Funktionalanalysis

Günther Hörmann
Fakultät für Mathematik
Universität Wien
guenther.hoermann@univie.ac.at

Sommersemester 2021

(korrigierte Fassung vom 26. August 2023)

An Vorkenntnissen setzen wir die Inhalte folgender Lehrveranstaltungen (gemäß Curriculum 2014) voraus: Analysis bzw. Lineare Algebra und Geometrie aus den ersten drei Semestern sowie die Grundbegriffe der Topologie (wie z.B. in [Hoe20] behandelt). Für das Auswahlaxiom und einige dazu äquivalente Aussagen verweisen wir auf den Appendix.

Für den Aufbau der Vorlesungsinhalte habe ich mich am stärksten an den beiden Büchern [Wer18] und [Heu06] orientiert.

Ich bedanke mich herzlich bei allen Studierenden, die mich (teilweise noch lange nach der Vorlesung im Sommersemester 2021) auf Ungereimtheiten und Fehler im Skriptum aufmerksam gemacht haben, was auch nachträglich zu Verbesserungen im Text geführt hat.

Inhaltsverzeichnis

1 Banachräume

Im Rahmen dieser VO betrachten wir ausschließlich Vektorräume über dem Grundkörper $\mathbb{R}$ oder $\mathbb{C}$, den wir immer dort mit $\mathbb{K}$ bezeichnen, wo der Unterschied nicht wichtig ist.

1.1. Definition: Ein *normierter Raum* ist ein $\mathbb{K}$-Vektorraum X zusammen mit einer *Norm*, d.h. einer Abbildung $X \to \mathbb{R}, x \mapsto \|x\|$, sodass für alle $x, y \in X$ und $\forall \lambda \in \mathbb{K}$ gilt:

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

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

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

Aus (iii) folgt leicht die sogenannte *umgekehrte Dreicksungleichung*: $|\|x\| - \|y\|| \leq \|x - y\|$.

Bekanntlich wird durch eine Norm auf X mittels $d(x, y) := \|x - y\|$ eine Metrik definiert und wir werden einen normierten Raum X stets auf diese Art als metrischen Raum auffassen, falls nichts anderes gesagt wird. Somit ist eine Folge (x_n) in X also genau dann eine *Cauchyfolge*, falls es zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ gibt, sodass für $n, m \geq n_0$ stets $\|x_m - x_n\| < \varepsilon$ gilt. *Konvergenz* der Folge (x_n) gegen $a \in X$ bedeutet, dass $\lim_{n\to\infty} \|x_n - a\| = 0$ gilt, was wiederum äquivalent dazu ist, dass für jedes $\varepsilon > 0$ die Folgenglieder schließlich in der *offenen Kugel*

$$U_\varepsilon(a) = \{x \in X \mid \|x - a\| < \varepsilon\}$$

vom Radius ε um a liegen. Letztere bilden ja die typischen metrischen ε-Umgebungen des Punktes a und somit eine Umgebungsbasis.

1.2. Kugeln und Sphären in normierten Räumen: Für $a \in X$ und $r > 0$ haben wir neben der offenen Kugel $U_r(a)$ mit Radius r um a auch noch die *abgeschlossene Kugel*

$$K_r(a) := \{x \in X \mid \|x - a\| \leq r\}.$$

Wir erinnern daran, dass die Offenheit von $U_r(a)$ und die Abgeschlossenheit von $K_r(a)$ aus der Stetigkeit der Abbildung $X \to \mathbb{R}$, $x \mapsto d(x, a) = \|x - a\|$ folgen. Während in allgemeinen metrischen Räumen nur $\overline{U_r(a)} \subseteq K_r(a)$ garantiert war, gilt in normierten Räumen stets

$$\overline{U_r(a)} = K_r(a)$$

(Beweis als UE). Somit können wir dies mit der *Sphäre* $S_r(a) := \{x \in X \mid \|x - a\| = r\}$ vom Radius r um a auch so schreiben: $\overline{U_r(a)} = K_r(a) = U_r(a) \cup S_r(a)$.

Der Beweis der folgenden grundlegenden Aussagen ist eine UE-Aufgabe.

1.3. Proposition: Sei X ein normierter Raum, dann sind folgende Abbildungen stetig:

(i) Addition $X \times X \to X$, $(x, y) \mapsto x+y$, d.h. $x_n \to x$ und $y_n \to y$ impliziert $x_n+y_n \to x+y$,

(ii) Multiplikation mit Skalaren $\mathbb{K} \times X \to X$, $(\lambda, x) \mapsto \lambda x$,
d.h. aus $\lambda_n \to \lambda$ (in $\mathbb{K}$) und $x_n \to x$ (in X) folgt $\lambda_n x_n \to \lambda x$,

(iii) die Norm als Abbildung $X \to \mathbb{R}$, $x \mapsto \|x\|$, d.h. $x_n \to x$ erzwingt $\|x_n\| \to \|x\|$.

1.4. Definition: Ein normierter Raum $(X, \| \;\|)$ heißt *Banachraum*, falls er (als metrischer Raum) vollständig ist, also jede Cauchyfolge in X konvergent ist.

1.5. Teilräume von normierten Räumen: Ein Teilraum U eines normierten Raumes X wird durch Einschränkung der Norm selbst zu einem normierten Raum. Weiters gilt:

(i) Der Abschluss $\overline{U}$ ist ebenfalls ein Teilraum von X.

(ii) Ist X ein Banachraum und U abgeschlossen, dann ist auch U ein Banachraum.

(iii) Ist U vollständig, so ist U auch abgeschlossen.

Die Beweise liegen auf der Hand: (i) folgt aus der oben bewiesenen Stetigkeit von Addition und Multiplikation mit Skalaren in Kombination mit der Teilraumeigenschaft, während (ii) und (iii) nur auf allgemeinen Eigenschaften von Teilmengen metrischer Räume beruhen.

