichung für d.

Wir zeigen, daß die Topologien von (X, d) und $(X, \mathcal{P})$ identisch sind. Sei $(x_\lambda)_{\lambda\in\Lambda}$ konvergent gegen x bezüglich d, d.h. $d(x_\lambda, x) \to 0$. Dann folgt $\frac{p_n(x_\lambda - x)}{1+p_n(x_\lambda - x)} \to 0$ für alle n, and damit $p_n(x_\lambda) \to p_n(x)$, d.h. $(x_\lambda)_{\lambda\in\Lambda}$ konvergiert gegen x bezüglich $\mathcal{P}$. Ist umgekehrt $(x_\lambda)_{\lambda\in\Lambda}$ konvergent gegen x bezüglich $\mathcal{P}$, dann gibt es zu $\epsilon > 0$ ein $N \in \mathbb{N}$ mit $\sum_{n=N+1}^{\infty} \frac{1}{2^n} < \frac{\epsilon}{2}$. Wähle für alle $n \in \{0, 1, \ldots, N\}$ ein $\mu_n \in \Lambda$ mit $p_n(x_\lambda - x) < \frac{\epsilon}{4}$ für alle $\lambda \geq \mu_n$. Wähle dann ein $\mu \geq \mu_n$ für alle $n \in \{0, 1, \ldots, N\}$. Es folgt

$$\begin{aligned} d(x_\lambda, x) &= \sum_{n=0}^{N} \frac{1}{2^n} \frac{p_n(x-y)}{1+p_n(x-y)} + \sum_{n=N+1}^{\infty} \frac{1}{2^n} \frac{p_n(x-y)}{1+p_n(x-y)} \\ &< \sum_{n=0}^{N} \frac{1}{2^n} \cdot \frac{\epsilon}{4} + \sum_{n=N+1}^{\infty} \frac{1}{2^n} \cdot 1 \leq \epsilon \ , \end{aligned}$$

d.h. Konvergenz bezüglich d von $(x_\lambda)_{\lambda\in\Lambda}$ gegen x. □

Definition 11.2 Ein metrisierbarer lokalkonvexer Vektorraum X, der bezüglich einer translations-invarianten Metrik vollständig ist, heißt *Fréchet-Raum*.

Bemerkung 11.3 Es läßt sich zeigen, daß diese Definition unabhängig von der Wahl einer translations-invarianten Metrik ist. Insbesondere kann die aus dem Halbnormensystem gewonnene Metrik genutzt werden.

Folgendes Kriterium ist nützlich zur Überprüfung der Vollständigkeit:

Lemma 11.4 *Sei X ein lokalkonvexer Vektorraum mit abzählbarem Halbnormensystem $\mathcal{P} = \{p_n : n \in \mathbb{N}\}$ und d die induzierte Metrik nach Satz 11.1. Dann sind für eine Folge $(x_k)_{k\in\mathbb{N}}$ äquivalent:*

i) *$(x_k)_{k\in\mathbb{N}}$ ist Cauchy-Folge bezüglich d.*

ii) *$(x_k)_{k\in\mathbb{N}}$ ist Cauchy-Folge bezüglich $\mathcal{P}$ in folgendem Sinn: Für alle $p_n \in \mathcal{P}$ und $\epsilon > 0$ gibt es ein (von p_n und ϵ abhängiges) $N \in \mathbb{N}$, so daß $p_n(x_k - x_l) < \epsilon$ für alle $k, l \geq N$.*

Beweis. Analog zu Satz 11.1.

Beispiel 11.5 $\mathcal{C}(\mathbb{R})$ versehen mit dem Halbnormensystem $p_n(f) := \sup_{t\in[-n,n]} |f(t)|$ ist Fréchet-Raum. Sei $(f_k)_{k\in\mathbb{N}}$ Cauchy-Folge und $t \in \mathbb{R}$ beliebig. Wähle ein $n > |t|$. Dann ist $|f_k(t) - f_l(t)| \leq p_n(f_k - f_l) < \epsilon$ für alle $k, l \geq N$, so daß $(f_k(t))_{k\in\mathbb{N}}$ in $\mathbb{K}$ konvergiert. Definiere eine Funktion $f : \mathbb{R} \to \mathbb{K}$ durch $f(t) := \lim_{k\to\infty} f_k(t)$. Dann konvergiert $(f_k)_{k\in\mathbb{N}}$ auf $[-n, n]$ gleichmäßig gegen f, also ist f auf $[-n, n]$ stetig. Dann ist aber $f \in \mathcal{C}(\mathbb{R})$, und $(f_k)_{k\in\mathbb{N}}$ konvergiert gleichmäßig gegen f auf jedem Intervall $[-n, n]$. Damit ist $(\mathcal{C}(\mathbb{R}), \mathcal{P})$ Fréchet-Raum. $\triangleleft$

Beispiel 11.6 Ohne Beweis erwähnen wir:

i) $\mathcal{C}^\infty(\mathbb{R})$ ist Fréchet-Raum bezüglich $\mathcal{P} = \{p_{n,m} : n, m \in \mathbb{N} : p_{n,m}(f) = \sup_{t\in[-n,n]} |f^{(m)}(t)|\}$

ii) $\mathcal{S}(\mathbb{R}^n)$ ist Fréchet-Raum bezüglich $\mathcal{P} = \{p_{m,\alpha} : m \in \mathbb{N}\ ,\ \alpha \in \mathbb{N}^n : p_{m,\alpha}(f) = \sup_{t\in\mathbb{R}^n} |(1 + \|t\|^m)(\partial^\alpha f)(t)|\}$. $\triangleleft$

Bemerkung 11.7 In Fréchet-Räumen gilt selbstverständlich der Satz von Baire. Damit übertragen sich auf Fréchet-Räume:

- das Prinzip der gleichmäßigen Beschränktheit (Theorem 6.6, Banach-Steinhaus),
- der Satz von der offenen Abbildung (Theorem 6.10),
- der Satz vom inversen Operator (Folgerung 6.12),
- der Satz vom abgeschlossenen Graphen (Theorem 6.16).

Bemerkung 11.8 Für Abbildungen zwischen Fréchet-Räumen X, Y läßt sich ein Ableitungsbegriff einführen, die Gâteaux-Ableitung. Sei $U \subset X$ offen, dann heißt $F : U \to Y$ Gâteau-differenzierbar in Richtung $x \in X$, falls für alle $u \in U$ der Grenzwert

$$(DF)(u; x) := \lim_{\tau\to 0} \frac{1}{\tau}\big(F(u + \tau x) - F(u)\big)$$

existiert. Dabei ist der Limes $\tau \to 0$ bezüglich der Topologie von Y zu verstehen. Existiert der Grenzwert für alle $x \in X$, so heißt F Gâteaux-differenzierbar. Die

Gâteaux-Ableitung ist homogen, $(DF)(u;\alpha x) = \alpha(DF)(u;x)$, aber nicht notwendig additiv (damit nicht notwendig linear) in x. Ist jedoch $DF : U \times X \to Y$ stetig (man sagt: F ist stetig Gâteaux-differenzierbar, kurz $\mathcal{C}^1$), so ist $(DF)(U,\ .\ ) : X \to Y$ linear (vgl. partielle/totale Differenzierbarkeit).

Für festgehaltenes x können höhere Ableitungen definiert werden:

$$(D^2F)(u;x_1,x_2) := \lim_{\tau\to 0} \frac{1}{\tau}\big(DF(u+\tau x_2;x_1) - DF(u;x_1)\big)\ .$$

Für (genügend oft) stetig Gâteaux-differenzierbare Funktionen gelten im wesentlichen alle Rechneregeln der klassischen Analysis (ohne Beweis):

i) Kettenregel: $(D(G\circ F))(u;x) = (DG)(F(u);(DF)(u;x))$

ii) Satz von Schwarz: $(DF)(u;x_1,x_2) = (DF)(u;x_2,x_1)$

iii) Hauptsatz der Differential- und Integralrechnung:
$$F(v) - F(u) = \int_0^1 d\tau\ (DF)(u+\tau(v-u);v-u)$$

iv) Taylor-Formel: $F(u+h) = F(u) + \sum_{l=1}^{k} \frac{1}{l!}(D^lF)(u;\underbrace{h,\dots,h}_{l}) + R_{k+1}$ mit
$$R_{k+1} = \frac{1}{k!}\int_0^1 d\tau\ (1-\tau)^k (D^{k+1}F)(u+\tau h;\underbrace{h,\dots,h}_{k+1})$$

Das Integral ist jeweils Grenzwert Riemannscher Summen mit Konvergenz in der Fréchet-Topologie.

Teil III
Hilbert-Räume

12 Grundlegende Eigenschaften

Wir verwenden die Konvention der Physik für das Skalarprodukt. Diese erweist sich später bei der Behandlung von Operatoralgebren als sinnvoller als die in der Mathematik übliche Konvention.

Definition 12.1 Ein *Skalarprodukt* auf einem $\mathbb{K}$-Vektorraum X ist eine Abbildung $\langle\ ,\ \rangle : X \times X \to \mathbb{K}$ mit

i) $\langle w, \lambda_1 v_1 + \lambda_2 v_2\rangle = \lambda_1\langle w, v_1\rangle + \lambda_2\langle w, v_2\rangle$ für alle $v_1, v_2, w \in X$ und $\lambda_1, \lambda_2 \in \mathbb{K}$.

ii) $\langle w, v\rangle = \overline{\langle v, w\rangle}$ für alle $v, w \in X$.

iii) $\langle v, v\rangle \geq 0$ für alle $v \in X$ und $\langle v, v\rangle = 0$ genau dann, wenn $v = 0$.

$(X, \langle\ ,\ \rangle)$ heißt *Prä-Hilbert-Raum* (oder *unitärer Vektorraum*).

Satz 12.2 (Cauchy-Schwarzsche Ungleichung) *In einem Prä-Hilbert-Raum $(X, \langle\ ,\ \rangle)$ gilt*

$$\big|\langle w, v\rangle\big|^2 \leq \langle v, v\rangle\langle w, w\rangle \qquad \text{für alle } v, w \in X$$

mit Gleichheit genau dann, wenn v und w linear abhängig sind.

Beweis. i) Sei $w \neq 0$ (für $w = 0$ ist die Ungleichung offensichtlich erfüllt), somit ist $\langle w, w\rangle \neq 0$. Dann gilt für beliebiges $\lambda \in \mathbb{K}$

$$\begin{aligned} 0 \leq \langle \lambda w - v, \lambda w - v\rangle &= \lambda\overline{\lambda}\langle w, w\rangle - \overline{\lambda}\langle w, v\rangle - \lambda\langle v, w\rangle + \langle v, v\rangle \\ &= \langle w, w\rangle\left|\lambda - \frac{\langle w, v\rangle}{\langle w, w\rangle}\right|^2 + \frac{1}{\langle w, w\rangle}\Big(\langle w, w\rangle\langle v, v\rangle - \big|\langle w, v\rangle\big|^2\Big)\ . \end{aligned}$$

Für $\lambda = \frac{\langle w, v\rangle}{\|w\|^2}$ wird direkt die Ungleichung von Cauchy-Schwarz erhalten.

ii) Gilt das Gleichheitszeichen (und sind $v, w \neq 0$), so folgt $v = \lambda w$ für $\lambda = \frac{\langle w, v\rangle}{\langle w, w\rangle}$. □

Folgerung 12.3 *Ein Prä-Hilbert-Raum $(X, \langle\ ,\ \rangle)$ wird durch $\|v\| := \sqrt{\langle v, v\rangle}$ zu einem normierten Vektorraum, und das Skalarprodukt ist in der Norm-Topologie beidseitig stetig.*

Beweis. Sei $v_n \to v$ und $w_n \to w$, dann

$$|\langle v_n, w_n\rangle - \langle v, w\rangle| = |\langle v_n - v, w_n\rangle + \langle v, w_n - w\rangle| \leq \|v_n - v\|\|w_n\| + \|v\|\|w_n - w\| \to 0\ .$$

□

Satz 12.4 *Es sei* $(X, \| \ \|)$ *ein normierter Vektorraum über* $\mathbb{K}$*. Die Norm* $\| \ \|$ *geht genau dann aus einem Skalarprodukt hervor, wenn die* Parallelogrammgleichung *gilt,*

$$\|v+w\|^2 + \|v-w\|^2 = 2\|v\|^2 + 2\|w\|^2 \qquad \textit{für alle } v,w \in X \; .$$

Beweisidee. Die Richtung ($\Rightarrow$) ist einfaches Nachrechnen. Für die Umkehrung ($\Leftarrow$) definiert man über die *Polarisationsformeln*

$$\begin{aligned} \langle v,w\rangle &= \tfrac{1}{4}\big(\|v+w\|^2 - \|v-w\|^2 - \mathrm{i}\|v+\mathrm{i}w\|^2 + \mathrm{i}\|v-\mathrm{i}w\|^2\big) && \text{für } \mathbb{K} = \mathbb{C}\;, \\ \langle v,w\rangle &= \tfrac{1}{4}\big(\|v+w\|^2 - \|v-w\|^2\big) && \text{für } \mathbb{K} = \mathbb{R} \end{aligned}$$

zunächst Abbildungen $\langle\ ,\ \rangle : X \times X \to \mathbb{K}$, für die man dann die Eigenschaften des Skalarprodukts zeigen muß, wobei die Parallelogrammgleichung eingeht. Für die Linearität ist das eine sehr mühsame Rechnung! Siehe [Hirzebruch-Scharlau].
□

Definition 12.5 Ein Prä-Hilbert-Raum $(X, \langle\ ,\ \rangle)$, der bezüglich der induzierten Norm vollständig (also Banach-Raum) ist, heißt *Hilbert-Raum*.

Beispiel 12.6 Hilbert-Räume (jeweils über $\mathbb{C}$) sind:

i) $\mathbb{C}^n$ mit $\langle v,w\rangle = \sum_{i=1}^n \overline{v_i} w_i$ wenn $v = (v_1,\dots,v_n)$, $w = (w_1,\dots,w_n)$.

ii) $\ell^2 = \ell^2(\mathbb{N})$ mit $\langle v,w\rangle = \sum_{i=0}^\infty \overline{v_i} w_i$ wenn $v = (v_i)_{i\in\mathbb{N}}$, $v = (w_i)_{i\in\mathbb{N}}$.

iii) $L^2(X,\mu)$ für beliebige Maßräume mit (X,μ) mit $\langle f,g\rangle = \displaystyle\int_X d\mu\ \bar{f} g$

iv) Sobolew-Raum $H^s(\mathbb{R}^n)$. Sei $\hat{f} \in L^2(\mathbb{R}^n)$ die kontinuierliche Fourier-Transformierte einer Funktion $f \in L^2(\mathbb{R}^n)$, d.h. $\hat{f}(k) := \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} d\lambda\ f(x)\ e^{\mathrm{i}\langle k,x\rangle}$. die kontinuierliche Fourier-Transformierte einer lokal-integrablen Funktion f. Der Sobolev-Raum zum Parameter $s > 0$ ist definiert als

$$H^s(\mathbb{R}^n) := \{ f \in L^2(\mathbb{R}^n) \ :\ (1+\|k\|^2)^{\frac{s}{2}} \hat{f}(k) \in L^2(\mathbb{R}^n)\}\;.$$

Er wird durch $\langle f,g\rangle := \displaystyle\int_{\mathbb{R}^n} d\lambda\ (1+\|k\|^2)^{\frac{s}{2}} \overline{\hat{f}(k)} \hat{g}(k)$ zu einem Hilbert-Raum. Die Sobolev-Räume spielen eine wichtige Rolle in der Theorie partieller Differentialgleichungen.

Lemma 12.7 *Sei* $(X, \langle\ ,\ \rangle)$ *Prä-Hilbert-Raum und* $\bar{X}$ *die Vervollständigung bezüglich* $\| \ \|$*. Dann ist* $\bar{X}$ *ein Hilbert-Raum.*

Beweis. Die Norm ist stetig auf $\bar{X}$. Damit setzt sich die Parallelogrammgleichung auf $\bar{X}$ fort und definiert somit ein Skalarprodukt über die Polarisationsformeln.
□

Wir werden Hilbert-Räume mit $(H, \langle\ ,\ \rangle)$ bezeichnen.

Satz 12.8 *Jeder Hilbert-Raum ist reflexiv und gleichmäßig konvex.* □

Lemma 12.9 *Sei $W \subset H$ abgeschlossene konvexe Teilmenge eines Hilbert-Raums H, dann gibt es zu $v \in H$ genau ein $w \in W$ mit $\|v - w\| = d(v, W) = \inf_{u \in W} \|v - u\|$.*

Beweis. Das ist Satz 5.12. Im komplexen Fall betrachte die Einschränkung von H auf den reellen gleichmäßig konvexen normierten Vektorraum $\mathrm{span}_{\mathbb{R}}(v, w_1, w_2)$. Daraus folgt die Eindeutigkeit. □

Definition 12.10 Sei $(H, \langle\ ,\ \rangle)$ (Prä-)Hilbert-Raum.

i) Zwei Vektoren $v, w \in H$ heißen *orthogonal*, geschrieben $v \perp w$, wenn $\langle v, w\rangle = 0$.

ii) Ist $W \subset H$ eine beliebige Teilmenge, dann heißt $W^\perp := \{v \in H \ : \ v \perp w \text{ für alle } w \in W\}$ der *Orthogonalraum* zu W.

Seien H_1, H_2 (Prä-)Hilbert-Räume. Durch das Skalarprodukt $\langle (v_1, v_2), (w_1, w_2)\rangle := \langle v_1, w_1\rangle + \langle v_2, w_2\rangle$, für $v_1, w_1 \in H_1$ und $v_2, w_2 \in H_1$ wird $H_1 \times H_2$ zu einem (Prä-)Hilbert-Raum, der mit $H_1 \oplus H_2$ bezeichnet wird. Diese Konstruktion heißt *direkte Summe von Hilbert-Räumen*.