1.6. Bemerkung: (i) Wie in der Topologie-VO bereits für metrische (oder uniforme) Räume diskutiert, kann ein nicht vollständiger normierter Raum X zunächst mit dichtem Bild und isometrisch in einen vollständigen metrischen Raum $\hat{X}$ eingebettet werden. Wir erinnern daran, dass $\hat{X}$ aus Äquivalenzklassen $[(x_n)]$ von Cauchyfolgen (x_n) konstruiert werden kann, wobei $(x_n) \sim (y_n)$ einfach $\|x_n - y_n\| = d(x_n, y_n) \to 0$ bedeutet. Auf $\hat{X}$ lässt sich durch komponentenweise Operationen leicht eine Vektorraumstruktur angeben und durch $\|[(x_n)]\|_{\hat{X}} := \lim \|x_n\|$ eine Norm definieren bzgl. der $\hat{X}$ ein Banachraum ist und als *Vervollständigung* des normierten Raumes X bezeichnet wird.

(ii) Sei $(X, \| \;\|)$ ein normierter Raum. In den UE beweisen wir folgende Äquivalenz:

X ist vollständig $\iff$ Jede absolut konvergente Reihe konvergiert in X.

(Die zweite Eigenschaft heißt im Detail: Zu jeder Folge (a_n) in X mit $\sum_{n=1}^{\infty} \|a_n\| < \infty$ gibt es ein $z \in X$ mit $z = \lim_{N\to\infty} \sum_{n=1}^{N} a_n$).

1.7. Vergleichbarkeit und Äquivalenz von Normen: Seien $\| \;\|_1$ und $\| \;\|_2$ zwei Normen auf dem $\mathbb{K}$-Vektorraum X. Die Norm $\| \;\|_2$ heißt *stärker* als die Norm $\| \;\|_1$, falls es eine Konstante $C > 0$ gibt, sodass $\|x\|_1 \le C\|x\|_2$ für alle $x \in X$ gilt. Dies ist gleichbedeutend damit, dass aus $\|x_n - a\|_2 \to 0$ stets $\|x_n - a\|_1 \to 0$ folgt, also Konvergenz einer Folge (x_n) gegen $a \in X$ bzgl. $\| \;\|_2$ auch jene bzgl. $\| \;\|_1$ erzwingt. (Die von $\| \;\|_2$ induzierte metrische Topologie ist feiner als jene von $\| \;\|_1$.)

Ist $\| \, \|_2$ stärker als $\| \, \|_1$ und auch $\| \, \|_1$ stärker als $\| \, \|_2$, dann sprechen wir von *äquivalenten Normen*. Das trifft genau dann zu, falls es Konstanten $C_1, C_2 > 0$ gibt, sodass

$$\forall x \in X : \quad C_1 \|x\|_2 \leq \|x\|_1 \leq C_2 \|x\|_2.$$

In diesem Fall stimmen also die Konvergenzbegriffe für beide Normen (Metriken) überein.

1.8. Proposition: Sei X ein endlichdimensionaler $\mathbb{K}$-Vektorraum, dann gilt:

(i) Auf X sind je zwei Normen äquivalent.

(ii) X ist mit jeder Norm ein Banachraum.

Beweis. Um (i) zu zeigen nehmen wir $\dim X = n$ an und wählen eine Basis $e_1, \ldots, e_n$ von X. Wegen der Eindeutigkeit und Linearität der Koeffizienten in jeder Basisentwicklung erhalten wir durch

$$\left\| \sum_{j=1}^{n} \alpha_j e_j \right\|_2 := \sqrt{\sum_{j=1}^{n} |\alpha_j|^2}$$

eine Norm auf X. Sei weiters $\| \, \|$ eine beliebige Norm auf X.

Behauptung 1: $\| \, \|_2$ ist stärker als $\| \, \|$.

Dies folgt mittels $C := \max(\|e_1\|, \ldots, \|e_n\|) > 0$, der Dreiecksungleichung und der Cauchy-Schwarz-Ungleichung (in $\mathbb{R}^n$) für $x = \sum_{j=1}^{n} \alpha_j e_j$ aus

$$\|x\| = \left\| \sum_{j=1}^{n} \alpha_j e_j \right\| \leq \sum_{j=1}^{n} |\alpha_j| \, \|e_j\| \leq \sqrt{\sum_{j=1}^{n} |\alpha_j|^2} \sqrt{\sum_{j=1}^{n} \|e_j\|^2} \leq \|x\|_2 \sqrt{\sum_{j=1}^{n} C^2} = \sqrt{n} C \|x\|_2.$$

Behauptung 2: $\| \, \|$ ist stärker als $\| \, \|_2$.

Gemäß Behauptung 1 ist $f \colon (X, \| \, \|_2) \to \mathbb{R}$, gegeben durch $f(x) := \|x\|$ für $x \in X$, eine stetige Abbildung, weil $|f(x) - f(y)| = |\|x\| - \|y\|| \leq \|x - y\| \leq \sqrt{n} C \|x - y\|_2$ gilt. Wir bezeichnen mit h die Einschränkung von f auf die Einheitssphäre $S_1(0) = \{x \in X \mid \|x\|_2 = 1\}$ bzgl. $\| \, \|_2$. Letztere ist gerade das Bild der euklidischen Sphäre $S^n \subseteq \mathbb{K}^n$ unter dem isometrischen Koordinatenisomorphismus $\kappa \colon \mathbb{K}^n \to X$, $(\alpha_1, \ldots, \alpha_n) \mapsto \sum_{j=1}^{n} \alpha_j e_j$, d.h. $S_1(0) = \kappa(S^n)$. Als isometrische Abbildung ist κ auch stetig. Daher folgt aus der Kompaktheit von S^n auch jene von $S_1(0)$. Somit nimmt die reellwertige Funtion h auf $S_1(0)$ ihr Minimum μ an, d.h. es gibt einen Punkt x_0 mit $\|x_0\|_2 = 1$ (daher $x_0 \neq 0$) und $0 < \|x_0\| = h(x_0) = \mu \leq \|x\|$ für alle $x \in X$ mit $\|x\|_2 = 1$. Für jedes $y \neq 0$ erhalten wir daraus wegen $y / \|y\|_2 \in S_1(0)$ die Ungleichung

$$\|y\|_2 \leq \frac{1}{\mu} \|y\|,$$

die trivialer Weise auch für $y = 0$ gilt.