Lemma 12.11 *Sei H Hilbert-Raum.*

i) *Für beliebige Teilmengen $W \subset H$ ist $W^\perp$ abgeschlossener Untervektorraum von H.*

ii) *Ist $X \subset H$ abgeschlossener Untervektorraum, so gilt $H = X \oplus X^\perp$ als direkte Summe von Hilbert-Räumen.*

Beweis. i) klar ist, daß $W^\perp$ Untervektorraum ist. Dieser ist abgeschlossen als $W^\perp = \bigcap_{w \in W} s_w^{-1}(\{0\})$ bezüglich der stetigen Abbildung $s_w : H \to \mathbb{K}$, $s_w(v) = \langle w, v\rangle$.

ii) Da X abgeschlossen und konvex ist, gibt es nach Lemma 12.9 genau ein $x_1 \in X$ mit $\|v - x_1\| = d(v, X)$. Wir behaupten: $x_2 := v - x_1 \in X^\perp$. Wegen der Minimalitätseigenschaft gilt für beliebige $0 \neq x \in X$:

$$\|x_2\|^2 = \|v - x_1\|^2 \leq \Big\| v - x_1 - \frac{\langle x, v - x_1\rangle}{\langle x, x\rangle} x \Big\| = \|x_2\|^2 - \frac{|\langle x, x_2\rangle|^2}{\|x\|^2} \,,$$

und damit $x \perp x_2$ für alle $x \in X$. Ist $x \in X \cap X^\perp$, so $\langle x, x\rangle = 0$, also $X \cap X^\perp = \{0\}$, so daß es nur eine Zerlegung $v = x_1 + x_2$ mit $x_1 \in X$ und $x_2 \in X^\perp$ gibt. Ist $w = y_1 + y_2$ die analoge Zerlegung für $w \in H$ in $y_1 \in X$ und $y_2 \in X^\perp$, so folgt $\langle v, w\rangle = \langle x_1, y_1\rangle + \langle x_2, y_2\rangle$. □

Satz 12.12 (Riesz) *Sei H ein Hilbert-Raum, $H' = (H, \|\ \|)'$ sein Dualraum und $j : H \to H'$ definiert durch $j(v)(w) := \langle v, w\rangle$. Dann gilt:*

i) *j ist antilinear, d.h.* $j(\lambda v + \mu w) = \overline{\lambda} j(v) + \overline{\mu} j(w)$*.*

ii) *j ist isometrisch, also injektiv.*

iii) *j ist surjektiv, also ein Anti-Isomorphismus. (gilt nicht für Prä-Hilbert-Räume!)*

Beweis. i) folgt aus der Definition.

ii) Zu zeigen ist $\|j(v)\| = \|v\|$ für alle $v \in H$. Das ist klar für $v = 0$. Ansonsten ($v \neq 0$) betrachten wir

$$\|j(v)\| = \sup_{w \in H\,,\; \|w\| \leq 1} |j(v)(w)| = \sup_{w \in H\,,\; \|w\| \leq 1} |\langle v, w\rangle| \leq \|v\| \;,$$

also $\|j(v)\| \leq \|v\|$. Andererseits ist für die spezielle Wahl $j(v)\big(\frac{v}{\|v\|}\big) = \|v\|$ und damit $\|j(v)\| \geq \|v\|$.

iii) Sei $0 \neq f \in H'$ und sei $X := \ker f \subset H$. Wegen der Stetigkeit von f ist X abgeschlossener Untervektorraum. Wegen $f \neq 0$ gibt es ein $0 \neq v \in H$ mit $f(v) \neq 0$. Damit ist $X \neq H$, nach Lemma 12.11 also $H = X \oplus X^\perp$ mit $X^\perp \neq \{0\}$, d.h. es gibt ein $v \in X^\perp$ mit $\|v\| = 1$. Setze $\lambda := f(v)$, dann gilt für alle $w \in H$

$$f\big(f(w)v - f(v)w\big) = f(w)f(v) - f(v)f(w) = 0 \;,$$

also $f(w)v - f(v)w \in X$. Weiter gilt

$$0 = \langle \underbrace{v}_{\in X^\perp}, \underbrace{f(w)v - f(v)w}_{\in X} \rangle = f(w)\|v\|^2 - f(v)\langle v, w\rangle \;,$$

also $f(w) = \lambda\langle v, w\rangle = \langle \overline{\lambda} v, w\rangle$ und damit $f = j(\overline{\lambda} v)$. $\square$

Insbesondere gilt in H' die Parallelogrammgleichung, so daß H' mit dem Skalarprodukt aus der Polarisationsformel selbst ein Hilbert-Raum ist. Der Rieszsche Darstellungssatz wird in vielen Konstruktionen benötigt, z.B. bei der Definition des adjungierten Operators.

13 Summierbarkeit und Orthonormalbasen

Wir zeigen, daß jeder Hilbert-Raum H eine Orthogonalbasis besitzt, durch die $H \simeq \ell^2(I)$ für eine geeignete Index-Menge wird. In separablen Hilbert-Räumen (I – abzählbar) kann eine solche konstruiert werden über das Verfahren von Gram-Schmidt. Für den allgemeinen Fall müssen wir zunächst einen allgemeinen Summierbarkeitsbegriff einführen.

Definition 13.1 Sei $(X, \|\;\|)$ normierter Vektorraum und $I \neq \varnothing$ eine Familie von Indizes. Sei $(x_i)_{i \in I}$ eine Familie von Elementen aus X. Sei $\Lambda := \{J \subset I \text{ endlich}\}$ versehen mit $J_1 \leq J_2 \quad \Leftrightarrow \quad J_1 \subset J_2$. Die Familie $(x_i)_{i \in I}$ heißt *summierbar*, falls es ein $x \in X$ gibt, so daß das Netz der Partialsummen $\{s_J\}_{J \in \Lambda}$ gegen x konvergent ist, mit $s_J := \sum_{j \in J} x_j$.

Anders gesagt ist $(x_i)_{i\in I}$ summierbar, wenn es zu jedem $\epsilon > 0$ eine endliche Teilmenge $J_\epsilon \subset I$ gibt, so daß für alle endlichen Teilmengen $J \subset I$ mit $J_\epsilon \subset J$ gilt $|x - \sum_{j\in J} x_j| < \epsilon$. Im folgenden bezeichnet $J, J_0, \dots, J_n, J_\epsilon, J' \dots$ immer eine endliche Teilfamilie einer Indexmenge I.

Lemma 13.2 *Sei $(X, \|\ \|)$ ein Banach-Raum.*

i) *Cauchy-Kriterium: Eine Familie $(x_i)_{i\in I}$ in X ist genau dann summierbar, wenn es für alle $\epsilon > 0$ ein $J_0 \subset I$ gibt mit $\|\sum_{j\in J'} x_j\| < \epsilon$ für alle $J' \subset I$ mit $J' \cap J_0 = \varnothing$.*

ii) *Die Familie $(x_i)_{i\in I}$ ist summierbar zu x genau dann, wenn höchstens abzählbar viele x_j ungleich 0 sind und wenn für jede Abzählung $x_1, x_2, \dots$ dieser $x_j \neq 0$ gilt $\lim_{N\to\infty} \sum_{n=1}^N x_n = x$.*

Beweis. Sei $(x_i)_{i\in I}$ summierbar zu x. Zu $\epsilon > 0$ gibt es ein J_0 mit $\|x - \sum_{j\in J_1} x_j\| < \frac{\epsilon}{2}$ für alle $J_1 \supset J_0$. Dann gilt für alle $J' \subset I$ mit $J' \cap J_0 = \varnothing$

$$\Big\|\sum_{j\in J'} x_j\Big\| = \Big\|\sum_{j\in J'\cup J_0} x_j - \sum_{j\in J_0} x_j\Big\| \le \Big\|\sum_{j\in J'\cup J_0} x_j - x\Big\| + \Big\|x - \sum_{j\in J_0} x_j\Big\| < \epsilon\ .$$

Umgekehrt gilt für beliebige $I \supset J_1, J_2 \supset J_0$:

$$\Big\|\sum_{j\in J_1} x_j - \sum_{j\in J_2} x_j\Big\| = \Big\|\sum_{j\in J_1\setminus J_0} x_j - \sum_{j\in J_2\setminus J_0} x_j\Big\| < 2\epsilon\ .$$

Somit bilden die Partialsummen $\big(\sum_{j\in J} x_j\big)_{J\supset J_0}$ ein Cauchy-Netz, welches wegen der Vollständigkeit von X gegen ein $x \in X$ konvergiert.

ii) Zu jedem $n \in \mathbb{N} \setminus \{0\}$ gibt es ein $J_n \subset I$, so daß für alle $J \subset I$ mit $J \cap J_n = \varnothing$ gilt: $\|\sum_{j\in J} x_j\| < \frac{1}{n}$. Insbesondere ist $\|x_i\| < \frac{1}{n}$ für alle $i \notin J_n$. Sei $I_0 := \bigcup_{n=1}^\infty J_n$, dann ist I_0 abzählbar, und für $i \notin I_0$ gilt $\|x_i\| < \frac{1}{n}$ für alle n, also $x_i = 0$.

Wir können $I = I_0$ abzählbar annehmen. Sei $\{J_n\}_{n=1}^\infty$ eine beliebige Familie endlicher Teilmengen von I_0 mit $J_n \subset J_{n+1}$ und $I_0 = \bigcup_{n=1}^\infty J_n$. Nach Definition der Summierbarkeit gibt es zu $\epsilon > 0$ ein $J_0 \subset I_0$ mit $\|x - \sum_{j\in J}\| < \epsilon$ für alle $J \supset J_0$. Da J_0 endlich ist, gilt $J_0 \subset J_N$ für ein $N \in \mathbb{N}$. Also ist $\|x - \sum_{j\in J_N}\| < \epsilon$, und damit die Summe unabhängig von der Anordnung der Elemente.

Die Umkehrung von ii) ist klar. □

Beispiel 13.3 Die alternierende harmonische Reihe $\big(\frac{(-1)^n}{n+1}\big)_{n\in\mathbb{N}}$ ist nicht summierbar. ◁

Absolute Summierbarkeit ist hinreichend:

Lemma 13.4 *Sei $(X, \|\ \|)$ Banach-Raum. Ist $(\|x_i\|)_{i\in I}$ summierbar in $\mathbb{R}$, so ist $(x_i)_{i\in I}$ summierbar in X, und es gilt $\|\sum_{i\in I} x_i\| \le \sum_{i\in I} \|x_i\|$.*

Beweis. Zu $\epsilon > 0$ gibt es ein $J_0 \subset I$, so daß für alle $J' \in I$ mit $J' \cap J_0 = \varnothing$ gilt $\sum_{j\in J'} \|x_j\| < \epsilon$. Da es endliche Summen sind, gilt $\| \sum_{j\in J'} x_j\| \leq \sum_{j\in J'} \|x_j\| < \epsilon$, d.h. $(x_i)_{i\in I}$ ist summierbar nach Lemma 13.2.i). Damit können wir nach Lemma 13.2.ii) eine Abzählung wählen, und die verallgemeinerte Dreiecksungleichung folgt dann wie üblich. □

Definition 13.5 Sei H ein (Prä-)Hilbert-Raum. Eine Familie $(v_i)_{i\in I}$ heißt *Orthogonalsystem*, falls $v_i \perp v_{i'}$ für alle $i, i' \in I$ mit $i \neq i'$, und *Orthonormalsystem* (ONS), falls zusätzlich $\|v_i\| = 1$.

Für Orthogonalsysteme in Hilbert-Räumen gilt die Umkehrung von Lemma 13.4:

Lemma 13.6 *Sei H ein Hilbert-Raum und $v_i \in H$.*

i) *Ist $(v_i)_{i\in I}$ summierbar zu v, so folgt $\langle w, v\rangle = \sum_{i\in I}\langle w, v_i\rangle$ für alle $w \in H$.*

ii) *Pythagoras: Ist $(v_i)_{i\in I}$ ein Orthogonalsystem, so ist $(v_i)_{i\in I}$ genau dann in H summierbar, wenn $(\|v_i|^2)_{i\in I}$ in $\mathbb{R}$ summierbar ist, und es gilt $\| \sum_{i\in I} v_i\|^2 = \sum_{i\in I} \|v_i\|^2$.*

Beweis. i) Für eine beliebige Abzählung der $v_i \neq 0$ konvergiert nach Lemma 13.2 die Folge ihrer Partialsummen gegen v. Das Skalarprodukt ist stetig, so daß Grenzwert und Skalarprodukt vertauschen.

ii) Wegen der Orthogonalität gilt $\| \sum_{j\in J'} v_j\|^2 = \sum_{j\in J'} \|v_i\|^2$ für alle endlichen Teilmengen $J' \subset I$. Sind ϵ, J_0 gegeben, dann gilt für alle J' mit $J' \cap J_0$:

$$(\|v_i\|^2)_{i\in I} \text{ summierbar } \Leftrightarrow \sum_{j\in J'} \|v_i\|^2 = \| \sum_{j\in J'} v_j\|^2 < \epsilon^2 \Leftrightarrow (v_i)_{i\in I} \text{ summierbar}\,.$$

Schließlich folgt für $v = \sum_{i\in I} v_i$ die Gleichung $\langle v, v\rangle = \sum_{i\in I}\langle v_i, v\rangle = \sum_{i,i'\in I}\langle v_i, v_{i'}\rangle = \sum_{i\in I}\langle v_i, v_i\rangle$. □

Satz 13.7 *Für ein ONS $(v_i)_{i\in I}$ in einem Hilbert-Raum H gilt:*

i) *Besselsche Ungleichung:* $\displaystyle\sum_{i\in I} |\langle v_i, v\rangle|^2 \leq \|v\|^2$ *für alle $v \in H$.*

ii) *Parsevalsche Gleichung:* $\displaystyle\sum_{i\in I} |\langle v_i, v\rangle|^2 = \|v\|^2$ *genau dann, wenn* $v = \displaystyle\sum_{i\in I}\langle v_i, v\rangle v_i$.

Beweis. Für alle endlichen Teilmengen $J \subset I$ gilt

$$0 \leq \Big\|v - \sum_{j\in J}\langle v_j, v\rangle v_j\Big\|^2 = \|v\|^2 - \sum_{j\in J} |\langle v_j, v\rangle|^2\,.$$

Damit existiert $s := \sup_{J\subset I} \sum_{j\in J} |\langle v_j, v\rangle|^2 \leq \|v\|^2$, und $(|\langle v_i, v\rangle|^2)_{i\in I}$ ist in $\mathbb{R}$ summierbar zu s, wobei wegen $s \leq \|v\|^2$ die Besselsche Ungleichung gilt. Gilt in der Besselschen Ungleichung Gleichheit, so folgt $v = \sum_{i\in I} \langle v_i, v\rangle v_i$. □

Es wird nicht behauptet, daß das ONS $(v_i)_{i\in I}$ summierbar ist (was auch nicht gilt)! Dennoch sind $(|\langle v_i, v\rangle|^2)_{i\in I}$ summierbar in $\mathbb{R}$ und $(\langle v_i, v\rangle v_i)_{i\in I}$ summierbar in H, und zwar für alle $v \in H$.

Satz/Definition 13.8 (Orthonormalbasis) *Für ein ONS $(v_i)_{i\in I}$ in einem Hilbert-Raum H sind äquivalent:*

i) *$(v_i)_{i\in I}$ ist maximal, d.h. für jedes ONS $(w_i)_{i\in I}$, welches $(v_i)_{i\in I}$ enthält, gilt $(v_i)_{i\in I} = (w_i)_{i\in I}$ als Mengen.*

ii) *Gilt $v \perp v_i$ für alle $v \in X$, so folgt $v = 0$.*

iii) *$H = \overline{\operatorname{span}\{v_i \;:\; i \in I\}}$ (Norm-Abschluß).*

iv) *Die Abbildung $T : \ell^2(I) \ni (x_i)_{i\in I} \mapsto \sum_{i\in I} x_i v_i \in H$ ist ein isometrischer Isomorphismus mit Umkehrabbildung $T^{-1}(v) = (\langle v_i, v\rangle)_{i\in I}$.*

v) *Für alle $v \in H$ gilt $v = \sum_{i\in I} \langle v_i, v\rangle v_i$ (Fourier-Entwicklung).*

vi) *Für alle $v, w \in H$ gilt $\langle v, w\rangle = \sum_{i\in I} \langle v, v_i\rangle\langle v_i, w\rangle$.*

vii) *Für alle $v \in H$ gilt $\|v\|^2 = \sum_{i\in I} |\langle v_i, v\rangle|^2$ (Parsevalsche Gleichung).*

Ein ONS $(v_i)_{i\in I}$, das eine (und damit alle) der Bedingungen i)–vii) erfüllt, heißt Orthonormalbasis (ONB).

Beweis. i)⇒ii) Wäre $v \neq 0$ und $v \perp v_i$, so wäre $(x_i)_{i\in I} \bigcup \{\frac{v}{\|v\|}\}$ ein größeres ONS.

ii⇒iii) $X := \overline{\operatorname{span}\{v_i \;:\; i \in I\}}$ ist abgeschlossener Untervektorraum von H. Wäre $X \neq H$, dann $H = X \oplus X^\perp$ mit $X^\perp \neq \{0\}$. Somit gibt es ein $0 \neq v \in X^\perp \subset H$ mit $v \perp v_i$ für alle $i \in I$, im Widerspruch zu ii).

iii⇒iv) Nach Lemma 13.6.ii) ist $(x_i v_i)_{i\in I}$ in H summierbar genau dann, wenn $(\|x_i v_i\|^2)_{i\in I}$ in $\mathbb{R}$ summierbar ist, d.h. wenn $x \in \ell^2(I)$ wegen $\|x_i x_i\|^2 = |x_i|^2$. Die Isometrie von T ist dann Lemma 13.6.ii) (Pythagoras). Es gilt $\operatorname{span}\{v_i \;:\; i \in I\} \subset T(\ell^2(I)) \subset H$, also ist $T(\ell^2(I))$ dicht in H. Weiter gilt $\ell^2(I) = L^2(I, \mu)$ für das Zählmaß μ, also ist $\ell^2(I)$ vollständig. Dann ist auch $T(\ell^2(I))$ vollständig, insbesondere Norm-abgeschlossen, somit gleich H.

iv⇒v) Jedes $v \in H$ besitzt eine eindeutige Darstellung $v = \sum_{i\in I} x_i v_i$ mit $(x_i)_{i\in I} \in \ell^2(I)$. Nach Lemma 13.6.i) gilt $\langle v_j, v\rangle = \sum_{i\in I} x_i\langle v_j, v_i\rangle = x_j$. Insbesondere ist $T^{-1}(v) = (\langle v_i, v\rangle)_{i\in I}$.

v⇒vi) Sei $v = \sum_{i\in I}\langle v_i, v\rangle v_i$ und $w = \sum_{i'\in I}\langle v_{i'}, w\rangle v_{i'}$, dann gilt nach Lemma 13.6.i): $\langle v, w\rangle = \sum_{i,i'\in I}\overline{\langle v_i, v\rangle}\langle v_{i'}, w\rangle\langle v_i, v_i'\rangle = \sum_{i\in I}\langle v, v_i\rangle\langle v_i, w\rangle$.

vi⇒vii) setze $w = v$.

vii⇒i) Existenz eines weiteren $v \in H$ mit $\|v\| = 1$ und $v \perp v_i$ für alle i ist im Widerspruch zu vii). □

Folgerung 13.9 *Jeder Hilbert-Raum $H \neq \{0\}$ besitzt eine Orthonormalbasis.*

Beweis. Die nichtleere Menge $\mathcal{B}$ der ONS in H ist geordnet. Sei K eine Kette, dann ist $S := \bigcup_{B\subset K} B$ eine obere Schranke von K. Nach dem Lemma 3.2 von Zorn besitzt $\mathcal{B}$ ein maximales Element, nach Satz/Definition 13.8.i) ist dieses eine ONB. □

Folgerung 13.10 *Jeder Hilbert-Raum $H \neq \{0\}$ ist isometrisch isomorph zu $\ell^2(I)$ für eine geeignete Index-Menge I.*

Beweis. Satz/Definition 13.8.iv) und Folgerung 13.9. □

Sei H separabel (in der Norm-Topologie), d.h. es gibt (nach Lemma 4.20) eine abzählbare Teilmenge $A = \{x_n : n \in \mathbb{N}\} \subset H$ mit $H = \overline{\mathrm{span}(x_n : n \in \mathbb{N}\}}$. Durch Entfernen überflüssiger Elemente kann $(x_n)_{n\in\mathbb{N}}$ als linear unabhängig angenommen werden. Das Orthonormalisierungsverfahren von Gram-Schmidt überführt $(x_n)_{n\in\mathbb{N}}$ dann in ein ONS $(v_n)_{n\in\mathbb{N}}$ mit $H = \overline{\mathrm{span}(v_n : n \in \mathbb{N}\}}$. Nach Satz/Definition 13.8 ist dann $(v_n)_{n\in\mathbb{N}}$ eine ONB. Somit besitzt jeder separable Hilbert-Raum eine abzählbare ONB und ist isometrisch isomorph zum Hilbertschen Folgenraum $\ell^2(\mathbb{N})$.

Satz 13.11 (Gram-Schmidt) *Sei H ein (Prä-)Hilbert-Raum und $(x_n)_{n\in\mathbb{N}}$ eine linear unabhängige Familie von Vektoren. Dann gibt es ein ONS $(v_n)_{n\in\mathbb{N}}$ mit* $\mathrm{span}\{x_n : n \in \mathbb{N}\} = \mathrm{span}\{v_n : n \in \mathbb{N}\}$.

Beweis. Aus der linearen Unabhängigkeit folgt $x_n \neq 0$ für alle n. Setze $v_0 := \frac{x_0}{\|x_0\|}$ und konstruiere v_n per Induktion. Seien $v_0, \dots, v_k$ bereits konstruiert, mit $\langle v_i, v_j\rangle = \delta_{ij}$ und $\mathrm{span}(x_0, \dots, x_k) = \mathrm{span}(v_0, \dots, v_k)$. Setze $\tilde{v}_{k+1} := x_{k+1} - \sum_{j=0}^{k}\langle v_k, x_{k+1}\rangle v_k$. Dann ist $\tilde{v}_{k+1} \neq 0$, denn ansonsten wäre $x_{k+1} \in \mathrm{span}(v_0, \dots, v_k) = \mathrm{span}(x_0, \dots, x_k)$. Nach Konstruktion gilt $\tilde{v}_{k+1} \perp v_j$ für alle $j = 0, \dots, k$. Setze $v_{k+1} := \frac{\tilde{v_{k+1}}}{\|\tilde{v_{k+1}}\|}$, dann ist $(v_0, \dots, v_{k+1})$ ein ONS. Nach Basisaustauschsatz (der linearen Algebra) gilt $\mathrm{span}(v_0, \dots, v_{k+1}) = \mathrm{span}(v_0, \dots, v_k, \tilde{v}_{k+1}) = \mathrm{span}(v_0, \dots, v_k, x_{k+1}) = \mathrm{span}(x_0, \dots, x_{k+1})$. □

Beispiel 13.12 In den Übungen wird gezeigt: $L^2([0, 2\pi])$ ist separabel, und eine mögliche Orthonormalbasis ist gegeben durch $(e_n)_{n\in\mathbb{Z}}$ mit $e_n(t) = e^{int}$. Die Dichtheit der stetigen 2π-periodischen Funktionen in $L^2([0, 2\pi])$ folgt daraus, daß die

Treppenfunktionen dicht in $L^2([0,2\pi])$ sind bezüglich $\| \ \|_2$, und jede Treppenfunktion läßt sich bezüglich $\| \ \|_2$ durch stetige Funktionen approximieren. Die 2π-Periodizität ist für $L^2([0,2\pi])$ unerheblich.

Somit gilt für beliebige Funktionen $[f] \in L^2([0,2\pi])$ die Fourier-Entwicklung $f = \sum_{n\in\mathbb{Z}} \langle e_n, f\rangle e_n$, d.h. die Reihe konvergiert in $L^2([0,2\pi])$ gegen f und damit fast überall punktweise. In den Übungen wurde als Folgerung des Prinzips der gleichmäßigen Beschränktheit gezeigt, daß selbst für stetige 2π-periodische Funktionen die Fourier-Reihe nicht überall punktweise konvergiert. Dagegen gilt: Ist f stetig auf $[0,2\pi]$ und stückweise stetig differenzierbar, so konvergiert die Fourier-Reihe von f sogar gleichmäßig gegen f. $\triangleleft$

Analog ist auch $L^2([a,b])$ separabel für beliebige $a,b \in \mathbb{R}$ mit $a < b$, und durch Translation und Skalierung der Fourier-ONB aus Beispiel 13.12 wird eine ONB in $L^2([a,b])$ erhalten. Es kann aber sinnvollere ONB geben, die sich z.B. durch *Orthogonale Polynome* ausdrücken:

Beispiel 13.13 i) $\big(\sqrt{n+\frac{1}{2}}P_n(t)\big)_{n\in\mathbb{N}}$ ist ONB in $L^2(]-1,1[)$, dabei sind $P_n(t) := \frac{1}{2^n n!}\frac{d^n}{dt^n}(t^2-1)^n$ die *Legendreschen Polynome*. Andere wichtige ONB in $L^2(]-1,1[)$ sind durch Jacobi-Polynome, Tschebyschew-Polynome und Gebenbauer-Polynome definiert.

ii) $\big(e^{-\frac{t}{2}}L_n(t)\big)_{n\in\mathbb{N}}$ ist ONB in $L^2(]0,\infty[)$, dabei sind $L_n(t) := \frac{1}{n!}e^t\frac{d^n}{dt^n}(t^n e^{-t})$ die *Laguerreschen Polynome*. Die zugeordneten Laguerre-Polynome definieren eine andere wichtige ONB in $L^2(]0,\infty[)$.

iii) $\big(\frac{1}{\sqrt{2^n\sqrt{\pi}n!}}e^{-\frac{t^2}{2}}H_n(t)\big)_{n\in\mathbb{N}}$ ist ONB in $L^2(\mathbb{R})$, dabei sind $H_n(t) := (-1)^n e^{t^2}\frac{d^n}{dt^n}(e^{-t^2})$ die *Hermiteschen Polynome*.

Diese ONB entstehen als Lösungen von Differentialgleichungen 2. Ordnung. Vereinfacht gesagt definieren diese Differentialgleichungen einen selbstadjungierten Operator $T : H \to H$, für den es ein Eigenwertproblem gibt. Die zugehörigen Eigenfunktionen sind orthogonal und bilden damit die ONB. $\triangleleft$

14 Operatoren auf Hilbert-Räumen

Sei H ein Hilbert-Raum. Wir nennen eine lineare stetige Abbildung $A \in \mathcal{L}(H) := \mathcal{L}(H,H)$ einen linearen stetigen Operator auf H, meistens nur "Operator". Unser Ziel ist die Konstruktion des zu A adjungierten Operators A^*. Wir betrachten die allgemeinere Situation $A \in \mathcal{L}(H_1,H_2)$.

Satz/Definition 14.1 *Seien H_1, H_2 Hilbert-Räume und $A \in \mathcal{L}(H_1,H_2)$. Dann gibt es genau ein $A^* \in \mathcal{L}(H_2,H_1)$, so daß für alle $v \in H_1$ und $w \in H_2$ gilt:*

$$\langle w, Av\rangle = \langle A^*w, v\rangle \tag{*}$$

Diese durch (*) eindeutig bestimmte lineare stetige Abbildung $A^* \in \mathcal{L}(H_2, H_1)$ *heißt der zu* A *adjungierte Operator. Es gilt* $\|A^*\| = \|A\|$.

Beweis. Aus der Stetigkeit der Abbildung $v \mapsto Av$ und der Stetigkeit des Skalarprodukts folgt, daß für alle $w \in H_2$ das lineare Funktional $f(v) := \langle w, Av\rangle_{H_2}$ auf H_1 stetig ist. Nach dem Rieszschen Darstellungssatz existiert genau ein Element $A^*w \in H_1$ mit $\langle w, Av\rangle_{H_2} = f(v) = \langle A^*w, v\rangle_{H_1}$. Ist $w = \lambda_1 w_1 + \lambda_2 w_2$, Dann gibt es auch zu w_1, w_2 eindeutig bestimmte Vektoren $A^*w_1, A^*w_2 \in H_1$ mit $\langle w_1, Av\rangle_{H_2} = \langle A^*w_1, v\rangle_{H_1}$ und $\langle w_2, Av\rangle_{H_2} = \langle A^*w_2, v\rangle_{H_1}$. Somit gilt

$$\langle A^*w, v\rangle_{H_1} = \langle w, Av\rangle_{H_2} = \overline{\lambda_1}\langle w_1, Av\rangle_{H_2} + \overline{\lambda_2}\langle w_2, Av\rangle_{H_2} = \langle \lambda_1 A^*w_1 + \lambda_2 A^*w_2, v\rangle_{H_1} .$$

Subtrahieren wir $\langle A^*w, v\rangle_{H_1}$ und setzen $v = \lambda_1 A^*w_1 + \lambda_2 A^*w_2 - A^*w$, so folgt die Linearität von $A^* : H_2 \to H_1$. Für die Operatornormen gilt mit $\|w\|_{H_2} \leq 1$ und Cauchy-Schwarz

$$\|A^*w\|^2_{H_1} = \langle A^*w, A^*w\rangle_{H_1} = \langle w, AA^*w\rangle_{H_2} \leq \|w\|_{H_2}\|AA^*w\|_{H_2} \leq \|A\|\|A^*w\|_{H_1} ,$$

also $\|A^*\| \leq \|A\|$. Durch analoge Abschätzung von $\|Aw\|_{H_2}$ beweist man die Umkehrung $\|A^*\| \leq \|A\|$. Damit ist A^* stetig. $\square$

Einige offensichtiche Konsequenzen sind:

i) Seien $A, A_1, A_2 \in \mathcal{L}(H_1, H_2)$ und $\lambda_1, \lambda_2 \in \mathbb{K}$, dann gilt

$$(\lambda_1 A_1 + \lambda_2 A_2)^* = \overline{\lambda_1} A_1^* + \overline{\lambda_2} A_2^* , \qquad (A^*)^* = A .$$

ii) Ist H_3 ein weiterer Hilbert-Raum und $A \in \mathcal{L}(H_1, H_2)$, $B \in \mathcal{L}(H_2, H_3)$, dann gilt $(BA)^* = A^*B^* \in \mathcal{L}(H_3, H_1)$. Dabei ist $AB := A \circ B$, und es gilt $\|AB\| \leq \|A\|\,\|B\|$.

Weniger offensichtlich ist:

Satz 14.2 *Seien* H_1, H_2 *Hilbert-Räume und* $A \in \mathcal{L}(H_1, H_2)$. *Dann gilt:*

i) $\|A^*A\| = \|A\|^2$ *(C^*-Eigenschaft)*

ii) $\ker A = (\operatorname{im} A^*)^\perp$

iii) $\ker A^* = (\operatorname{im} A)^\perp$

Dabei bezeichent $\operatorname{im} A \subset H_2$ *das Bild von* $A \in \mathcal{L}(H_1, H_2)$.

Beweis. i) Nach Dreiecksungleichung gilt $\|A^*A\| \leq \|A^*\|\|A\| = \|A\|^2$. Für die Umkehrung gilt mit Cauchy-Schwarz

$$\|Av\|^2 = \langle Av, Av\rangle = \langle A^*Av, v\rangle \leq \|A^*Av\|\,\|v\| \leq \|A^*A\|\,\|v\|^2$$

und dann $\|A\| = \sup_{v \in B_{H_1}} \|Av\| \leq \sqrt{\|A^*A\|}$.

ii) Sei $v \in H_1$ mit $Av = 0$. Für alle $w \in H_2$ gilt $0 = \langle w, Av \rangle = \langle A^* w, v \rangle$, also $\ker A \subset (\mathrm{im}\ A^*)^\perp$. Sei umgekehrt $v \in (\mathrm{im}\ A^*)^\perp$, d.h. für alle $w \in H_2$ gilt $0 = \langle A^* w, v \rangle$. Dann folgt $\langle w, Av \rangle = 0$ für alle $w \in H_2$, insbesondere für $w = Av$, somit $Av = 0$.

iii) aus $A^{**} = A$. □

Lemma 14.3 *Sei H Hilbert-Raum und $X \subset H$ Unterverktorrraum. Dann gilt $\bar{X} = (X^\perp)^\perp =: X^{\perp\perp}$.*

Beweis. Sei $w \in X^\perp$ und $v \in \bar{X}$. Zu jedem $\epsilon > 0$ gibt es ein $v_\epsilon \in X$ mit $\|v - v_\epsilon\| < \epsilon$. Dann folgt $|\langle v, w \rangle| = |\langle v - v_\epsilon, w \rangle| \leq \epsilon \|w\|$, also auch $w \in \bar{X}^\perp$. Somit $\bar{X}^\perp = X^\perp$. Nach Lemma 12.11 gilt: $X^\perp$ ist abgeschlossen, und $H = X^\perp \oplus X^{\perp\perp}$. Andererseits ist $H = \bar{X}^\perp \oplus \bar{X}$, und aus der Eindeutigkeit der Zerlegung folgt die Behauptung. □

Folgerung 14.4 *Für $A \in \mathcal{L}(H_1, H_2)$ gilt $\overline{\mathrm{im}\ A} = (\ker A^*)^\perp$ und $\overline{\mathrm{im}\ A^*} = (\ker A)^\perp$.* □

Folgerung 14.5 *$A \in \mathcal{L}(H_1, H_2)$ injektiv $\Leftrightarrow$ $\overline{\mathrm{im}\ A^*} \subset H$ dicht.* □

Die folgenden Definitionen besonderer Operatoren sind von großer Bedeutung.

Definition 14.6 Sei H Hilbert-Raum und $A, N, E, P \in \mathcal{L}(H)$. Mit $\mathbf{1}_H \in \mathcal{L}(H)$ werde der identische Operator bezeichnet, $\mathbf{1}_H v = v$ für alle $v \in H$.

i) A heißt *selbstadjungiert* (oder *hermitesch*), falls $A = A^*$.

ii) N heißt *normal*, falls $N^* N = N N^*$.

iii) E heißt *idempotent*, falls $E^2 = E$.

iv) P heißt *(Orthogonal-)Projektor* oder *(orthogonale) Projektion*, falls $P = P^2 = P^*$.

Seien H_1, H_2 Hilbert-Räume und $U, S \in \mathcal{L}(H_1, H_2)$.

vi) U heißt *unitär*, falls $U^* U = \mathbf{1}_{H_1}$ *und* $U U^* = \mathbf{1}_{H_2}$.

vii) S heißt *isometrisch* oder *Isometrie*, falls $S^* S = \mathbf{1}_{H_1}$.

viii) S heißt *partielle Isometrie*, falls $S^* S \in \mathcal{L}(H_1)$ Projektor ist.

Einige offensichtliche Konsequenzen sind:

i) Jeder selbstadjungierte Operator (somit jeder Projektor) ist normal.

ii) Für jeden Operator $B \in \mathcal{L}(H_1, H_2)$ ist $B^* B \in \mathcal{L}(H_1)$ selbstadjungiert. Für $B \in \mathcal{L}(H)$ ist $B + B^*$ selbstadjungiert.

iii) Ist H komplexer Hilbert-Raum, dann besitzt jeder Operator $B \in \mathcal{L}(H)$ eine eindeutige Zerlegung $B = A_1 + \mathrm{i}A_2$ mit A_1, A_2 selbstadjungiert: Setze $A_1 := \frac{1}{2}(B + B^*)$ und $A_2 := \frac{1}{2\mathrm{i}}(B - B^*)$.

iv) Jeder unitäre Operator ist isometrisch.

v) Ist $U \in \mathcal{L}(H)$ unitär, so ist U auch normal.