Behauptungen 1 und 2 zeigen, dass jede Norm auf X zu $\| \ \|_2$ äquivalent ist. Mittels elementarer Ungleichungsketten folgt, dass je zwei beliebige Normen auf X äquivalent sind.

(ii) folgt nun aus der Beobachtung, dass der oben erwähnte Koordinatenisomorphismus κ als Isometrie die Vollständigkeit des $\mathbb{K}^n$ mit der euklidischen Metrik auf $(X, \| \ \|_2)$ überträgt und somit wegen (i) auch auf X mit jeder anderen Norm. □

Im Verlauf des obigen Beweises hatten wir die Kompaktheit einer Einheitssphäre im Endlichdimensionalen mittels Stetigkeit des Koordinatenismorphismus erkannt. Mit derselben Technik und einfachen Translationen und Skalierungen ergibt sich folgende Aussage.

1.9. Korollar: Jede abgeschlossene Kugel und jede Sphäre eines endlichdimensionalen normierten Raumes ist kompakt.

Der folgende Satz beinhaltet sogar eine Umkehrung davon.

1.10. Theorem: Ein normierter Raum ist genau dann endlichdimensional, wenn seine abgeschlossene Einheitskugel kompakt ist.

Beweis. Es ist nur noch zu zeigen, dass aus der Kompaktheit von $K_1(0) = \{x \in X \mid \|x\| \leq 1\}$ im normierten Raum $(X, \| \ \|)$ stets $\dim X < \infty$ folgt. Durch $(U_{1/2}(x))_{x \in K_1(0)}$ erhalten wir eine offene Überdeckung von $K_1(0)$, für die es eine endliche Teilüberdeckung geben muss, d.h. wir können $z_1, \ldots, z_m \in K_1(0)$ finden, sodass $K_1(0) \subseteq \bigcup_{j=1}^m U_{1/2}(z_j)$ gilt. Es sei $E := \operatorname{span}\{z_1, \ldots, z_m\}$, dann ist also zu zeigen, dass $X = E$ gilt. Dafür genügt es wiederum, $K_1(0) \subseteq E$ nachzuweisen, weil die Vielfachen von $K_1(0)$ natürlich ganz X erreichen.

Zunächst bemerken wir, dass E als endlichdimensionaler normierter Raum gemäß Proposition 1.8 ein Banachraum ist, also nach 1.5(iii) insbesondere auch abgeschlossen.

Aus der anfangs konstruierten Überdeckung erhalten wir zunächst

$$K_1(0) \subseteq \bigcup_{j=1}^m U_{1/2}(z_j) = \bigcup_{j=1}^m \big(z_j + U_{1/2}(0)\big) \subseteq E + U_{1/2}(0) \subseteq E + \frac{1}{2} K_1(0)$$

und dann iterativ in n Schritten

$$K_1(0) \subseteq E + \frac{1}{2} K_1(0) \subseteq E + \frac{1}{2}\left(E + \frac{1}{2} K_1(0)\right) = E + \frac{1}{4} K_1(0) \subseteq \ldots \subseteq E + \frac{1}{2^n} K_1(0).$$

Sei nun $x \in K_1(0)$ beliebig. Für jedes $n \in \mathbb{N}$ gibt es nach obiger Inklusionsrelation ein $x_n \in E$ mit $\|x - x_n\| \leq 1/2^n$; daher folgt $x_n \to x \in \overline{E} = E$. Somit gilt also $K_1(0) \subseteq E$. □

1.11. Beispiel: Die sogenannten *p-Normen* auf $\mathbb{K}^n$ sind gegeben durch

$$\|x\|_p := \left(\sum_{j=1}^n |x_j|^p \right)^{1/p} \quad \text{für } 1 \leq p < \infty \text{ und } \|x\|_\infty := \max(|x_1|, \ldots, |x_n|).$$

Nur der Nachweis der Dreiecksungleichung für die Fälle $1 \leq p < \infty$ ist nichtrivial und entspricht genau der aus der Analysis bekannten Minkowski-Ungleichung. Gemäß Proposition 1.8 sind all diese Normen äquivalent und machen $\mathbb{K}^n$ jeweils zum Banachraum.

1.12. Beispiel: Für $1 \leq p < \infty$ bezeichnen wir mit l^p die Menge aller Folgen $x = (x_n)_{n\in\mathbb{N}}$ in $\mathbb{K}$ mit der Eigenschaft $\sum_{n=1}^{\infty} |x_n|^p < \infty$ und setzen

$$\|x\|_p := \left(\sum_{n=1}^{\infty} |x_n|^p\right)^{1/p}.$$

Die üblichen komponentenweisen Operationen geben l^p eine $\mathbb{K}$-Vektorraumstruktur, wobei z.B. für $x = (x_n), y = (y_n) \in l^p$ die Summenfolge $x + y = (x_n + y_n)$ auf folgende Art als Element von l^p nachgewiesen werden kann:

$$\begin{aligned}\sum_{n=1}^{\infty} |x_n + y_n|^p &\leq \sum_{n=1}^{\infty} (|x_n| + |y_n|)^p \leq \sum_{n=1}^{\infty} (2\max(|x_n|, |y_n|))^p = 2^p \sum_{n=1}^{\infty} \max(|x_n|^p, |y_n|^p) \\ &\leq 2^p \sum_{n=1}^{\infty}(|x_n|^p + |y_n|^p) = 2^p \sum_{n=1}^{\infty} |x_n|^p + 2^p \sum_{n=1}^{\infty} |y_n|^p < \infty.\end{aligned}$$

Beim Nachweis, dass $\| \; \|_p$ eine Norm auf l^p ergibt, steckt wieder die Minkowski-Ungleichung hinter der Dreiecksungleichung und die anderen Eigenschaften sind klar.

Wir zeigen nun auch noch die Vollständigkeit, d.h. dass l^p ein Banachraum ist: Sei $(x^{(k)})_{k\in\mathbb{N}}$ eine Cauchyfolge in l^p, wobei $x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \ldots) = (x_n^{(k)})$ mit $x_n^{(k)} \in \mathbb{K}$ ist. Bei beliebigem fixen $n \in \mathbb{N}$ ist $(x_n^{(k)})_{k\in\mathbb{N}}$ eine Cauchyfolge in $\mathbb{K}$, weil für $y = (y_m) \in l^p$ stets $|y_m| \leq \|y\|_p$ für jedes $m \in \mathbb{N}$ gilt und somit also speziell $|x_n^{(k)} - x_n^{(l)}| \leq \left\|x^{(k)} - x^{(l)}\right\|_p$ folgt (mit $m = n$ und $y = x^{(k)} - x^{(l)}$). Wir[1] setzen $x_n := \lim_{l\to\infty} x_n^{(l)}$ und betrachten die Folge $x := (x_n)$.
Wir behaupten, dass $x \in l^p$ und $x^{(l)} \to x$ gilt.
Zu vorgegebenem $\varepsilon > 0$ wählen wir gemäß Cauchyfolgeneigenschaft ein $k_0 \in \mathbb{N}$, sodass für alle $k, l \geq k_0$ stets $\left\|x^{(k)} - x^{(l)}\right\|_p < \varepsilon$ garantiert ist. Daher gilt für $k, l \geq k_0$ und Partialsummen mit beliebiger Obergrenze $N \in \mathbb{N}$ die Abschätzung

$$\left(\sum_{n=1}^{N} |x_n^{(k)} - x_n^{(l)}|^p\right)^{1/p} \leq \left\|x^{(k)} - x^{(l)}\right\|_p < \varepsilon.$$

Wir dürfen also ganz links zum Grenzwert für $k \to \infty$ übergehen (endliche nichtnegative Summe mit oberer Schranke ε unabhängig von k) und erhalten

$$\left(\sum_{n=1}^{N} |x_n - x_n^{(l)}|^p\right)^{1/p} \leq \varepsilon,$$

[1] verwenden die Vollständigkeit von $\mathbb{K}$ und

was wegen der Beliebigkeit von N (und der Unabhängigkeit der oberen Schranke ε von N) auch noch

$$\left(\sum_{n=1}^{\infty} |x_n - x_n^{(l)}|^p\right)^{1/p} \leq \varepsilon$$

ergibt. Daher ist $x - x^{(l)} \in l^p$ und weiter $x = (x - x^{(l)}) + x^{(l)} \in l^p$. Schließlich zeigt die obige Abschätzung $\left\|x - x^{(l)}\right\|_p \leq \varepsilon$ für alle $l \geq k_0$. Somit konvergiert $(x^{(l)})$ gegen x.

1.13. Beispiel: Sei T eine nichtleere Menge und $B(T)$ die Menge aller beschränkten Funktionen $x\colon T \to \mathbb{K}$. Es ist ganz leicht nachzuweisen, dass $B(T)$ mit den üblichen punktweise definierten Operationen ein Vektorraum ist und

$$\|x\|_\infty := \sup_{t \in T} |x(t)|$$

eine Norm ergibt. Die Konvergenz einer Funktionenfolge (x_n) bzgl. dieser Supremumsnorm entspricht natürlich genau der gleichmäßigen Konvergenz auf T.

Wir zeigen, dass $(B(T), \| \ \|_\infty)$ ein Banachraum ist: Sei (x_n) eine $\| \ \|_\infty$-Cauchyfolge, dann folgt aus der für alle $y \in B(T)$ gültigen Abschätzung $|y(t)| \leq \|y\|_\infty$ für jedes $t \in T$, dass $(x_n(t))_{n\in\mathbb{N}}$ eine Cauchyfolge in $\mathbb{K}$ ist für jedes $t \in T$. Wir[2] definieren die Funktion $x\colon T \to \mathbb{K}$ durch $x(t) := \lim_{n\to\infty} x_n(t)$.
Wir behaupten, dass $x \in B(T)$ und $x_n \to x$ gilt.
Sei $\varepsilon > 0$ vorgegeben. Wir wählen gemäß Cauchyfolgeneingenschaft ein $n_0 \in \mathbb{N}$, sodass für $m, n \geq n_0$ stets $\|x_m - x_n\|_\infty < \varepsilon$ gilt. Sei $t \in T$ beliebig, dann gibt es wegen $x_m(t) \to x(t)$ ein $m_0 \in \mathbb{N}$ (abhängig von ε und t) mit $|x_m(t) - x(t)| < \varepsilon$ für alle $m \geq m_0$. Zusammen erhalten wir nun für ein fixes $m_1 \geq \max(n_0, m_0)$ und alle $n \geq n_0$

$$\begin{aligned}(1.1)\quad |x(t) - x_n(t)| &= |x(t) - x_{m_1}(t) + x_{m_1}(t) - x_n(t)| \\ &\leq |x(t) - x_{m_1}(t)| + |x_{m_1}(t) - x_n(t)| \leq \varepsilon + \|x_{m_1} - x_n\|_\infty \leq \varepsilon + \varepsilon = 2\varepsilon,\end{aligned}$$

wobei wir also speziell für $|x(t) - x_{n_0}(t)|$ eine obere Schranke unabhängig von $t \in T$ erhalten haben. Somit gilt

$$\forall t \in T : |x(t)| = |x(t) - x_{n_0}(t) + x_{n_0}(t)| \leq |x(t) - x_{n_0}(t)| + |x_{n_0}(t)| \leq 2\varepsilon + \|x_{n_0}\|_\infty,$$

daher dürfen wir hier das Supremum über alle $t \in T$ nehmen, was $x \in B(T)$ zeigt. Die Abschätzung (1.1) lehrt nun, dass $\|x - x_n\|_\infty \leq 2\varepsilon$ für alle $n \geq n_0$ gilt. Somit folgt $x_n \to x$.