Lemma 14.7 i) *$S \in \mathcal{L}(H_1, H_2)$ ist genau dann Isometrie, wenn $\langle Sv_1, Sv_2 \rangle_{H_2} = \langle v_1, v_2 \rangle_{H_1}$ für alle $v_1, v_2 \in H_1$, bzw. wenn $\|Sv\| = \|v\|$ für alle $v \in H_1$.*

ii) *$U \in \mathcal{L}(H_1, H_2)$ ist genau dann unitär, wenn U surjektiv ist und $\langle Uv_1, Uv_2 \rangle_{H_2} = \langle v_1, v_2 \rangle_{H_1}$ für alle $v_1, v_2 \in H_1$, bzw. wenn U surjektiv ist und $\|Uv\| = \|v\|$ für alle $v \in H_1$.*

iii) *$A \in \mathcal{L}(H)$ ist genau dann selbstadjungiert, wenn $\langle Av_1, v_2 \rangle = \langle v_1, Av_2 \rangle$ für alle $v_1, v_2 \in H$.*

iv) *$N \in \mathcal{L}(H)$ ist genau dann normal, wenn $\langle Av_1, Av_2 \rangle = \langle A^*v_1, A^*v_2 \rangle$ für alle $v_1, v_2 \in H$.*

Beweis. i) ($\Rightarrow$) ist klar. Gilt umgekehrt $\langle v_1, v_2 \rangle_{H_1} = \langle Sv_1, Sv_2 \rangle_{H_2} = \langle S^*Sv_1, v_2 \rangle_{H_1}$ für alle $v_1, v_2 \in H_1$, so folgt $\langle (S^*Sv_1 - v_1), v_2 \rangle_{H_1} = 0$ für alle $v_2 \in H_1$, insbesondere für $v_2 = S^*Sv_1 - v_1$. Damit ist $S^*Sv_1 = v_1$ für alle $v_1 \in H_1$. Die weitere Äquivalenz ergibt sich aus den Polarisationsformeln.

ii) Ist U isometrisch und surjektiv, so ist U bijektiv, und nach dem Satz vom inversen Operator (Folgerung 6.12) ist $U^{-1} \in \mathcal{L}(H_2, H_1)$. Für alle $v \in H_1$ und $w \in H_2$ gilt

$$\langle v, U^{-1}w \rangle_{H_1} = \langle Uv, UU^{-1}w \rangle_{H_2} = \langle Uv, w \rangle_{H_2} = \langle v, U^*w \rangle_{H_1} \ .$$

also $U^{-1} = U^*$. Ist umgekehrt $UU^* = \mathbf{1}_{H_2}$, dann ist $w = U(U^*w)$ für alle $w \in H_2$, also U surjektiv.

iii) ($\Rightarrow$) ist klar. Gilt umgekehrt $\langle v_1, Av_2 \rangle = \langle Av_1, v_2 \rangle$ für alle $v_1, v_2 \in H_1$, so folgt $\langle (A^*v_1 - Av_1), v_2 \rangle = 0$ für alle $v_2 \in H$, somit $A^* = A$. Analog folgt iv). □

In endlich-dimensionalen Vektorräumen ist jede injektive lineare Abbildung auch surjektiv. Im Unendlich-dimensionalen gilt das nicht!

Beispiel 14.8 (einseitiger Shift) Sei H ein separabler unendlich-dimensionaler Hilbert-Raum mit Orthonormalbasis $(v_n)_{n \in \mathbb{N}}$. Der durch $Sv_n := v_{n+1}$ für alle $n \in \mathbb{N}$ und lineare Fortsetzung definierte lineare Operator $S : H \to H$ heißt *einseitiger Shift*. Ist $v = \sum_{n=0}^{\infty} \lambda_n v_n$, mit $\|v\|^2 = \sum_{n=0}^{\infty} |\lambda_n|^2$, dann ist $Sv = \sum_{n=0}^{\infty} \lambda_n v_{n+1}$ und damit $\|Sv\| = \|v\|$ für alle $v \in H$, d.h. S ist Isometrie. Es gilt $\langle S^*v_n, v_m \rangle = \langle v_n, Sv_m \rangle = \delta_{m,n+1}$, also $S^*v_n = v_{n-1}$ für

$n \geq 1$ und $S^* v_0 = 0$. Damit ist $\ker S^* = \mathbb{K} v_0$, und nach Satz 14.2.iii) folgt $(\operatorname{im} S)^\perp = \mathbb{K} v_0$. Somit ist S nicht surjektiv, also auch nicht unitär.

Sei jetzt $P \in \mathcal{L}(H)$ definiert durch $Pv = \langle v_0, v\rangle v_0$. Dann ist $P^2 = P$, außerdem gilt für beliebige $v, w \in H$

$$\langle w, Pv\rangle = \langle w, \langle v_0, v\rangle v_0\rangle = \langle w, v_0\rangle\langle v_0, v\rangle = \langle\langle v_0, w\rangle v_0, v\rangle$$

und damit $P = P^*$ nach Lemma 14.7.iii), d.h. P ist Projektor. Dann ist auch $\mathbf{1}_H - P$ Projektor (Übungen), und es gilt $SS^* = \mathbf{1}_H - P$. Wegen $SS^* = (S^*)^* S^*$ ist S^* eine partielle Isometrie. $\triangleleft$

Lemma 14.9 *Sei $P \in \mathcal{L}(H)$ Projektor. Dann gilt:*

i) $P = 0$ *oder* $\|P\| = 1$

ii) $\operatorname{im} P = \ker(\mathbf{1}_H - P)$, *insbesondere ist* $\operatorname{im} P$ *abgeschlossen.*

iii) $H = \operatorname{im} P \oplus \ker P$.

Beweis. i) folgt aus $\|P\| = \|P^2\| = \|P^* P\| = \|P\|^2$.

ii) Sei $w \in \operatorname{im} P$, also $w = Pv$ für ein $v \in H$. Dann gilt $(\mathbf{1}_H - P)w = (P - P^2)v = 0$, also $\operatorname{im} P \subset \ker(\mathbf{1}_H - P)$. Ist umgekehrt $(\mathbf{1}_H - P)w = 0$, dann $w = Pw \in \operatorname{im} P$.

iii) $H = \ker P \oplus (\ker P)^\perp = \ker P \oplus \overline{\operatorname{im} P^*} = \ker P \oplus \operatorname{im} P$. $\square$

Bemerkung 14.10 Sei $S \in \mathcal{L}(H_1, H_2)$ eine partielle Isometrie, d.h. $S^* S = P$ für eine Projektion $P \in \mathcal{L}(H_1)$. Ist $v \in \operatorname{im} P$, dann ist $Pv = v$ und damit $S^* S v = v$ für alle $v \in \operatorname{im} P = (\ker P)^\perp = \operatorname{im} S^*$. Somit ist $\operatorname{im} S^*$ abgeschlossen und dann $(\ker S)^\perp = \operatorname{im} S^* = (\ker P)^\perp$ nach Satz 14.2.ii) und Lemma 14.3. Dadurch erklärt sich die Bezeichnung "partielle Isometrie": Die Einschränkung von S auf $(\ker S)^\perp$ ist eine Isometrie. Der Untervektorraum $(\ker S)^\perp$ ist abgeschlossen in H, damit vollständig. Wegen $\|Sv\| = \|v\|$ für alle $v \in (\ker S)^\perp$ ist dann auch $\operatorname{im}(S)$ vollständig, somit abgeschlossen. Damit ist $S : (\ker S)^\perp \to \operatorname{im} S = \overline{\operatorname{im} S} = (\ker S^*)^\perp$ surjektiv, also (in diesen Einschränkungen) unitär, und S^* ist die inverse Abbildung $S^* : (\ker S^*)^\perp \to \operatorname{im} S^* = (\ker S)^\perp$. Somit ist auch $SS^* = \mathrm{id}_{(\ker S^*)^\perp}$, d.h. auch SS^* ist eine partielle Isometrie.

Wir diskutieren nun selbstadjungierte Operatoren. Unser Ziel ist zu zeigen:

Satz 14.11 *In einem komplexen Hilbert-Raum H ist $A \in \mathcal{L}(H)$ genau dann selbstadjungiert, wenn $\langle v, Av\rangle$ reell ist für alle $v \in H$.*

Beweis. $(\Rightarrow)$ $\langle v, Av\rangle = \langle Av, v\rangle = \overline{\langle v, Av\rangle}$.

$(\Leftarrow)$ $\langle v, Av\rangle = \overline{\langle v, Av\rangle} = \langle Av, v\rangle = \langle v, A^* v\rangle$, also $\langle v, (A - A^*)v\rangle = 0$ für alle $v \in H$. Die Behauptung folgt dann aus den folgenden beiden Lemmata von eigenem Interesse. $\square$

Lemma 14.12 *Es sei H ein Hilbert-Raum, $A \in \mathcal{L}(H_1, H_2)$ und $v \in H_1$ sowie $w \in H_2$. Dann gilt*

i) $|\langle w, Av\rangle| \leq \|A\|\|w\|\|v\|$.

ii) *Ist $|\langle w, Av\rangle| \leq c\|w\|\|v\|$ für alle $v \in H_1$ und $w \in H_2$, dann ist $\|A\| \leq c$.*

iii) $\|A\| = \sup\{|\langle w, Av\rangle| \ : \ v \in B_{H_1}\ , \ w \in B_{H_2}\}$.

Beweis. i) folgt aus Cauchy-Schwarz und der Definition der Operatornorm.

ii) Es gilt $\|Av\|^2 = \langle Av, Av\rangle \leq c\|Av\|\|v\| \leq c\|A\|\|v\|^2$ und damit $\|A\| := \sup\limits_{v\in B_{H_1}} \|Av\| \leq \sqrt{c\|A\|}$.

iii) Insbesondere gilt ii) auch für $\|v\| = 1 = \|w\|$, d.h. ist $|\langle v, Aw\rangle| \leq c$ für alle $v \in B_{H_1}$ und $w \in B_{H_2}$, dann ist $\|A\|_{op} \leq c$. Das ist erfüllt für $c := \sup\{|\langle w, Av\rangle| \ : \ v \in B_{H_1}\ , \ w \in B_{H_2}\}$. Die Umkehrung $c \leq \|A\|$ folgt aus i). □

Lemma 14.13 *Gilt in einem komplexen Hilbert-Raum $\langle v, Av\rangle = 0$ für alle $v \in H$, so folgt $A = 0$.*

Beweis. Wir zeigen, daß dann auch $\langle v, Aw\rangle = 0$ für alle $v, w \in H$ gilt. Nach Lemma 14.12 ist dann $\|A\| = 0$, also $A = 0$. Zunächst gilt

$$0 = \langle v + w, A(v + w)\rangle - \langle v - w, A(v - w)\rangle = 2\langle v, Aw\rangle + 2\langle w, Av\rangle \ .$$

Ersetzung von $v \mapsto iv$ liefert $0 = -2i\langle v, Aw\rangle + 2i\langle w, Av\rangle$ für alle $v, w \in H$ und damit $A = 0$. □

Damit ist Satz 14.11 bewiesen. Lemma 14.13 wäre falsch in reellen Hilbert-Räumen! Für selbstadjungierte Operatoren vereinfacht sich Lemma 14.12 wie folgt:

Satz 14.14 *Ist $A \in \mathcal{L}(H)$ selbstadjungiert, dann gilt $\|A\| = \sup\limits_{v\in B_H} |\langle v, Av\rangle|$.*

Beweis. Es sei $a := \sup_{v\in B_H} |\langle v, Av\rangle|$. Dann gilt $|\langle v, Av\rangle| \leq a\|v\|^2$ für alle $v \in H$. Nach Lemma 14.12 ist $a \leq \|A\|$. Wegen

$$\langle (v + w), A(v + w)\rangle - \langle (v - w), A(v - w)\rangle = 2\langle v, Aw\rangle + 2\langle w, Av\rangle = 4\mathrm{Re}(\langle w, Av\rangle)$$

gilt mit der Parallelogrammgleichung für alle $v, w \in H$

$$\begin{aligned} 4\mathrm{Re}|\langle w, Av\rangle| &\leq |\langle (v + w), A(v + w)\rangle| + |\langle (v - w), A(v - w)\rangle| \\ &\leq a(\|v + w\|^2 + \|v - w\|^2) = 2a(\|v\|^2 + \|w\|^2) \end{aligned}$$

und damit $\sup_{v,w\in B_H} \mathrm{Re}|\langle w, Av\rangle| \leq a$. Für $\mathbb{K} = \mathbb{R}$ ist das wegen Lemma 14.12.iii) die Behauptung. Für $\mathbb{K} = \mathbb{C}$ können wir durch Drehung $w \mapsto e^{i\alpha}w$ stets $\langle w, Av\rangle \in \mathbb{R}$ erreichen, so daß die Behauptung wieder aus Lemma 14.12.iii) folgt. □

Definition 14.15 Ein Operator $A \in \mathcal{L}(H)$ heißt *positiv*, geschrieben $A \geq 0$, falls $\langle v, Av\rangle \geq 0$ für alle $v \in H$. Für selbstadjungierte Operatoren $A, B \in \mathcal{L}(H)$ schreiben wir $A \geq B$, falls $A - B \geq 0$.

Insbesondere gilt $A = B \quad \Leftrightarrow \quad (A \geq B$ und $B \geq A)$. Aus der Definition folgt, daß in einem komplexen Hilbert-Raum jeder positive Operator auch selbstadjungiert ist. Sind A, B positiv und $\alpha, \beta \in \mathbb{R}_+$, so ist auch $\alpha A + \beta B \geq 0$. Für alle $B \in \mathcal{L}(H)$ ist $A = B^*B$ positiv, denn $\langle v, B^*Bv\rangle = \langle Bv, Bv\rangle \geq 0$ für alle $v \in H$. Insbesondere ist jeder Projektor positiv wegen $P = P^*P$. Die Ordnungsrelation $\geq$ aus Definition 14.15 definiert damit eine partielle Ordnung auf der Menge der Projektoren, die viele wichtige Informationen enthält. Zunächst charakterisieren wir Projektoren wie folgt:

Lemma 14.16 *Durch $X = \operatorname{im} P$ wird eine $1:1$-Korrespondenz zwischen abgeschlossenen Untervektorräumen $X \subset H$ und Projektoren $P = P^* = P^2 \in \mathcal{L}(H)$ erklärt.*

Beweis. Ist P Projektor, dann ist $X = \operatorname{im} P$ abgeschlossener Untervektorraum nach Lemma 14.9.ii). Sei umgekehrt $X \subset H$ abgeschlossener Untervektorraum, dann gibt es nach Lemma 12.11 eine eindeutige Zerlegung $H = X \oplus X^\perp$, d.h. jeder Vektor $v \in H$ besitzt eine eindeutige Zerlegung $v = v_1 + v_2$ mit $v_1 \in X$ und $v_2 \in X^\perp$. Definieren wir $Pv := v_1$, dann gilt $P^2 = P$. Mit der analogen Zerlegung $w = w_1 + w_2$ folgt $\langle w, Pv\rangle = \langle w_1 + w_2, v_1\rangle = \langle w_1, v_1\rangle = \langle w_1, v_1 + v_2\rangle = \langle Pw, v\rangle$. Damit ist P selbstadjungiert. □

Satz 14.17 *Seien $P, Q \in \mathcal{L}(H)$ Projektoren und X_P, X_Q die zugehörigen abgeschlossenen Untervektorräume von H nach Lemma 14.16. Dann sind äquivalent:*

i) $X_P \subset X_Q$

ii) $QP = P = PQ$

iii) $\|Pv\| \leq \|Qv\|$ *für alle* $v \in H$

iv) $P \leq Q$

In diesem Fall gilt: Auch $Q - P$ ist Projektor, und zwar auf den Orthogonalraum zu $P(H)$ in $Q(H)$.

Beweis. i)⇒ii) Für alle $v \in H$ ist $Pv \in X_P \subset X_Q$ und damit $QPv = Pv$, also $QP = P$. Es folgt $P = P^* = (QP)^* = P^*Q^* = PQ$.

ii)⇒iii) $\|Pv\| = \|PQv\| \leq \|P\|\|Qv\| \leq \|Qv\|$.

iii)⇒iv) $\langle v, (Q-P)v\rangle = \langle v, Qv\rangle - \langle v, Pv\rangle = \langle Qv, Qv\rangle - \langle Pv, Pv\rangle = \|Qv\|^2 - \|Pv\|^2 \geq 0$.

iv)⇒i) Für $w \subset X_P$ gilt $\|w\|^2 = \langle w, Pw\rangle \leq \langle w, Qw\rangle \leq \|Qw\|^2 \leq \|w\|^2$, d.h. sogar $\|Qw\|^2 = \|w\|^2$. Da $(\mathbf{1}_H - Q)w$ und Qw orthogonal sind, gilt $\|w\|^2 =$

$\|((\mathbf{1}_H - Q) + Q)w\|^2 = \|(\mathbf{1}_H - Q)w\|^2 + \|Qw\|^2$, also $\|(\mathbf{1}_H - Q)w\|^2 = 0$ und damit $Qw = w$ und $X_P \subset X_Q$.

Auch $Q - P$ ist selbstadjungiert, und aus ii) folgt $(Q - P)^2 = Q^2 - QP - QP + P^2 = Q - P$. □

Für $\alpha \in \mathbb{R}$ ist auch $\alpha\,\mathbf{1}_H$ selbstadjungiert, so daß sich selbstadjungierte Operatoren und reelle Zahlen vergleichen lassen. Dabei gilt $-\|A\| \leq A \leq \|A\|$. Wir können dann monotone und beschränkte Netze selbstadjungierter Operatoren betrachten:

Satz 14.18 *Sei $(A_\lambda)_{\lambda\in\Lambda}$ ein monotones und beschränktes Netz selbstadjungierter Operatoren $A_\lambda \in \mathcal{L}(H)$, d.h. es gilt (für monoton wachsend) $A_\lambda \leq A_\mu$ für alle $\lambda, \mu \in \Lambda$ mit $\lambda \leq \mu$ sowie $A_\lambda \leq B$ für alle $\lambda \in \Lambda$. Dann ist $(A_\lambda)_{\lambda\in\Lambda}$ in der schwachen Operatortopologie konvergent gegen einen selbstadjungierten Operator $A \in \mathcal{L}(H)$, d.h. für alle $v, w \in H$ gilt $\langle w, A_\lambda v\rangle \to \langle w, Av\rangle$.*

Beweis (für monoton wachsend). Für $v \in H$ ist das Netz $(\langle v, A_\lambda v\rangle)_{\lambda\in\Lambda}$ beschränkt in $\mathbb{R}$, besitzt also ein Supremum s. Somit gibt es zu $\epsilon > 0$ ein $\mu \in \Lambda$ mit $s - \epsilon \leq \langle v, A_\mu v\rangle \leq s$. Aus der Monotonie des Netzes folgt: $s - \epsilon \leq \langle v, A_\lambda c\rangle \leq s$ für alle $\lambda \geq \mu$, d.h. $(\langle v, A_\lambda v\rangle)_{\lambda\in\Lambda}$ ist konvergent gegen s. Aus den Polarisationsformeln folgt, daß für alle $v, w \in H$ das Netz $(\langle w, A_\lambda v\rangle)_{\lambda\in\Lambda}$ in $\mathbb{K}$ konvergiert. Da die schwache Operator-Topologie eine Fréchet-Topologie ist, wird durch $\langle w, A_\lambda v\rangle \to \langle w, Av\rangle$ ein Operator $A : H \to H$ erklärt. Ähnlich wie beim Rieszschen Darstellungssatz beweist man die Linearität von A, und die Beschränktheit ist Teil der Voraussetzungen. Analog zur Linearität folgt $A = A^*$. □

Die Aussage gilt auch in der starken Operatortopologie, benötigt dann aber Beweistechniken, die uns noch nicht zur Verfügung stehen.

Wir wenden diesen Satz auf Projektoren an:

Satz 14.19 *Jedes monoton wachsendes Netz $(P_\lambda)_{\lambda\in\Lambda}$ von Projektoren ist konvergent in der schwachen Operatortopologie, und der Limes P ist wieder Projektor. Es gilt* $\ker P = \bigcap_{\lambda\in\Lambda} \ker P_\lambda$.

Beweis. Für jeden Projektor P_λ gilt $P_\lambda \leq \mathbf{1}_H$. Nach Satz 14.18 konvergiert das Netz $(P_\lambda)_{\lambda\in\Lambda}$ in der schwachen Operatortopologie gegen einen selbstadjungierten positiven und durch $\mathbf{1}_H$ beschränkten Operator $P \in \mathcal{L}(H)$. Das bedeutet insbesondere: Für alle $\epsilon > 0$ und $v, w \in H$ gibt es $\mu \in \Lambda$, so daß die drei Abschätzungen

$$|\langle w, (P - P_\lambda)v\rangle| < \epsilon\,, \quad |\langle Pw, (P - P_\lambda)v\rangle| < \epsilon\,, \quad |\langle (P - P_\lambda)w, P_\lambda v\rangle| < \epsilon$$

gelten für alle $\lambda \geq \mu$. Es folgt

$$|\langle w, (P^2 - P)v\rangle| = |\langle w, P(P - P_\lambda) + (P - P_\lambda)P_\lambda + (P_\lambda - P))v\rangle| < 3\epsilon$$

für alle $v, w \in H$ und $\epsilon > 0$. Damit gilt $P^2 = P$.

Ist $v \in \ker P$, so auch $v \in \ker P_\lambda$ für alle λ wegen $P_\lambda \leq P$. Ist umgekehrt $v \in \ker P_\lambda$ für alle λ, dann ist $\langle Pv, Pv\rangle = \lim_\lambda \langle Pv, P_\lambda v\rangle = 0$ und damit $v \in \ker P$. □

Der Satz spielt eine wichtige Rolle bei der Konstruktion von *Spektralscharen*.

15 Das Spektrum

Die Spektraltheorie verallgemeinert die Theorie der Eigenwertprobleme in endlich-dimensionalen Vektorräumen. Ergebnisse der Spektraltheorie sind von fundamentaler Bedeutung in der Theorie der Operatoralgebren. Uns interessieren vor allem Operatoren auf Hilbert-Räumen. Manche Aussagen gelten sogar allgemeiner auf Banach-Räumen.

Definition 15.1 Sei X ein Banach-Raum und $A \in \mathcal{L}(X) := \mathcal{L}(X, X)$. Eine komplexe Zahl $\lambda \in \mathbb{C}$ heißt *Spektralwert* von A, wenn die lineare stetige Abbildung $A - \lambda \mathbf{1}_X : X \to X$ nicht bijektiv ist. Die Menge aller Spektralwerte von A heißt das *Spektrum von* A, bezeichnet mit $\mathrm{sp}(A)$. Dabei heißt λ *Eigenwert*, falls $\ker(A - \lambda \mathbf{1}_X) \neq \{0\}$, und ein Vektor $0 \neq v \in \ker(A - \lambda \mathbf{1}_X)$ heißt Eigenvektor.

Das Komplement $R(A) := \mathbb{C} \setminus \mathrm{sp}(A)$ ist die *Resolventenmenge* von A, und Werte $\mu \in R(A)$ heißen *reguläre Werte von* A. Die Funktion $R(A) \ni \mu \mapsto R_\mu(A) := (A - \mu \mathbf{1}_X) \in \mathcal{L}(X)$ heißt *Resolventenfunktion*.

Nach dem Satz vom inversen Operator (Folgerung 6.12) ist $\mu \in R(A)$ genau dann, wenn $A - \mu \mathbf{1}_X$ beideitig beschränkt invertierbar ist.

Jeder Eigenwert ist Spektralwert, insbesondere auch die Eigenwerte quadratischer Matrizen. Es gibt aber auch andere Arten von Spektralwerten:

Beispiel 15.2

i) Sei (Y, μ) ein Maßraum und $X = L^p(Y, \mu)$. Dann definiert jede stetige beschränkte Funktion $f \in C_b(Y)$ durch Multiplikation einen linearen beschränkten Operator $f : X \to X$. Das Spektrum von f besteht dann aus den Funktionswerten: Denn ist $f(y) = \lambda$ für ein $y \in Y$, so gibt es zu $f - \lambda 1$ keine inverse Funktion. Aber λ ist (außer für f konstant) kein Eigenwert.

ii) Sei H unendlich-dimensionaler separabler Hilbert-Raum mit ONB $\{v_n\}$ und $Sv_n := v_{n+1}$ der Shift-Operator. Dann ist 0 Spektralwert, da S nicht surjektiv ist. Dagegen hat S keinen Eigenwert. ◁

Bemerkung 15.3 Man unterscheidet für $A \in \mathcal{L}(X)$ folgende Spektraltypen:

i) $\mathrm{sp}_p(A) = \{\lambda \in \mathbb{K} \,:\, A - \lambda \mathbf{1}_X \text{ nicht injektiv }\}$ heißt *Punktspektrum*.

ii) $\mathrm{sp}_c(A) = \{\lambda \in \mathbb{K} \,:\, A - \lambda \mathbf{1}_X \text{ injektiv, nicht surjektiv, aber mit dichtem Bild }\}$ heißt *kontinuierliches Spektrum*.

iii) $\operatorname{sp}_r(A) = \{\lambda \in \mathbb{K} \,:\, A - \lambda \mathbf{1}_X$ injektiv, ohne dichtes Bild $\}$ heißt *Residualspektrum.*

Lemma 15.4 *Seien X, Y Banach-Räume, $A \in \mathcal{L}(X,Y)$ bijektiv und $B \in \mathcal{L}(X,Y)$ mit $\|A-B\| < \|A^{-1}\|^{-1}$. Dann ist auch B bijektiv.*

Beweis. Nach dem Satz vom inversen Operator ist A^{-1} stetig. Wir betrachten die Abbildung $C := \sum_{n=0}^{\infty} (A^{-1}(A-B))^n A^{-1} : Y \to X$ mit $(A^{-1}(A-B))^0 = \mathbf{1}_X$. Die Reihe konvergiert nach dem Cauchy-Kriterium und der Vollständigkeit von X, denn sei $\|A-B\| \le q\|A^{-1}\|^{-1}$ mit $0 \le q < 1$, so gilt

$$\begin{aligned}
\Big\| \sum_{n=N}^{M} (A^{-1}(A-B))^n A^{-1} \Big\| &\le \sum_{n=N}^{M} \|(A^{-1}(A-B))^n A^{-1}\| \\
&\le \sum_{n=N}^{M} (\|A^{-1}\|\|A-B\|)^n \|A^{-1}\| \le q^N \frac{1-q^{M-N+1}}{1-q} \|A^{-1}\| \to 0\,.
\end{aligned}$$

Weiter gilt

$$\begin{aligned}
B\sum_{n=0}^{N} (A^{-1}(A-B))^n A^{-1} &= \sum_{n=0}^{N} AA^{-1}(B-A+A)(A^{-1}(A-B))^n A^{-1}\,, \\
&= A\Big(\sum_{n=0}^{N} (A^{-1}(A-B))^n A^{-1} - \sum_{n=1}^{N+1} (A^{-1}(A-B))^n A^{-1} \Big) \\
&= \mathbf{1}_X - A(A^{-1}(A-B))^{N+1} A^{-1} \\
\Big(\sum_{n=0}^{N} (A^{-1}(A-B))^n A^{-1} \Big) B &= \sum_{n=0}^{N} (A^{-1}(A-B))^n A^{-1}(B-A+A)A^{-1}A \\
&= \Big(\sum_{n=0}^{N} (A^{-1}(A-B))^n A^{-1} - \sum_{n=1}^{N+1} (A^{-1}(A-B))^n A^{-1} \Big) A \\
&= \mathbf{1}_X - (A^{-1}(A-B))^{N+1}\,.
\end{aligned}$$

Für $N \to \infty$ folgt $BC = \mathbf{1}_X = CB$. □

Folgerung 15.5 *Für $A \in \mathcal{L}(X)$ ist die Resolventenmenge $R(A)$ offen, das Spektrum $\operatorname{sp}(A)$ abgeschlossen.*

Bweis. Ist $A - \mu_1 \mathbf{1}_X$ bijektiv, also beidseitig beschränkt invertierbar, und ist $|\mu_2 - \mu_1|$ genügend klein, so ist nach Lemma 15.4 auch $A - \mu_2 \mathbf{1}_X$ beidseitig beschränkt invertierbar. □

Satz 15.6 *Sei X komplexer Banach-Raum und $A \in \mathcal{L}(X)$. Dann ist $\operatorname{sp}(A)$ kompakt, nicht leer, und enthalten in der Kreisscheibe in $\mathbb{C}$ mit Mittelpunkt 0 und Radius $\|A\|$.*

Beweis. Sei $|\lambda| > \|A\|$. Wegen $\lambda \neq 0$ ist $B := -\lambda \mathbf{1}_X$ bijektiv mit $\|B^{-1}\| = |\lambda|^{-1}$. Damit gilt $\|(A - \lambda\mathbf{1}_X) - B\| = \|A\| < |\lambda| = \|B^{-1}\|^{-1}$, und nach Lemma 15.4 ist $A - \lambda\mathbf{1}_X$ beschränkt invertierbar, also λ regulärer Wert. Somit ist $\mathrm{sp}(A)$ beschränkt (durch $\|A\|$), außerdem abgeschlossen nach Folgerung 15.5, d.h. kompakt.

Ist $\zeta \in \mathbb{C}$ regulärer Wert und $z \in \mathbb{C}$ mit $|z - \zeta| < \|(A - \zeta\mathbf{1}_X)^{-1}\|^{-1}$. Dann ist $A - z\mathbf{1}_X$ beschränkt invertierbar nach Lemma 15.4, und es gilt $R_z(A) = (A - z\mathbf{1}_X)^{-1} = \sum_{n=0}^{\infty}(R_\zeta(A)(z-\zeta))^n R_\zeta(A) = \sum_{n=0}^{\infty}(R_\zeta(A))^{n+1}(z-\zeta)^n$. Somit ist die Resolventenfunktion $\zeta \mapsto R_\zeta(A)$ in jedem Punkt $\zeta \in R(A)$ holomorph (in eine Potenzreihe entwickelbar).

Angenommen, $\mathrm{sp}(A)$ wäre leer, die Resolventenfunktion somit eine ganze Funktion (auf ganz $\mathbb{C}$ holomorph). Außerhalb der kompakten Kreisscheibe K mit Mittelpunkt 0 und Radius $2\|A\|$ ist $|R_\zeta(A)| < \frac{1}{\|A\|}$, und auf K ist $R_\zeta(A)$ beschränkt als stetige Funktion. Somit ist $R_\zeta(A)$ eine beschränkte ganze Funktion, nach dem Satz von Liouville (der für Potenzreihen mit Koeffizienten aus $\mathcal{L}(X)$ genauso gilt wie in $\mathbb{C}$) dann konstant, und wegen $\lim_{\zeta\to\infty} R_\zeta(A) = 0$ identisch 0. Somit ist auch $R_0(A) = A^{-1} = 0$, Widerspruch. □

Nach dieser allgemeinen Theorie des Spektrums von Operatoren auf Banach-Räumen steuern wir nun auf die besondere Situation für normale Operatoren auf einem Hilbert-Raum zu. Es stellt sich heraus, daß Spektralwerte “nahezu” Eigenwerte sind. Als Vorbereitung dieses Ergebnisses benötigen wir Aussagen zum inversen Operator, die wir allgemein in normierten Vektorräumen zeigen.

Lemma 15.7 *Seien X, Y normierte Vektorräume, $S : X \to Y$ linear und stetig und $T : Y \to X$ linear mit $ST = \mathbf{1}_Y$. Es sei $a := \inf\{\|Sx\| \ : \ x \in \mathrm{im}(T)\ ,\ \|x\| = 1\}$. Dann gilt: T ist genau dann stetig, wenn $a > 0$, und in diesem Fall gilt $\|T\| = a^{-1}$.*

Beweis. Wegen $ST = \mathbf{1}_Y$ ist T injektiv, und betrachtet als $T : Y \to \mathrm{im}(T)$ bijektiv mit Inversem $S\big|_{\mathrm{im}(T)}$. Es gilt:

$$\begin{aligned}
\|T\| &= \sup\Big\{\frac{\|Ty\|}{\|y\|} \ : \ y \in Y \setminus \{0\}\Big\} = \sup\Big\{\frac{\|Ty\|}{\|STy\|} \ : \ y \in Y \setminus \{0\}\Big\} \\
&= \sup\Big\{\frac{\|x\|}{\|Sx\|} \ : \ x \in \mathrm{im}(T) \setminus \{0\}\Big\} \\
&= \sup\big\{(\|Sx\|)^{-1} \ : \ x \in \mathrm{im}(T)\ ,\ \|x\| = 1\big\} = a^{-1}\ .
\end{aligned}$$

Insbesondere ist T unbeschränkt für $a = 0$. □

Ist S bijektiv, so kann T als Inverses von S angenommen werden, und die Stetigkeits-Bedingung reduziert sich auf $a := \inf_{x \in B_X} \|Sx\| > 0$ und dann

$\|T\| = a^{-1}$. Hier ist zu beachten, daß der Satz vom inversen Operator (Folgerung 6.12) nicht gilt, wenn X, Y nicht vollständig sind.

Lemma 15.8 *Seien H Hilbert-Raum, $A \in \mathcal{L}(H)$ normal und $a := \inf_{v\in B_H} \|Av\| > 0$. Dann ist A beidseitig beschränkt invertierbar mit $\|A^{-1}\| = a^{-1}$.*

Beweis. Für $a > 0$ ist $\ker A = \{0\}$, d.h. $A : H \to \mathrm{im}(A)$ ist bijektiv mit Inversem $T : \mathrm{im}(A) \to H$, mit $\|T\| = a^{-1}$. Ist (v_n) Cauchy-Folge in $\mathrm{im}(A)$, so ist (Tv_n) Cauchy-Folge in H, somit konvergent gegen ein $w = Tv \in H$ mit $Aw = v$. Damit ist $\mathrm{im}(A)$ vollständig, insbesondere abgeschlossen. Angenommen, es gibt ein $0 \neq v \in (\mathrm{im}(A))^\perp$ mit $\|v\| = 1$. Dann gilt

$$0 = \langle v, AA^*v\rangle = \langle v, A^*Av\rangle = \|Av\|^2 \geq a^2 > 0\;,$$

Widerspruch. Somit ist $\mathrm{im}(A) = H$, und A ist auf ganz H beidseitig beschränkt invertierbar. □

Satz 15.9 *Sei H Hilbert-Raum und $A \in \mathcal{L}(H)$ normal. Eine Zahl $\lambda \in \mathbb{C}$ ist genau dann Spektralwert, wenn es eine Folge (v_n) von Einheitsvektoren in H gibt mit $\|(A - \lambda\mathbf{1}_H)v_n\| \to 0$ für $n \to \infty$.*

Beweis Mit A ist auch $A - \lambda\mathbf{1}_H$ normal. Aus Lemma 15.8 folgt:

$$\begin{aligned}
\lambda \text{ Spektralwert} \quad &\Leftrightarrow \quad A - \lambda\mathbf{1}_H \text{ nicht beschränkt invertierbar}\\
&\Leftrightarrow \quad a := \inf_{v\in B_H} \|(A - \lambda\mathbf{1}_H)v\| = 0\\
&\Leftrightarrow \quad \text{es gibt } (v_n) \text{ mit den genannten Eigenschaften.}
\end{aligned}$$

□

Theorem 15.10 *Sei H Hilbert-Raum und $A, P, U \in \mathcal{L}(H)$.*

i) *Ist $A = A^*$ selbstadjungiert, so ist $\mathrm{sp}(A)$ eine kompakte Teilmenge des reellen Intervalls $[-\|A\|, \|A\|]$.*

ii) *Ist $A \geq 0$ positiv, so ist $\mathrm{sp}(A)$ eine kompakte Teilmenge des reellen Intervalls $[0, \|A\|]$.*

iii) *Ist $P = P^* = P^2$ Projektor, so besteht $\mathrm{sp}(P) \subset \{0, 1\}$ höchstens aus den Zahlen 0 und 1.*

iv) *Ist U unitär, so ist $\mathrm{sp}(U)$ eine kompakte Teilmenge des Einheitskreises $S^1 \subset \mathbb{C}$.*

Ferner gilt (allgemein in $\mathcal{L}(X)$): $\mathrm{sp}(A^*) = \{\bar\lambda \;:\; \lambda \in \mathrm{sp}(A)\}$. Ist A invertierbar, dann gilt $\mathrm{sp}(A^{-1}) = \{\lambda^{-1} \;:\; \lambda \in \mathrm{sp}(A)\}$.