1.14. Beispiel: Als Spezialfall des vorigen Beispiels mit $T = \mathbb{N}$ erhalten wir den Banachraum $l^\infty := B(\mathbb{N})$ der beschränkten Folgen $x = (x_n)_{n\in\mathbb{N}}$ in $\mathbb{K}$ mit der Norm

$$\|x\|_\infty = \sup_{n\in\mathbb{N}} |x_n|.$$

[2] verwenden wiederum die Vollständigkeit von $\mathbb{K}$ und

1.15. Beispiel: Wir betrachten die Teilräume c und c_0 von $(l^\infty, \| \|_\infty)$, wobei c die Menge der konvergenten Folgen in $\mathbb{K}$ bezeichne und c_0 jene davon mit Grenzwert 0. Klarerweise ist $c_0 \subseteq c$ und $c_0 \neq c$. Wir werden beweisen, dass beide Teilräume abgeschlossen sind, also nach 1.5 selbst Banachräume sind (bzgl. $\| \|_\infty$).

Sei $(x^{(k)})_{k\in\mathbb{N}}$ eine Folge in c, die bzgl. $\| \|_\infty$ gegen $x \in l^\infty$ konvergiert. Wir müssen zeigen, dass $x \in c$ gilt.

Es sei $x^{(k)} = (x_n^{(k)})_{n\in\mathbb{N}}$, $x = (x_n)_{n\in\mathbb{N}}$ und für jedes $k \in \mathbb{N}$ jeweils $a_k := \lim_{n\to\infty} x_n^{(k)}$. Die in l^∞ konvergente Folge $(x^{(k)})$ ist auch eine $\| \|_\infty$-Cauchyfolge, daher ist wegen

$$|a_k - a_l| = |\lim_{n\to\infty}(x_n^{(k)} - x_n^{(l)})| \leq \sup_{n\in\mathbb{N}} |x_n^{(k)} - x_n^{(l)}| = \left\|x^{(k)} - x^{(l)}\right\|_\infty$$

die Folge $(a_k)_{k\in\mathbb{N}}$ eine Cauchyfolge in $\mathbb{K}$; es existiert also $a := \lim_{k\to\infty} a_k \in \mathbb{K}$.
Wir behaupten, dass auch $\lim_{n\to\infty} x_n = a$ gilt, somit insbesondere $x = (x_n) \in c$ wie erhofft. Sei $\varepsilon > 0$ vorgegeben. Wir wählen $k_0, n_0 \in \mathbb{N}$ mit den Eigenschaften

$$\left\|x - x^{(k_0)}\right\|_\infty < \frac{\varepsilon}{3}, \quad |a_{k_0} - a| < \frac{\varepsilon}{3}, \quad \forall n \geq n_0 \colon |x_n^{(k_0)} - a_{k_0}| < \frac{\varepsilon}{3}.$$

Dann gilt für jedes $n \geq n_0$

$$|x_n - a| \leq |x_n - x_n^{(k_0)}| + |x_n^{(k_0)} - a_{k_0}| + |a_{k_0} - a| \leq \left\|x - x^{(k_0)}\right\|_\infty + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} < \varepsilon;$$

also folgt $x_n \to a$.

Die Abgeschlossenheit von c_0 beweisen wir wie oben mit der Beobachtung, dass dann $a_k = 0$ für jedes k ist und somit auch $a = 0$, woraus $x \in c_0$ folgt.

1.16. Beispiel: Sei T ein metrischer Raum und

$$C_b(T) := \{f \colon T \to \mathbb{K} \mid f \text{ stetig und beschränkt}\}.$$

Für kompaktes T ist einfach $C_b(T) = C(T, \mathbb{K}) := \{f \colon T \to \mathbb{K} \mid f \text{ stetig}\}$. In jedem Fall ist $C_b(T)$ ein Teilraum des Banachraumes $(B(T), \| \|_\infty)$ (vgl. Beispiel 1.13). Er ist auch abgeschlossen, weil der gleichmäßige Limes einer Folge stetiger Funktionen stetig ist; dies wurde in der Analysis wohl bewiesen, wenn T ein reelles Intervall ist; auch im allgemeineren Fall ist hier nichts anderes als der „typische $\frac{\varepsilon}{3}$-Beweis“ nötig. Für $f_n \in C_b(T)$ mit $f_n \to f$ in $B(T)$ wollen wir also zeigen, dass f in einem beliebig vorgebenen Punkt $t_0 \in T$ stetig ist: Sei $\varepsilon > 0$ vorgegeben. Wir wählen $N \in \mathbb{N}$ mit $\|f - f_n\|_\infty < \varepsilon/3$ für $n \geq N$ und zusätzlich gemäß Stetigkeit von f_N ein passendes $\delta > 0$, sodass $d(t, t_0)$ stets $|f_N(t) - f_N(t_0)| < \varepsilon/3$ nach sich zieht. Solcherart vorbereitet können wir $f(t) - f(t_0) = f(t) - f_N(t) + f_N(t) - f_N(t_0) + f_N(t_0) - f(t_0)$ schreiben und gewinnen die Abschätzung

$$\begin{aligned}|f(t) - f(t_0)| &\leq |f(t) - f_N(t)| + |f_N(t) - f_N(t_0)| + |f_N(t_0) - f(t_0)| \\ &\leq \|f - f_N\|_\infty + \frac{\varepsilon}{3} + \|f_N - f\|_\infty < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.\end{aligned}$$

Somit ist auch $(C_b(T), \| \|_\infty)$ ein Banachraum.