Beweis. In i)–iv) handelt es sich um normale Operatoren, so daß Satz 15.9 anwendbar ist. Die Kompakatheit und Beschränktheit durch $\|A\|$ war in Satz 15.6 gezeigt.

i) Ist A selbstadjungiert und $\lambda \in \mathrm{sp}(A)$, dann gibt es eine Folge v_n von Einheitsvektoren in H mit $\|(A - \lambda \mathbf{1}_H)v_n\| \to 0$ für $n \to \infty$. Dann ist aber auch $\langle v_n, (A - \lambda \mathbf{1}_H)v_n\rangle = \langle v_n, Av_n\rangle - \langle v_n, \lambda v_n\rangle \to 0$, d.h. $(\langle v_n, Av_n\rangle)$ konvergiert gegen λ. Aber $\langle v_n, Av_n\rangle$ ist für alle n reell, somit ist $\lambda \in \mathbb{R}$.

ii) Analog wie i), zusätzlich ist $\langle v_n, Av_n\rangle \geq 0$ und damit $\lambda \geq 0$.

iii) Es gilt $\|P(P-\lambda\mathbf{1}_H)v_n\| = \|(1-\lambda)Pv_n\| = \|(1-\lambda)(P-\lambda\mathbf{1}_H+\lambda\mathbf{1}_H)v_n\| \to 0$ für $n \to \infty$, somit $\|(1-\lambda)\lambda v_n\| \to 0$. Wegen $\|v_n\| = 1$ für alle n muß $\lambda(1-\lambda) = 0$ gelten.

iv) Es gilt $1 = \langle v_n, v_n\rangle = \langle Uv_n, Uv_n\rangle$. Die rechte Seite konvergiert gegen $|\lambda|^2\langle v_n, v_n\rangle$, somit $|\lambda|^2 = 1$.

Ferner ist klar: $A-\lambda\mathbf{1}_X$ ist genau dann beschränkt invertierbar, wenn $A^*-\bar{\lambda}\mathbf{1}_X$ beschränkt invertierbar ist.

Sei A invertierbar (dann ist 0 regulärer Wert) und $\lambda \neq 0$. Dann gilt $A^{-1} - \lambda^{-1}\mathbf{1}_X = (\lambda\mathbf{1}_X - A)(A^{-1}\lambda^{-1})$. Da $A^{-1}\lambda^{-1}$ bijektiv ist, gilt: $\lambda^{-1} \in \mathrm{sp}(A^{-1}) \Leftrightarrow \lambda \in \mathrm{sp}(A)$. □

Bemerkung 15.11 Es gelten alle Umkehrungen von Theorem 15.10.i)-iv) in folgendem Sinn: ist $A \in \mathcal{L}(H)$ mit $\mathrm{sp}(A) \subset \mathbb{R}$ (und H komplexer Hilbert-Raum!), so ist $A = A^*$. Analog für ii)–iv). Diese Umkehrungen ergeben sich aus dem *Spektralsatz für normale Operatoren.*

Wir geben nun stetigen Funktionen $f(A)$ selbstadjungierter Operatoren einen Sinn. Eine besonders wichtige Funktion wird die Quadratwurzel positiver Operatoren sein, so daß wir den Betrag eines Operators erklären können als $|A| = \sqrt{A^*A}$.

Satz 15.12 *Sei X komplexer Banach-Raum, $A \in \mathcal{L}(X)$. Jedes Polynom $p = p(t) = \sum_{k=0}^n c_k t^k$ mit $c_k \in \mathbb{C}$ einer Variablen t defininiert durch $p(A) = \sum_{k=0}^n c_k A^k$, mit $A^0 := \mathbf{1}_X$, einen linearen stetigen Operator $p(A) \in \mathcal{L}(X)$. Es gilt* $\mathrm{sp}(p(A)) = \{p(\lambda) \ :\ \lambda \in \mathrm{sp}(A)\} =: p(\mathrm{sp}(A))$.

Beweis. $p(A) \in \mathcal{L}(X)$ ist klar. Sei $\lambda \in \mathrm{sp}(A)$. Dann gilt

$$p(A) - p(\lambda\mathbf{1}_X) = \sum_{k=1}^n c_k(A^k - \lambda^k\mathbf{1}_X) .$$

Für $k \geq 1$ gilt $(A^k-\lambda^k\mathbf{1}_X) = \Big(\sum_{j=0}^{k-1} A^j\lambda^{k-1-j}\Big)(A-\lambda\mathbf{1}_X)$, so daß $p(A)-p(\lambda\mathbf{1}_X) = B(A-\lambda\mathbf{1}_X)$ für ein $B \in \mathcal{L}(X)$. Nach Voraussetzung ist $(A-\lambda\mathbf{1}_X)$ nicht bijektiv, damit ist auch $p(A) - p(\lambda\mathbf{1}_X)$ nicht bijektiv, somit $p(\lambda)$ Spektralwert von $p(A)$.

Sei umgekehrt $\mu \in \mathrm{sp}(p(A))$. Die Gleichung $p(\lambda) - \mu = 0$ hat (nach Fundamentalsatz der Algebra) n mit Vielfachheit gezählte Lösungen $\lambda_1, \ldots, \lambda_n$. Da komplexe Polynome in Linearfaktoren zerfallen, gilt

$$p(A) - \mu\mathbf{1}_X = c_n(A - \lambda_1\mathbf{1}_X)(A - \lambda_2\mathbf{1}_X)\cdots(A - \lambda_n\mathbf{1}_X) .$$

Da μ Spektralwert von $p(A)$ ist, können nicht alle $A-\lambda_k \mathbf{1}_X$ bijekiv sein. Also gilt $\lambda_j \in \mathrm{sp}(A)$ für mindestens ein $j \in \{1,\dots,n\}$. Aber $\mu = p(\lambda_j)$, d.h. $\mathrm{sp}(p(A)) \subset p(\mathrm{sp}(A))$. □

Folgerung 15.13 *Ist $A = A^* \in \mathcal{L}(H)$ selbstadjungiert, so ist zumindest eine der Zahlen $\|A\|$ oder $-\|A\|$ Spektralwert.*

Beweis. Das ist klar für $A = 0$. Ansonsten können wir durch Skalieren mit $\|A\|^{-1}$ annehmen: $A = A^* \in \mathcal{L}(H)$ mit $\|A\| = 1$. Also gibt es eine Folge (v_n) von Einheitsvektoren in H mit $\|Av_n\| \to 1$ für $n \to \infty$. Für diese Folge gilt:

$$\begin{aligned} 0 \le \|(\mathbf{1}_H - A^2)v_n\|^2 &= \|v_n\|^2 + \|A^2 v_n\|^2 - 2\mathrm{Re}\langle v_n, A^2 v_n\rangle \\ &= 1 + \|A^2 v_n\|^2 - 2\|Av_n\| \le 2 - 2\|Av_n\| \to 0\ . \end{aligned}$$

Nach Satz 15.9 ist 1 Spektralwert von A^2, damit $1 \in (\mathrm{sp}(A))^2$ nach Satz 15.12 und schließlich $1 \in \mathrm{sp}(A)$ oder $-1 \in \mathrm{sp}(A)$. □

Diese Aussage führt (zusammen mit der C^*-Identität) auf eine rein algebraische Definition der Norm von $A \in \mathcal{L}(H)$: Es gilt

$$\|A\| = \sup\{\lambda > 0 \ :\ A^*A - \lambda^2 \mathbf{1}_H \text{ ist nicht bijektiv}\}\ .$$

Definition 15.14 Für $A \in \mathcal{L}(X)$ heißt $r(A) := \sup\{|\lambda| \ :\ \lambda \in \mathrm{sp}(A)\}$ *Spektralradius* von A.

Es gilt $r(A) \le \|A\|$, für selbstadjungierte $A = A^* \in \mathcal{L}(H)$ sogar $r(A) = \|A\|$ nach Folgerung 15.13.

Bemerkung 15.15 Über den *holomorphen Funktionalkalkül* beweist man für $A \in \mathcal{L}(H)$ die sehr wichtige *Spektralradiusformel* $r(A) = \lim_{n\to\infty} \|A^n\|^{\frac{1}{n}}$. Insbesondere existiert dieser Limes und kann dann durch $n = 2^k$ verdichtet werden. Ist A normal, so folgt

$$\begin{aligned} r(A) &= \lim_{k\to\infty} \|A^{2^k}\|^{\frac{1}{2^k}} = \lim_{k\to\infty} \|(A^*)^{2^k} A^{2^k}\|^{\frac{1}{2^{k+1}}} = \lim_{k\to\infty} \|(A^*A)^{2^k}\|^{\frac{1}{2^{k+1}}} \\ &= \|(A^*A)\|^{\frac{1}{2}} = \|A\|\ . \end{aligned}$$

Theorem 15.16 (stetiger Funktionalkalkül) *Ist H komplexer Hilbert-Raum, $A = A^* \in \mathcal{L}(H)$ selbstadjungiert und f eine stetige Funktion auf $\mathrm{sp}(A) \subset [-\|A\|, \|A\|]$, dann gibt es genau ein $f(A) \in \mathcal{L}(H)$ derart, daß die Zuordnung $C(\mathrm{sp}(A)) \ni f \mapsto f(A) \in \mathcal{L}(H)$ stetig ist und für Polynome die elementare Bedeutung aus Satz 15.12 hat.*

Beweis. $\mathrm{sp}(A)$ ist kompakter Hausdorff-Raum. Nach dem Satz von Stone-Weierstraß gibt es eine gegen f in der Supremums-Norm konvergente Folge (p_k) von Polynomen (vom Grad n_k), die somit eine Cauchy-Folge auf $\mathrm{sp}(A)$ bilden: zu

$\epsilon > 0$ gibt es ein $N \in \mathbb{N}$, so daß für alle $k, l \geq N$ gilt: $\sup_{t \in \mathrm{sp}(A)} |p_k(t) - p_l(t)| < \epsilon$. Andererseits gilt für ein beliebiges Polynom p nach Satz 15.12

$$\|p(A)\| = r(p(A)) = \max\{|\mu| \,:\, \mu \in \mathrm{sp}(p(A))\} = \max\{|p(\lambda)| \,:\, \lambda \in \mathrm{sp}(A)\} = \|p\| \,.$$

Somit gilt auch $\|p_k(A) - p_l(A)\| = \sup_{t \in \mathrm{sp}(A)} |p_k(t) - p_l(t)| < \epsilon$ für alle $k, l \geq N$, d.h. $(p_k(A))$ ist Cauchy-Folge in $\mathcal{L}(H)$, die einen eindeutigen Grenzwert $f(A) \in \mathcal{L}(H)$ besitzt. Wieder durch polynomiale Approximation folgt, daß die Abbildung $f \mapsto f(A)$ stetig ist. $\square$

Für $f, g \in \mathcal{C}(\mathrm{sp}(A))$ und $\alpha, \beta \in \mathbb{C}$ gilt:

i) $\|f(A)\| = \|f\|$

ii) $(\alpha f + \beta g)(A) = \alpha f(A) + \beta g(A)$

iii) $(fg)(A) = f(A)g(A)$

iv) $(f(A))^* = \bar{f}(A)$

v) $f(A)$ ist normal

vi) $\|f(A)\| = \|f\| = \sup_{t \in \mathrm{sp}(A)} |f(t)|$

Dabei ergibt sich vi) durch polynomiale Approximation und $\|p(A)\| = \|p\|$, die anderen Eigenschaften sind klar.

Theorem 15.16 hat fundamentale Bedeutung für die Theorie der Operatoralgebren. Wir geben eine Reihe von Anwendungen an, die z.T. benutzen, daß die Umkehrungen von Theorem 15.10 gelten (Bemerkung 15.11).

Satz 15.17 *Sei H komplexer Hilbert-Raum. Jeder selbstadjungierte Operator $A = A^* \in \mathcal{L}(H)$ besitzt eine eindeutige Zerlegung $A = A_+ - A_-$ mit $A_+, A_- \geq 0$ und $A_+A_- = A_-A_+ = 0$. Es gilt $\|A\| = \max(\|A_+\|, \|A_-\|)$. Insbesondere ist jeder lineare stetige Operator eine eindeutige Linearkombination von (höchstens) 4 positiven Operatoren.*

Beweis. Wir betrachten die stetigen Funktion $f(t) = t$, $f_+(t) = \max(t, 0)$, $f_-(t) = \max(-t, 0)$. Für diese gilt $f = f_+ - f_-$ und $f_+f_- = f_-f_+ = 0$. Dann haben $A_+ := f(A)$ und $A_- := f_-(A)$ die behaupteten Eigenschaften, und ihr Spektrum ist positiv. Angenommen, es gibt eine weitere Zerlegung $A = B - C$ mit $B, C \geq 0$ und $BC = CB = 0$. Dann folgt $A^n = B^n + (-C)^n$, somit $p(A) = p(B) + p(-C)$ für jedes Polynom p ohne konstante Komponente. Es gibt eine gegen f_+ konvergente Folge solcher Polynome, so daß gilt $f_+(A) = f_+(B) + f_+(-C)$. Aber $f_+(-C) = 0$ und $f_+(B) = B$, somit ist die Zerlegung eindeutig.

Theorem 15.18 *Sei H komplexer Hilbert-Raum. Jeder positive Operator $0 \leq A \in \mathcal{L}(H)$ hat eine eindeutige positive Quadratwurzel $0 \leq A^{\frac{1}{2}} \in \mathcal{L}(H)$, mit $A = (A^{\frac{1}{2}})^2$.*

Beweis. Wegen $\mathrm{sp}(A) \subset [0, \|A\|]$ ist $f(t) := \sqrt{t}$ stetig auf $\mathrm{sp}(A)$. Wir setzen $A^{\frac{1}{2}} := f(A)$, dann ist $f(A)$ positiv und $(f(A))^2 = A$. Angenommen, es gäbe ein weiteres $0 \leq B \in \mathcal{L}(H)$ mit $B^2 = A$. Es gibt eine Folge (p_n) von Polynomen, die auf $\mathrm{sp}(A)$ gleichmäßig gegen f konvergiert. Sei $q_n(t) := p_n(t^2)$. Wegen $\mathrm{sp}(A) = \mathrm{sp}(B^2) = \{t^2 \ : \ t \in \mathrm{sp}(B)\}$ gilt $\lim_{n\to\infty} q_n(t) = \lim_{n\to\infty} p_n(t^2) = f(t^2) = t$ gleichmäßig auf $\mathrm{sp}(B)$, und es folgt $B = \lim_{n\to\infty} q_n(B) = \lim_{n\to\infty} p_n(B^2) = \lim_{n\to\infty} p_n(A) = f(A)$. □

Definition 15.19 Sei H komplexer Hilbert-Raum und $A \in \mathcal{L}(H)$, dann heißt $|A| := (A^*A)^{\frac{1}{2}} \in \mathcal{L}(H)$ der *Betrag* von A.

Theorem 15.20 (Polarzerlegung) *Sei H komplexer Hilbert-Raum und $A \in \mathcal{L}(H)$, dann existiert eine partielle Isometrie $S \in \mathcal{L}(H)$ von $(\ker A)^\perp$ nach $(\ker A^*)^\perp$, so daß gilt $A = S|A|$.*

Beweis. Wir definieren einen Operator $T : \mathrm{im}\,|A| \to \mathrm{im}\,A$ durch $T(|A|v) := Av$ für alle $v \in H$. Dann gilt $\||A|v\|^2 = \langle v, |A|^2 v\rangle = \langle v, A^*Av\rangle = \|Av\|^2 = \|T(|A|v)\|^2$, d.h. T ist eine Isometrie. Diese setzt sich stetig fort zu $V : \overline{\mathrm{im}\,|A|} \to \overline{\mathrm{im}\,A} = (\ker A^*)^\perp$: Wähle eine gegen $w \in \overline{\mathrm{im}\,|A|}$ konvergente Folge w_n aus $\mathrm{im}\,|A|$ und definiere $Tw := \lim_{n\to\infty} Tw_n$. Die rechte Seite ist konvergent in $\overline{\mathrm{im}\,A}$. Diese stetige Fortsetzung setzen wir zu $S \in \mathcal{L}(H)$ fort durch $S\big|_{\ker A} = 0$ und $S\big|_{(\ker A)^\perp} = T$. Dabei ist $(\mathrm{im}\,|A|)^\perp = \ker|A|$, und wegen $\||A|v\|^2 = \|Av\|^2$ ist $\ker|A| = \ker A$, somit S die partielle Isometrie zu $T : (\ker A)^\perp \to (\ker A^*)^\perp$. □

Es läßt sich zeigen, daß diese Polarzerlegung eindeutig ist.