Wir führen nun recht kursorisch als Beispiele auch jene äußerst wichtigen Banachräume ein, die aus dem *Lebesgue-Integral* (bzw. aus der *Maßtheorie*) entspringen. Falls die Details dazu noch nicht aus der Analysis oder anderen Zusammenhängen bekannt sind, bieten sich als Referenz dafür entsprechende Abschnitte in folgenden Quellen an (sortiert ungefähr nach Umfang und Allgemeinheit): [Mus14, Wer18, Cons16, Fried82, Foll99, El18].

1.17. Beispiel: Sei $I \subseteq \mathbb{R}$ ein Intervall und $1 \leq p < \infty$. Zunächst wird der Vektorraum aller messbaren Funktionen $f\colon I \to \mathbb{K}$, sodass $|f|^p$ Lebesgue-integrierbar ist, in Klassen modulo Gleichheit fast überall eingeteilt (das entspricht dem Quotientenvektorraum nach dem Teilraum aller Funktionen, die fast überall null sind). Wir folgen dem üblichen „abuse of notation" und bezeichnen auch diese Klassen wieder wie die Funktionen selbst mit f (wobei f also eigentlich ein gewählter Repräsentant seiner Klasse ist) und den so entstandenen Vektorraum (der Klassen) mit $L^p(I)$. Wir erhalten durch

$$\|f\|_p := \left(\int\limits_I |f(x)|^p \, dx \right)^{1/p}$$

eine Norm auf $L^p(I)$ (basierend auf der Minkowsi-Ungleichung) und die starken Konvergenzsätze der Lebesgueschen Integrationstheorie erlauben den Beweis, dass dadurch ein Banachraum definiert wird. Es lässt sich schließlich auch noch zeigen, dass $C(I) \cap L^p(I)$ bzgl. $\| \ \|_p$ dicht in $L^p(I)$ ist, was für $p = 1$ gerade darauf hinweist, dass durch Lebesgue die älteren Integralbegriffe von Cauchy und Riemann „vervollständigt" wurden.
Allgemeiner, aber im Grundkonzept analog, können die Banachräume $L^p(\Omega, \mu)$ konstruiert werden, wobei μ ein Maß auf (einer σ-Algebra von Teilmengen von) einer Menge Ω ist. Übrigens ergibt sich dann für das Beispiel des Zählmaßes auf $\Omega = \mathbb{N}$ gerade $L^p(\mathbb{N}, \mu) = l^p$.

1.18. Beispiel: Es sei wieder $I \subseteq \mathbb{R}$ ein Intervall und wir gehen diesmal aus vom Vektorraum aller messbaren Funktionen $f\colon I \to \mathbb{K}$, sodass f außerhalb einer Lebesgue-Nullmenge beschränkt ist. Wiederum nehmen wir den Quotienten nach dem Teilraum der Funktionen, die fast überall verschwinden, und erhalten so den Vektorraum $L^\infty(I)$ (ebenso mit dem üblichen „abuse of notation", dass Funktionen für ihre Klassen stehen). Durch

$$\|f\|_\infty := \inf\{ \sup_{t \in I \setminus N} |f(t)| \mid N \subseteq I \text{ ist Lebesgue-Nullmenge}\}$$

erhalten wir eine Norm, die sogenannte *essentielle Supremumsnorm*, auf $L^\infty(I)$ und der Beweis der Vollständigkeit ist, abgesehen von technischem Hantieren mit Nullmengen, zurückführbar auf jene der Supremumsnorm für beschränkte Funktionen in Beispiel 1.13. Also ist $L^\infty(I)$ ein Banachraum.
Maßtheoretisch allgemeiner kann wieder $L^\infty(\Omega, \mu)$ als Banachraum eingeführt werden und l^∞ ergibt sich wieder als Spezialfall aus dem Zählmaß auf $\Omega = \mathbb{N}$.

Wir erinnern daran, dass ein topologischer Raum X *separabel* heißt, wenn er eine abzählbare dichte Teilmenge A besizt, d.h. A ist abzählbar und $\overline{A} = X$. Für einen normierten Raum können wir dies mit Hilfe der Dichtheit von gewissen Teilräumen umformulieren.

1.19. Lemma: Für einen ein normierten Raum X sind folgende Aussagen äquivalent:

(i) X ist separabel,

(ii) Es gibt eine abzählbare Teilmenge $A \subseteq X$, sodass $\operatorname{span} A$ dicht ist in X.

Beweis. (i) $\Rightarrow$ (ii) ist wegen $A \subseteq \operatorname{span} A$ klar.

(ii) $\Rightarrow$ (i): Details hier nur für $\mathbb{K} = \mathbb{R}$; im Komplexen sind die Überlegungen analog, wobei dann $\mathbb{Q}$ durch $\mathbb{Q} + i\mathbb{Q}$ ersetzt werden muss. Wir betrachten die Menge $B := \operatorname{span}_{\mathbb{Q}} A$ der Linearkombinationen von Elementen aus A mit rationalen Koeffizienten. Da Linearkombinationen stets aus endlichen Teilmengen von A gebildet werden, ist B abzählbar. Es ist noch zu zeigen, dass B dicht in X ist.

Sei $x \in X$ und $\varepsilon > 0$. Laut (ii) existieren $n \in \mathbb{N}$, $a_1, \ldots, a_n \in A$ und $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ mit $\|x - a\| < \varepsilon/2$, wobei $a := \lambda_1 a_1 + \ldots + \lambda_n a_n \in \operatorname{span} A$. Es ist OBdA $\alpha := \|a_1\| + \ldots + \|a_n\| > 0$. Für jedes $l = 1, \ldots, n$ wählen wir $r_l \in \mathbb{Q}$ mit $|\lambda_l - r_l| < \varepsilon/(2\alpha)$ und setzen $b := r_1 a_1 + \ldots r_n a_n$. Dann ist $b \in B$ und

$$\|x - b\| \leq \|x - a\| + \|a - b\| < \frac{\varepsilon}{2} + \sum_{l=1}^{n} |\lambda_l - r_l| \|a_l\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2\alpha} \sum_{l=1}^{n} \|a_l\| = \varepsilon.$$

□

Wir erinnern daran, dass sich in *metrischen* Räumen die Separabilität auf die Spurtopologien vererbt. Daher sind Teilräume separabler normierter Räume stets selbst separabel.