Theorem 15.21 *Sei komplexer Hilbert-Raum. Jeder Operator $A \in \mathcal{L}(H)$ ist Linearkombination von höchstens 4 unitären Operatoren.*

Beweis. Es genügt zu zeigen: jeder selbstadjungierte Operator $A = A^*$ mit $\|A\| \leq 1$ ist Linearkombination von 2 unitären. In diesem Fall ist $\mathrm{sp}(A) \subset [-1, 1]$, so daß $f(t) = t + \mathrm{i}\sqrt{1-t^2}$ auf $\mathrm{sp}(A)$ stetig ist. Es gilt $t = \frac{1}{2}(f(t) + \bar{f}(t))$ und $f(t)\bar{f}(t) = \bar{f}(t)f(t) = 1$. Setzen wir $U := f(A)$, also $U^* = \bar{f}(A)$, so folgt $A = \frac{1}{2}(U + U^*)$ und $U^*U = U^*U = \mathbf{1}_H$. □

16 Kompakte Operatoren

Viele Anwendungen führen auf sogenannte kompakte Operatoren, für die sich Eigenschaften zeigen lassen, die analog zur endlich-dimensionalen Situation sind. Das Spektrum kompakte Operatoren hat eine sehr einfache Struktur.

Definition 16.1 Seien X, Y Banach-Räume. Eine lineare Abbildung $T : X \to Y$ heißt *kompakt*, falls folgende äquivalente Bedingungen erfüllt sind:

i) $\overline{T(B_X)} \subset X$ (Norm-Abschluß) ist (Norm-)kompakt.

ii) Ist (x_n) eine beschränkte Folge aus X, so enthält (Tx_n) eine in Y (Norm-)konvergente Teilfolge.

Mit $\mathcal{K}(X,Y)$ werde die Menge der kompakten Operatoren $T : X \to Y$ bezeichnet.

In metrischen Räumen sind Kompaktheit und Folgenkompaktheit äquivalent; daraus folgt die Äquivalenz. Es gelten:

i) Jeder kompakte Operator ist stetig (da kompakte Mengen beschränkt sind).

ii) Ist X endlich-dimensional und T stetig, so ist T kompakt (da $\overline{T(B_X)}$ endlich-dimensional, beschränkt und abgeschlossen).

iii) Ist $T(X)$ endlich-dimensional und T stetig, so ist T kompakt (da $\overline{T(B_X)}$ endlich-dimensional, beschränkt und abgeschlossen).

iv) Ist X unendlich-dimensional, so ist $\mathbf{1}_X$ nicht kompakt (da B_X nicht kompakt ist).

Lemma 16.2 *Seien X, Y, Z Banach-Räume.*

i) *$\mathcal{K}(X,Y)$ ist abgeschlossener Untervektorraum von $\mathcal{L}(X,Y)$.*

ii) *Sei $S \in \mathcal{L}(Y,Z)$ und $T \in \mathcal{L}(X,Y)$. Ist S oder T kompakt, so auch $S \circ T$.*

Insbesondere gilt: $\mathcal{K}(X) := \mathcal{K}(X,X)$ ist norm-abgeschlossene Unteralgebra von $\mathcal{L}(X)$.

Beweis. i) Seien $T_1, T_2 \in \mathcal{K}(X,Y)$ und (x_n) beschränkt. Dann hat $(T_1 x_n)$ eine konvergente Teilfolge $(T_1 x_{n_k})$, und $(T_2 x_{n_k})$ hat konvergente Teilfolge $(T_2 x_{n_{kl}})$. Somit hat $((T_1 + T_2)x_n)$ die konvergente Teilfolge $((T_1 + T_2)x_{n_{kl}})$. Da λT_i trivialerweise kompakt, ist $\mathcal{K}(X,Y) \subset \mathcal{L}(X,Y)$ Untervektorraum.

Verbleibt die Abgeschlossenheit. Sei $T \in \overline{\mathcal{K}(X,Y)}$. Dann gibt es zu $\epsilon > 0$ ein $S \in \mathcal{K}(X,Y)$ mit $\|T - S\| < \epsilon$. Sei $Y_1 := S(B_X)$. Da $\overline{Y_1}$ kompakt ist, wird $\overline{Y_1}$, und damit Y_1, von endlich vielen ϵ-Umgebungen aus Y mit Mittelpunkten $\{Sx_1, \ldots, Sx_n\}$ überdeckt. Somit folgt: für jedes $x \in B_x$ gibt es ein $i \in \{1, \ldots, n\}$ mit $\|Sx - Sx_i\| < \epsilon$. Für dieses i folgt

$$\|Tx - Tx_i\| \leq \|Tx - Sx\| + \|Sx - Sx_i\| + \|Sx_i - Tx_i\| < 3\epsilon ,$$

d.h. endlich viele 3ϵ-Umgebungen überdecken $T(B_X)$ und damit $\overline{T(B_X)}$. Somit ist T kompakt.

ii) folgt direkt aus der Definition. □

Folgerung 16.3 *Ist $T \in \mathcal{L}(X,Y)$ mit $T = \lim_{n\to\infty} T_n$ (Norm-Limes) und $\dim(T_n(X)) < \infty$ für alle $n \in \mathbb{N}$, so ist T kompakt.* □

Satz 16.4 *Seien H_1, H_2 Hilbert-Räume und $T \in \mathcal{L}(H_1, H_2)$. Dann sind äquivalent:*

i) *T ist kompakt.*

ii) *T bildet schwach-konvergente Folgen in Norm-konvergente Folgen ab, d.h. ist (v_n) Folge in H_1 mit $\langle w, v_n\rangle \overset{n\to\infty}{\longrightarrow} \langle w, v\rangle$ für alle $w \in H_1$, so folgt $\|Tv_n - Tv\| \overset{n\to\infty}{\longrightarrow} 0$.*

iii) *T^*T ist kompakt.*

iv) *T^* ist kompakt.*

Beweis. i)⇒ii) Nach Folgerung 6.8 aus Banach-Steinhaus ist jede schwach-konvergente Folge beschränkt. Damit besitzt Tv_n eine in H_2 Norm-konvergente Teilfolge $(\tilde{v}_{n_k})_{k\in\mathbb{N}}$. Sei $\tilde{w}$ der Limes. Zu zeigen ist: Die gesamte Folge (Tv_n) konvergiert gegen $\tilde{w}$. Angenommen, es gibt ein $\epsilon > 0$ und eine Teilfolge $(v_{n_k})_{k\in\mathbb{N}}$ mit $\|Tv_{n_k} - \tilde{w}\| > \epsilon$ für alle k. Da v_{n_k} beschränkt ist, besitzt (Tv_{n_k}) eine konvergente Teilfolge, die wir vereinfachend als (Tv_{n_k}) selbst annehmen können. Sei w ihr Limes. Dann gilt:

$$\begin{aligned}\|w-\tilde{w}\|^2 &= \lim_{k\to\infty}\langle Tv_{n_k} - \tilde{w}, w-\tilde{w}\rangle = \lim_{k\to\infty}\langle v_{n_k}, T^*(w-\tilde{w})\rangle - \langle \tilde{w}, w-\tilde{w}\rangle \\ &= \langle v, T^*(w_1-\tilde{w})\rangle - \langle \tilde{w}, w-\tilde{w}\rangle = \lim_{k\to\infty}\langle \tilde{v}_{n_k}, T^*(w-\tilde{w})\rangle - \langle \tilde{w}, w-\tilde{w}\rangle \\ &= \lim_{k\to\infty}\langle T\tilde{v}_{n_k}, w-\tilde{w}\rangle - \langle \tilde{w}, w-\tilde{w}\rangle = 0\ ,\end{aligned}$$

Widerspruch.

ii)⇒i) Da H_1 reflexiv ist, besitzt jede in H_1 beschränkte Folge (v_n) nach Satz 4.29 eine schwach-konvergente Teilfolge v_{n_k}. Sei v der Limes. Nach Voraussetzung ist $\lim_{k\to\infty}\|Tv_{n_k} - Tv\| = 0$, also T kompakt.

i)⇒iii) und iv)⇒iii) ist Lemma 16.2.ii).

iii)⇒ii) Ist (v_n) schwach-konvergent gegen v, so ist T^*Tv_n Norm-konvergent gegen T^*Tv nach i)⇔ii). Es folgt

$$\|T(v_n - v)\|^2 = |\langle v_n - v, T^*T(v_n-v)\rangle| \leq \|v_n - v\|\ \|T^*T(v_n-v)\| \overset{n\to\infty}{\longrightarrow} 0$$

wegen $(v_n - v)$ beschränkt nach Folgerung 6.8.

i)⇒iv) Nach Lemma 16.2.ii) ist TT^* kompakt, dann T^* nach iii)⇒ii)⇒i) für T^* statt T. □

Der folgende Satz charakterisiert eine wichtige Beispielklasse von kompakten Operatoren: die Theorie der Integralgleichungen.

Satz 16.5 *Es sei (X,μ) und $K \in L^2(X\times X, \mu\times\mu)$. Dann definiert*

$$(Tf)(x) := \int_X d\mu(y)\ K(x,y) f(y)\ , \qquad f \in L^2(X,\mu)$$

einen kompakten Operator $T \in \mathcal{L}(L^2(X,\mu))$.

Beweis. Je nach Wahl der Integrationsvariable schreiben wir dx bzw. dy für $d\mu$. Zunächst existiert nach Fubini das Integral $\displaystyle\int_X dx\Big(\int_X dy\ |K(x,y)|^2\Big)$, d.h. bis auf eine Nullmenge $N \subset X$ ist $\displaystyle\int_X dy\ |K(x,y)|^2 < \infty$ für alle $x \in X \setminus N$. Damit existiert nach der Hölderschen Ungleichung das Integral $\displaystyle\int_X dy\ |K(x,y)f(y)|$ für alle $x \in X \setminus N$, und außerhalb einer Nullmenge gilt mit Cauchy-Schwarz

$$|(Tf)(x)|^2 = \Big|\int_X dy\ K(x,y)f(y)\Big|^2 \leq \int_X dy\ |K(x,y)|^2 \int_X dy\ |f(y)|^2\ .$$

Damit gilt

$$\|Tf\|_2^2 = \int_X dx\ |(Tf)(x)|^2 \leq \|K\|_2^2 \|f\|_2^2\ ,$$

d.h. $\|T\| \leq \|K\|_2$ und $T \in \mathcal{L}(L^2(X,\mu))$.

Nach Definition der Integrierbarkeit gibt es zu K eine Folge (K_n) endlicher Treppenfunktionen mit $\lim_{n\to\infty} \|K - K_n\|_2 = 0$. Nach Fubini dürfen wir $K_n(x,y) = \sum_{j=1}^{r_n} c_j \mathbf{1}_{A_j}(x)\mathbf{1}_{B_j}(y)$ annehmen für Borelmengen $A_j, B_j \subset \mathcal{B}(X)$ mit $\mu(A_j), \mu(B_j) < \infty$. Dann ist

$$\int_X dy\ K_n(x,y)f(y) = \sum_{j=1}^{r_n} c_j \mathbf{1}_{A_j}(x) \int_{B_j} dy\ f(y)$$

wieder eine endliche Treppenfunktion auf X. Der Vektorraum der endlichen Treppenfunktionen ist endlich-dimensional. Somit ist das Bild einer beliebigen beschränkten Teilmenge in $L^2(X,\mu)$ unter T_n endlich-dimensional und beschränkt, sein Abschluß damit kompakt. Damit ist T_n ein kompakter Operator. Da (K_n) in $L^2(X \times X, \mu \times \mu)$ norm-konvergent gegen K ist und die Abbildungen $K \mapsto T$ und $K_n \mapsto T_n$ stetig sind, ist auch T kompakt nach Folgerung 16.3. □

Satz 16.6 *Definiert $K(x,y)$ den kompakten Operator $T \in \mathcal{K}(L^2(X,\mu))$ entsprechend Satz 16.5, so definiert $\overline{K(y,x)}$ den adjungierten Operator T^*. Gilt also $K(x,y) = \overline{K(y,x)}$ für alle $x,y \in X$, dann ist T selbstadjungiert.*

Beweis. Für $f,g \in L^2(X,\mu)$ ist die Funktion $(x,y) \mapsto g(x)f(y)$ in $L^2(X \times X, \mu \times \mu)$. Damit ist die Funktion $(x,y) \mapsto K(x,y)f(y)g(x)$ integrierbar, und nach Fubini gilt

$$\begin{aligned}\langle g, Tf\rangle &= \int_X dx\ \overline{g(x)} \int_X dy\ K(x,y)f(y) = \int_X dy\ \Big(\int_X dx\ K(x,y)\overline{g(x)}\Big) f(y) \\ &= \int_X dx\ \overline{\Big(\int_X dy\ \overline{K(y,x)}g(y)\Big)}\ f(x) \stackrel{!}{=} \langle T^*g, f\rangle\ .\end{aligned}$$ □

Damit lassen sich Integralgleichungen der Form

$$\lambda f(x) = \int_X dy\; K(x,y)f(y)\;, \qquad K(x,y) = \overline{K(y,x)} \in L^2(X \times X, \mu \times \mu)\;,$$

welche sich z.B. aus Differentialgleichungen ergeben, auf das *Eigenwertproblem* $Tf = \lambda f$ für kompakte selbstadjungierte Operatoren zurückführen. Insbesondere wird es nicht für alle λ eine Lösung geben. Eine erste Aussage zum Spektrum ist schnell zu erhalten:

Satz 16.7 *Sei H Hilbert-Raum mit $\dim(H) = \infty$ und $T \in \mathcal{K}(H)$ kompakt. Dann ist $0 \in \mathrm{sp}(T)$, d.h. T ist nicht invertierbar.*

Beweis. Wäre $0 \notin \mathrm{sp}(T)$, so ist $T^{-1} \in \mathcal{L}(H)$ und dann $\mathbf{1}_H = TT^{-1}$ kompakt, somit H endlich-dimensional. □

Es gibt viele Beispiele von Differentialoperatoren D_x, deren Inverses ein kompakter Operator ist, z.B. ein Integraloperator. Das Inverse ist dann zu verstehen als $D_x \int_X dy\; K(x,y)f(y) = f(x)$. In diesen Fällen liefert die Spektraltheorie des kompakten Operators zu K wichtige Daten über das Spektrum des Differentialoperators.

Lemma 16.8 *Seien H_1, H_2 Hilbert-Räume, $T \in \mathcal{K}(H_1, H_2)$ und $\delta > 0$. Sei $(v_i)_{i\in I}$ ein ONS in H_1. Dann gilt: $J_\delta := \{i \in I \;:\; \|Tv_i\| \geq \delta\}$ ist endlich, somit $I_0 := \{i \in I \;:\; Tv_i \neq 0\}$ abzählbar.*

Beweis Wäre J_δ nicht endlich, so gibt es eine Folge (i_n) in J_δ mit $i_n \neq i_m$ für alle $n \neq m$. Nach Besselscher Ungleichung gilt für alle $v \in H_1$

$$\sum_{n=0}^{\infty} |\langle v, v_{i_n}\rangle|^2 \leq \|v\|^2\;.$$

Die Summe kann aber nur konvergieren, wenn $\langle v, v_{i_n}\rangle \to 0$ für $n \to \infty$, d.h. (v_{i_n}) ist schwach-konvergent gegen 0. Nach Satz 16.4.ii) ist Tv_{i_n} Norm-konvergent gegen 0, im Widerspruch zu $\|Tv_i\| \geq \delta$.

Die Aussage zu I_0 folgt aus $I_0 = \bigcup_{n=1}^{\infty} J_{\frac{1}{n}}$. □

Folgerung 16.9 *Seien H_1, H_2 Hilbert-Räume und $T \in \mathcal{K}(H_1, H_2)$. Dann ist $(\ker T)^\perp \subset H_1$ separabel.*

Beweis. Sei $(v_i)_{i\in I}$ eine ONB in H_1 und $I_0 = \{i \in I \;:\; Tv_i \neq 0\}$. Dann gilt $H_1 = \ker T \oplus (\ker T)^\perp$. Da $(v_i)_{i\in I\setminus I_0}$ ONB von $\ker T$ ist, ist I_0 ONB von $(\ker T)^\perp$, damit $(\ker T)^\perp$ separabel. □

Satz 16.10 *Für Hilbert-Räume H_1, H_2 bezeichne $\mathcal{F}(H_1, H_2) = \{T \in \mathcal{L}(H_1, H_2) \;:\; \dim(\mathrm{im}\, T) < \infty\}$ die Menge der Operatoren mit endlichem Rang. Dann gilt $\mathcal{K}(H_1, H_2) = \overline{\mathcal{F}(H_1, H_2)}$ (Norm-Abschluß).*

Beweis. $\supset$ ist Folgerung 16.3. Sei umgekehrt $T \in \mathcal{K}(H_1, H_2)$. Wir können uns in der Zerlegung $H_1 = (\ker T)^\perp \oplus (\ker T)$ auf $(\ker T)^\perp$ beschränken und deshalb annehmen: H_1 ist separabel und unendlich-dimensional (sonst ist bereits $T \in \mathcal{F}(H_1, H_2)$. Sei $(v_n)_{n\in\mathbb{N}}$ ONS in H_1 und P_k die Projektion auf $\mathrm{span}(v_1, \dots, v_k)$. Für beliebige $v \in H_1$ gilt nach der Parsevalschen Gleichung $v - P_k v \to 0$ für $k \to \infty$. Zu jedem $k \in \mathbb{N}$ wählen wir ein $w_k \in B_{H_1}$ mit $\|(TP_k - T)w_k\| \geq \frac{1}{2}\|TP_k - T\|$. Dann ist $w_k - P_k w_k$ schwach-konvergent gegen 0, damit $\|(TP_k - T)w_k\|$ Norm-konvergent gegen 0, somit $\|TP_k - T\|$ Norm-konvergent gegen 0. Wegen $TP_k \in \mathcal{F}(H_1, H_2)$ folgt die Behauptung. $\square$

Theorem 16.11 *Sei H komplexer Hilbert-Raum, $D \subset \mathbb{C}$ offen und zusammenhängend und $T : D \to \mathcal{L}(H)$ eine holomorphe Operator-wertige Funktion, so daß $T(z)$ kompakt ist für alle $z \in D$. Dann gilt* <u>*entweder*</u>

i) $(\mathbf{1}_H - T(z))^{-1}$ *existiert für kein $z \in D$,*

<u>*oder*</u>

ii) $(\mathbf{1}_H - T(z))^{-1}$ *existiert für alle $z \in \mathcal{D} \setminus S$, wobei $S \subset D$ eine diskrete Menge ohne Häufungspunkt ist. In diesem Fall ist $(\mathbf{1}_H - T(z))^{-1}$ holomorph auf $D \setminus S$, und $T(z_p)w = w$ hat für jedes $z_p \in S$ eine Lösung $0 \neq w \in H$.*

Beweis. Wir zeigen, daß in einer Umgebung jedes Punktes $z_0 \in D$ entweder i) oder ii) gilt. Da D offen und zusammenhängend, somit wegzusammenhängend ist, wobei eine offene Umgebung des Weges in D liegt, gilt die Alternative i) oder ii) konstant auf D.