1.20. Beispiel: 1) $\mathbb{K}^n$ ist mit jeder Norm separabel.

2) c_0 ist separabel (UE).

3) l^∞ ist nicht separabel (UE).

4) l^p mit $1 \leq p < \infty$ ist separabel: Bezeichne e_m die Folge mit lauter Einträgen 0, außer einem Einser in der m-ten Komponente und betrachte $A := \{e_m \mid m \in \mathbb{N}\}$. Dann ist $d := \operatorname{span} A \subseteq l^p$ gerade der Teilraum der endlichen Folgen und ein beliebig vorgelegtes $x = (x_k)_{k \in \mathbb{N}} \in l^p$ wird durch die Elemente $x^{(n)} := \sum_{m=1}^{n} x_m e_m = (x_1, \ldots, x_n, 0, 0, \ldots) \in d$ approximiert, denn

$$\lim_{n \to \infty} \left\| x - x^{(n)} \right\|_p^p = \lim_{n \to \infty} \sum_{k=n+1}^{\infty} |x_k|^p = 0.$$

5) In $(C([0,1]), \| \|_\infty)$ liegt der Teilraum aller Polynomfunktionen (eingeschränkt auf das Intervall $[0,1]$) dicht, was aus dem klassischen Approximationssatz von Weierstraß geschlossen werden kann (vgl. [Hoe20, 8.10] bzw. [Wer18, Satz I.2.11]). Da die Polynome gerade die lineare Hülle der abzählbaren Menge der Monome x^k ($k \in \mathbb{N}$) sind, folgt die Separabilität von $C([0,1])$. (Ebenso gilt dies allgemeiner für $C(K)$, wenn K eine kompakte Teilmenge von $\mathbb{K}^n$ ist.)

6) Für $1 \leq p < \infty$ ist $L^p([0,1])$ separabel: Zunächst ist, wie gegen Ende von 1.17 bemerkt, $C([0,1])$ dicht in $L^p([0,1])$, also wird jede Funktion f in $L^p([0,1])$ bzgl. $\| \ \|_p$ beliebig gut durch eine stetige Funktion g approximiert. Approximieren wir nun g gemäß 5) in der Supremumsnorm durch ein Polynom h, so erhalten wir wegen $\|f-h\|_p \leq \|f-g\|_p + \|g-h\|_p \leq \|f-g\|_p + \|g-h\|_\infty$ insgesamt eine beliebig gute Approximation von f durch h; d.h. die Polynomfunktionen (eingeschränkt auf $[0,1]$) liegen auch dicht in $L^p([0,1])$.

7) $L^\infty([0,1])$ ist nicht separabel: Für $0 \leq \delta \leq 1$ sei $f_\delta(x) := 1$ für $0 \leq x \leq \delta$ und $f_\delta(x) := 0$ sonst. Dann ist $M := \{f_\delta \mid 0 \leq \delta \leq 1\}$ eine überabzählbare Menge in $L^\infty([0,1])$ mit der Eigenschaft $\|f_\gamma - f_\delta\|_\infty = 1$ für $\delta \neq \gamma$. Es kann daher keine abzählbare Menge $A \subseteq L^\infty([0,1])$ geben, aus der wir an jedes Element von M auf $\| \ \|_\infty$-Nähe $1/3$ herankommen.

1.21. Quotienten normierter Räume nach abgeschlossenen Teilräumen: Es sei $(X, \| \ \|)$ ein normierter Raum und U ein *abgeschlossener* Teilraum von X. Der Quotientenvektorraum X/U besteht dann also aus den Klassen $x+U$ für $x \in X$ und die (wohldefinierten) Vektorraumoperationen sind $\lambda\cdot(x+U) = \lambda x + U$ und $(x+U)+(y+U) = (x+y)+U$. Wir behaupten, dass auf X/U wie folgt eine Norm eingeführt wird:

$$\|x+U\|' := \inf_{u \in U} \|x-u\|.$$

Unabhängigkeit von der Wahl des Repräsentanten einer Klasse folgt, weil $x_1 + U = x_2 + U$ einfach $x_1 = x_2 + u$ mit $u \in U$ bedeutet und damit das Infimum über den Teilraum U nicht verändert wird. Klarerweise ist $\|x+U\|' \geq 0$ und $\|0+U\|' = 0$. Ist $\|x+U\|' = 0$, dann ergibt sich durch Wahl einer das Infimum approximierenden Folge $u_n \in U$, dass $x \in \overline{U} = U$ ist, also $x + U = U = 0 + U$ gilt. Für $\lambda \in \mathbb{K}$, $\lambda \neq 0$ ist (wegen $u \in U \Leftrightarrow \lambda u \in U$)

$$\|\lambda(x+U)\|' = \|\lambda x + U\| = \inf_{u \in U} \|\lambda x - u\| = \inf_{u \in U} \|\lambda x - \lambda u\| = |\lambda| \|x+U\|';$$

für $\lambda = 0$ ist die Homogenität trivial. Schließlich ist die Dreiecksungleichung nachzuweisen. Bei gegebenen $x_1, x_2 \in X$ und beliebigem $\varepsilon > 0$ wähle $u_1, u_2 \in U$ mit $\|x_j - u_j\| < \|x_j + U\|' + \varepsilon/2$. Dann folgt

$$\begin{aligned} &\|(x_1+U)+(x_2+U)\|' = \|(x_1+x_2)+U\|' \leq \|(x_1+x_2)-(u_1+u_2)\| \\ &\leq \|x_1-u_1\| + \|x_2-u_2\| < \|x_1+U\|' + \frac{\varepsilon}{2} + \|x_2+U\|' + \frac{\varepsilon}{2} \leq \|x_1+U\|' + \|x_2+U\|' + \varepsilon \end{aligned}$$

und somit $\|(x_1+U)+(x_2+U)\|' \leq \|x_1+U\|' + \|x_2+U\|'$, weil ε beliebig war.