Nach Holomorphie-Voraussetzung gibt es zu $z_0 \in D$ ein $r > 0$, so daß $\|T(z) - T(z_0)\| < \frac{1}{2}$ für alle $z \in U_r(z_0)$. Nach Satz 16.10 gibt es einen Operator $R \in \mathcal{F}(H)$ mit $\|T(z_0) - R\| < \frac{1}{2}$. Dann ist $\|T(z) - R\| < 1$, so daß nach Lemma 15.4 der Operator $\mathbf{1}_H - T(z) + R$ für $z \in U_r(z_0)$ invertierbar ist mit holomorphem Inversen.

Da R endlichen Rang hat, gilt $R(v) = \sum_{i=1}^N R_i(v)e_i$, wobei $\mathrm{im}\, R = \mathrm{span}(e_1, \dots, e_N)$ und die R_i lineare Funktionale auf H sind. Nach dem Rieszschen Darstellungssatz ist $R_i(v) = \langle w_i, v\rangle$ für ein $w_i \in H$, somit $R(v) = \sum_{i=1}^N \langle w_i, v\rangle e_i$. Jedes w_i definiert eine Funktion $U_r(z_0) \ni z \mapsto \phi_i(z) := ((\mathbf{1}_H - T(z) + R)^{-1})^* w_i \in H$. sowie eine Operator-wertige Funktion $R(z) := R(\mathbf{1}_H - T(z) + R)^{-1})$, für die gilt

$$R(z)v = R((\mathbf{1}_H - T(z) + R)^{-1})v) = \sum_{i=1}^N \langle w_i, (\mathbf{1}_H - T(z) + R)^{-1})v\rangle e_i = \sum_{i=1}^N \langle \phi_i(z), v\rangle e_i\,.$$

Es gilt $\mathbf{1}_H - T(z) = (\mathbf{1}_H - R(z))(\mathbf{1}_H - T(z) + R)$, so daß folgt:

a) $\mathbf{1}_H - T(z)$ ist bijektiv genau dann, wenn $(\mathbf{1}_H - R(z))$ bijektiv ist,

b) $T(z)w = w$ ist genau dann lösbar mit $w \neq 0$, wenn $R(z)\tilde{w} = \tilde{w}$ lösbar ist mit $\tilde{w} \neq 0$.

Sei nun $R(z)\tilde{w} = \tilde{w} = \sum_{i=1}^N \langle \phi_i(z), \tilde{w}\rangle e_i$, also $\tilde{w} = \sum_{i=1}^N \beta_i e_i$ und $\beta_n = \sum_{i=1}^N \langle \phi_n(z), e_i\rangle \beta_i$. Für festes z ist das ein lineares Gleichungssystem zur Bestimmung der β_i, welches genau dann lösbar ist, wenn für die Determinante $d(z) := \det(\delta_{mn} - \langle \phi_m(z), e_n\rangle)$ gilt: $d(z) = 0$. Sei $S_r := \{z \in U_r(z_0) \;:\; d(z) = 0\}$. Da auch $d(z)$ holomorph ist, gilt nach dem Identitätssatz für holomorphe Funktionen: Entweder es gilt $S_r = U_r(z_0)$ und damit Behauptung i), oder S_d hat keinen Häufungspunkt. Im zweiten Fall ist $d(z) \neq 0$ für $z \in U_r(z_0)\setminus S_r$, und die Gleichung $(\mathbf{1}_H - R(z))\tilde{w} = \tilde{v}$ ist für alle $\tilde{v} \in H$ lösbar durch den Ansatz $\tilde{w} = \tilde{v} + \sum_{i=1}^N \alpha_i e_i$, der auf

$$\begin{aligned} &\tilde{v} - \sum_{i=1}^N \langle \phi_i(z), \tilde{v}\rangle e_i + \sum_{i=1}^N \alpha_i e_i - \sum_{i,j=1}^N \alpha_j \langle \phi_i(z), e_j\rangle e_i = \tilde{v} \\ \Rightarrow \qquad &\alpha_n = \langle \phi_n(z), \tilde{v}\rangle + \sum_{j=1}^N \langle \phi_n(z), e_j\rangle \alpha_j \end{aligned}$$

führt. Wegen $d(z) \neq 0$ ist dieses Gleichungssystem für alle $\tilde{v} \in H$ lösbar, d.h. $(\mathbf{1}_H - R(z))$ und damit $(\mathbf{1}_H - T(z))$ ist bijektiv für alle $z \in U_r(z_0) \setminus S_r$. Das ist die Behauptung ii). □

Folgerung 16.12 (Fredholmsche Alternative) *Ist $T \in \mathcal{K}(H)$ kompakt, dann ist entweder $(\mathbf{1}_H - T)$ bijektiv, oder $\ker T \neq \{0\}$.*

Beweis. Theorem 16.11 mit $T(z) = zT$ für $D = \mathbb{C}$. Der Fall i) ist für $z = 0$ ausgeschlossen. □

Satz 16.13 *Ist H komplexer Hilbert-Raum und $T \in \mathcal{K}(H)$ kompakter Operator, dann gilt:*

i) $\mathrm{sp}(T)$ *ist diskrete Teilmenge von $\mathbb{C}$ ohne Häufungspunkt außer eventuell* 0.

ii) *Jedes $0 \neq \lambda \in \mathrm{sp}(T)$ ist Eigenwert, und die zugehörigen Eigenräume $E_\lambda = \ker(T - \lambda \mathbf{1}_H)$ sind für $\lambda \neq 0$ endlich-dimensional.*

iii) *Das Spektrum von T ist abzählbar oder endlich.*

Beweis. i) Die Operator-wertige Funktion $\mathbb{C} \ni z \mapsto T(z) := zT \in \mathcal{K}(H)$ ist auf ganz $\mathbb{C}$ holomorph. Nach Theorem 16.11 ist die Menge $S := \{z \in \mathbb{C} \;:\; zTv = v$ hat Lösung $v \neq 0\}$ diskret (da $0 \notin S$) und ohne Häufungspunkt. Ist $\frac{1}{\lambda} \notin S$, dann existiert $(T - \lambda \mathbf{1}_H)^{-1} = -\frac{1}{\lambda}\big(\mathbf{1}_H - \frac{1}{\lambda}T)^{-1}$, und für $\frac{1}{\lambda} \in S$ gibt es nach Folgerung 16.12 einen Eigenvektor.

ii) Die Einschränkung von T auf den Eigenraum $E_\lambda = \ker(T - \lambda \mathbf{1}_H)$ ist ebenfalls kompakt. Diese Einschränkung ist aber $\lambda \mathbf{1}_{E_\lambda}$, so daß $\dim(E_\lambda) < \infty$ für $\lambda \neq 0$.

iii) Nach Folgerung 16.9 ist $(\ker T)^\perp$ separabel, so daß es höchstens abzählbar viele linear unabhängige Eigenvektoren zu $\lambda \neq 0$ geben kann, somit höchstens abzählbar viele $0 \neq \lambda \in \mathrm{sp}(T)$. □

Lemma 16.14 *Ist H komplexer Hilbert-Raum und $T \in \mathcal{K}(H)$ kompakt und normal. Seien $\lambda \neq \mu$ verschiedene Eigenwerte von T (einer möglicherweise $= 0$), so gilt $E_\lambda \perp E_\mu$.*

Beweis. Ist T normal, so gilt $\|(T - \lambda \mathbf{1}_H)v\| = \|(T^* - \bar{\lambda}\mathbf{1}_H))v\|$ nach Lemma 14.7.iv), somit $E_\lambda(T) = E_{\bar{\lambda}}(T^*)$. Ist $0 \neq v_\lambda \in E_\lambda$ und $0 \neq w_\mu \in E_\mu$, so folgt

$$\mu\langle v_\lambda, w_\mu\rangle = \langle v_\lambda, Tw_\mu\rangle = \langle T^* v_\lambda, w_\mu\rangle = \lambda\langle v_\lambda, w_\mu\rangle$$

und damit $\langle v_\lambda, w_\mu\rangle = 0$. □

Theorem 16.15 (Spektralsatz für selbstadjungierte kompakte Operatoren) *Es sei H ein Hilbert-Raum und $T = T^* \in \mathcal{K}(H)$ ein selbstadjungierter kompakter Operator. Sei $\mathrm{sp}_p(T) := \{\lambda \in \mathbb{R} \;:\; E_\lambda = \ker(T - \mathbf{1}_H) \neq \{0\}\}$ das Eigenwertspektrum und $P_\lambda : H \to E_\lambda$ die orthogonale Projektion auf den Eigenraum E_λ. Dann gilt:*

i) *$H = \overline{\bigoplus_{\lambda \in \mathrm{sp}_p(T)} E_\lambda}$ und $T = \sum_{\lambda \in \mathrm{sp}(T)} \lambda P_\lambda$ mit Konvergenz der Reihe in der starken Operator-Topologie.*

ii) *H besitzt eine Orthonormalbasis aus Eigenvektoren von T, wobei höchstens abzählbar viele Basiselemente zu Eigenvektoren aus $\mathrm{sp}_p(T)\backslash\{0\}$ gehören.*

Beweis. i) Nach Lemma 16.14 gilt $E_\mu \perp E_\lambda$ für $\lambda \neq \mu$. Sei dann $\tilde{H} := \overline{\bigoplus_{\lambda \in \mathrm{sp}_p(T)} E_\lambda} \subset H = \tilde{H} \oplus \tilde{H}^\perp$. Wir zeigen $\tilde{H}^\perp = \{0\}$. Wegen $TE_\lambda \subset E_\lambda$ ist $T\tilde{H} \subset \tilde{H}$. Da T selbstadjungiert ist, ist $T\tilde{H}^\perp \subset \tilde{H}^\perp$, denn sei $v \in \tilde{H}$ und $w \in \tilde{H}^\perp$, dann gilt $\langle Tw, v\rangle = \langle w, Tv\rangle = 0$. Sei nun $\tilde{T} := T\big|_{\tilde{H}^\perp} : \tilde{H}^\perp \to \tilde{H}^\perp$ die Einschränkung. Wegen der Abgeschlossenheit von $\tilde{H}^\perp$ ist $\tilde{T}$ ein kompakter selbstadjungierter Operator, sein Spektrum ist nicht-leer, und jeder Spektralwert $\neq 0$ ist Eigenwert von $\tilde{T}$ und damit von T. Damit ist nach Konstruktion 0 der einzige Spektralwert von $\tilde{T}$, d.h. der Spektralradius ist $r(\tilde{T}) = 0$. Da für selbstadjungierte Operatoren $\tilde{T}$ gilt $r(\tilde{T}) = \|\tilde{T}\|$, folgt $\tilde{T} = 0$. Somit ist $\tilde{T}w = 0w$ für alle $w \in \tilde{H}^\perp$, d.h. $\tilde{H}^\perp \subset \tilde{H}$.

Somit gibt es zu $v \in H$ eine eindeutige Zerlegung $v = \sum_{\lambda \in \mathrm{sp}_p(T)} v_\lambda = \sum_{\lambda \in \mathrm{sp}_p(T)} P_\lambda v$, mit Norm-Konvergenz der Reihe. Damit folgt

$$Tv = T\Big(\sum_{\lambda \in \mathrm{sp}_p(T)} v_\lambda\Big) = \sum_{\lambda \in \mathrm{sp}_p(T)} Tv_\lambda = \sum_{\lambda \in \mathrm{sp}(T)} \lambda v_\lambda = \sum_{\lambda \in \mathrm{sp}(T)} \lambda P_\lambda v = \Big(\sum_{\lambda \in \mathrm{sp}(T)} \lambda P_\lambda\Big) v\,.$$

Damit gilt $T = \sum_{\lambda \in \mathrm{sp}(T)} \lambda P_\lambda$ in der starken Operator-Toplogie. Der mögliche Spektralwert 0 liefert keinen Beitrag zur Summe.

ii) Ist $\mathcal{B}_\lambda = \{v_1^{(\lambda)}, \ldots, v_{r_\lambda}^{(\lambda)})\}$ Orthonormalbasis von E_λ für $0 \neq \lambda \in \mathrm{sp}(T)$ und $\mathcal{B}_0 = \{v_i \;:\; i \in I\}$ eine Orthonormalbasis für E_0 im Fall $0 \in \mathrm{sp}(T)$, so ist $\mathcal{B} = \bigcup_{\lambda \in \mathrm{sp}(T)} \mathcal{B}_\lambda$ eine Orthonormalbasis von H aus Eigenvektoren von T. □

Da $\mathrm{sp}(T)$ höchstens abzählbar ist, gilt für eine beliebige Abzählung $\{\lambda_n \;:\; n \in \mathbb{N})$ der Spektralwerte $T = \sum_{n=0}^{\infty} \lambda_n P_{\lambda_n}$. Ordnen wir $|\lambda_0| > |\lambda_1| > \ldots$, dannn gilt $|\lambda_0| = \|T\|$ und $\lambda_n \to 0$ für $n \to \infty$.

Eine wichtige Anwendung ist der *beschränkte Funktionalkalkül* für kompakte selbstadjungierte Operatoren $T = \sum_{\lambda \in \mathrm{sp}(T)} \lambda P_\lambda$. Ist $f : \Lambda \to \mathbb{C}$ eine beschränkte Funktion, so werden durch

$$f(T) := \sum_{\lambda \in \mathrm{sp}(T)} f(\lambda) P_\lambda \tag{*}$$

beschränkte Funktionen des kompakten Operators T definiert, mit Konvergenz in der starken Operator-Topologie. Für stetige Funktionen stimmt diese Konstruktion mit dem stetigen Funktionalkalkül aus Theorem 15.16 überein: wegen der Orthogonalität $P_\lambda P_\mu = 0$ für $\lambda \neq \mu$ folgt $T^n = \sum_{\lambda \in \mathrm{sp}(T)} \lambda^n P_\lambda$, d.h. (*) gilt für Polynome und damit nach Stone-Weierstraß auch für stetige Funktionen, dann sogar mit Norm-Konvergenz der Reihe. Dagegen gilt (*) auch für bloß beschränkte Funktionen, aber die Reihe konvergiert i.a. nur in der starken Operator-Topologie.

Der Spektralsatz verallgemeinert sich auf weit größere Klassen von Operatoren.

Definition 16.16 Eine Abbildung $\mathbb{R} \ni \lambda \mapsto P_\lambda \in \mathcal{L}(H)$ heißt *Spektralschar*, wenn gilt

i) P_λ ist Projektor für jedes $\lambda \in \mathbb{R}$

ii) $P_\lambda \leq P_\mu$ für $\lambda \leq \mu$

iii) $\lim_{\lambda \to -\infty} P_\lambda = 0$ und $\lim_{\lambda \to +\infty} P_\lambda = \mathbf{1}_H$ in der schwachen Operator-Toplogie.

iv) $\lambda \mapsto P_\lambda$ ist rechtsseitig stetig, d.h. $\lim_{\mu \searrow \lambda} P_\mu = P_\lambda$

Gilt statt iii) bereits $P_a = 0$ und $P_b = \mathbf{1}_H$ für $a < b \in \mathbb{R}$, so hat die Spektralschar kompakten Träger.

Für $v \in \mathbb{H}$ ist die Funktion $\lambda \mapsto \langle v, P_\lambda v\rangle$ mononon wachsend und definiert ein Lebesque-Stieltjes-Maß $d\langle v, P_\lambda v\rangle$. Ist f eine beschränkte Funktion auf dem Träger der Spektralschar, so kann man in sinnvoller Weise einen beschränkten linearen Operator A_f definieren durch

$$\langle v, A_f v\rangle := \int_a^b f(\lambda)\, d\langle v, P_\lambda v\rangle \qquad \text{oder} \quad A_f := \int_a^b f(\lambda)\, dP_\lambda$$

im Sinne der schwachen Operator-Topologie. Man zeigt folgendes: Ist $A = A^* \in \mathcal{L}(H)$ selbstadjungiert und P_λ^A die Projektion auf $\ker((A - \lambda \mathbf{1}_H)_+)$ (dabei ist B_+ der positive Anteil von B), so bildet $\{P_\lambda^A\}$ eine Spektralschar, und für stetige Funktionen f stimmt das Integral $f(A) = \displaystyle\int_a^b f(\lambda)\; dP_\lambda^A$ mit der Konstruktion aus Theorem 15.16 überein. Insbesondere gilt $A = \displaystyle\int_a^b \lambda\; dP_\lambda^A$. Für jeden unitären Operator U findet man einen selbstadjungierten Operator A mit $\mathrm{sp}(A) \in [-\pi, \pi]$, so daß gilt $U = \exp(\mathrm{i}A)$. Dann lassen sich Funktionen unitärer Operatoren erklären als

$$f(U) = \int_{-\pi}^{\pi} f(e^{\mathrm{i}\lambda})\; dP_\lambda^A\ .$$

Dadurch läßt sich die Spektraldarstellung auf unbeschränkte selbstadjungierte Operatoren $T : \mathrm{dom}(T) \to H$ erweitern. Man zeigt, daß $U_T := (T - \mathrm{i}\mathbf{1}_H)(T + \mathrm{i}\mathbf{1}_H)^{-1}$ unitär ist, somit $U_t = \exp(\mathrm{i}A)$. Dann gilt für $v \in \mathrm{dom}(T)$

$$\langle v, Tv \rangle = \int_{]-\pi,\pi[} \tan\frac{\lambda}{2}\; d\langle v, P_\lambda^A v\rangle\ .$$