Bemerkung: (i) Offensichtlich gilt $\|x+U\|' \leq \|x\|$ (weil ja $0 \in U$ im Infimum mitspielt).

(ii) Die Norm $\| \ \|'$ auf X/U erzeugt gerade die Quotiententopologie bzgl. der Relation $x \sim y \Leftrightarrow x - y \in U$ und der kanonischen Surjektion $q\colon X \to X/U$, d.h. $G \subseteq X/U$ ist genau dann offen, wenn $q^{-1}(G)$ offen in X ist (siehe [Conw10, Chapter III, Theorem 4.2.(c)]). Insbesondere ist q stetig. Daraus folgt übrigens im Falle der Separabilität von X auch sofort, dass X/U separabel ist: Sei $A \subseteq X$ abzählbar und dicht, dann ist $\overline{q(A)} \supseteq q(\overline{A}) = q(X) = X/U$, also $q(A)$ eine abzählbare dichte Teilmenge in X/U.

Proposition: Wenn X ein Banachraum ist, dann auch X/U.

Beweis. Wir verwenden die Charakterisierung der Vollständigkeit in 1.6(ii) mittels absolut konvergenter Reihen. Seien $x_k \in X$ gegeben, die eine absolut konvergente Reihe in X/U bilden, d.h. $\sum_{k=1}^{\infty} \|x_k + U\|' < \infty$. Gemäß Infimumskonstruktion in der Quotientennorm gibt es zu jedem $k \in \mathbb{N}$ ein $u_k \in U$, sodass $\|x_k - u_k\| < \|x_k + U\|' + 1/2^k$ gilt; wegen $x_k + U = x_k - u_k + U$ dürfen wir OBdA gleich $\|x_k\| < \|x_k + U\|' + 1/2^k$ annehmen. Nach Summation über k folgt daraus nun $\sum_{k=1}^{\infty} \|x_k\| < \infty$. Daher existiert $x := \sum_{k=1}^{\infty} x_k$ und wir erhalten für die Partialsummen

$$\left\| x + U - \sum_{k=1}^{n} (x_k + U) \right\|' = \left\| x - \sum_{k=1}^{n} x_k + U \right\|' \leq \left\| x - \sum_{k=1}^{n} x_k \right\|,$$

was für $n \to \infty$ natürlich gegen 0 strebt. Daher ist $\sum_{k=1}^{\infty} (x_k + U)$ konvergent in X/U. $\square$

1.22. Normierte Algebren, Potenzreihen und die Neumannsche Reihe: Sei X eine $\mathbb{K}$-Algebra mit Einselement $e \in X$, d.h. X ist ein $\mathbb{K}$-Vektorraum mit einer zusätzlichen Multiplikation $X \times X \to X$, die assoziativ und bilinear ist und mit e als neutralem Element. Weiters sei auf X eine Norm $\| \ \|$ definiert, die zusätzlich folgende Eigenschaften hat:

$$\forall x, y \in X: \quad \|xy\| \leq \|x\| \|y\| \quad \text{und} \quad \|e\| = 1.$$

Wir nennen X eine *normierte Algebra* und im Falle der Vollständigkeit *Banachalgebra*. Zusätzlich zur Stetigkeit der Vektorraumoperationen ist in einer normierten Algebra auch die Multiplikationsabbildung $X \times X \to X$, $(x, y) \mapsto xy$ stetig.

Beispiele für kommutative Banachalgebren sind (jeweils mit punktweisen bzw. komponentenweisen Operationen) l^∞, $C_b(T)$ und $L^\infty(\Omega, \mu)$. Als endlichdimensionale nichtkommutative Beispiele kennen Sie aus den Einführungsvorlesungen bereits die Matrixalgebren $M_n(\mathbb{K})$ ($n \geq 2$) mit der *Operatornorm* $\|A\| := \sup\{\|Ax\|' \mid x \in \mathbb{K}^n, \|x\|' = 1\}$, wobei $\| \ \|'$ eine beliebige fest gewählte unter den äquivalenten Normen auf $\mathbb{K}^n$ ist. Unendlichdimensionale nichtkommutative Beispiele ergeben sich aus dem nächsten Kapitel.

In einer normierten Algebra X ergibt sich induktiv für jedes $x \in X$ und $k \in \mathbb{N}_0$ (wir setzen $x^0 := e$) die Ungleichung

$$\left\| x^k \right\| \leq \|x\|^k$$

und für vertauschbare Elemente $x, y \in X$, d.h. $xy = yx$, die Gültigkeit des Binomischen Lehrsatzes. Wir nehmen nun für das Weitere an, *dass X vollständig, also eine Banachalgebra, ist.* Wir erhalten z.B. für jedes $x \in X$ die bzgl. $\| \ \|$ absolut konvergente Reihe

$$\sum_{k=0}^{\infty} \frac{1}{k!} x^k,$$

die gemäß Bemerkung 1.6(ii) in X konvergent ist und somit als Limes ein eindeutiges Element in X definiert, das wir mit $\exp(x)$ bezeichnen. Mittels Cauchyprodukt und Binomischem Lehrsatz gewinnen wir für vertauschbare $x, y \in X$ dann auch die Formel

$$\exp(x + y) = \exp(x) \exp(y).$$

Ein bisschen allgemeiner können wir Potenzreihen in X betrachten, also mit Koeffizientenfolgen $\alpha_k \in \mathbb{K}$ ($k \in \mathbb{N}_0$) definierte Reihen

$$\sum_{k=0}^{\infty} \alpha_k x^k,$$

die jedenfalls für jene $x \in X$ mit $\|x\| < R := 1/(\limsup \sqrt[k]{|\alpha_k|})$ absolut in X konvergiert.

Speziell können wir für jedes $x \in X$ mit $\|x\| < 1$ die *Neumannsche Reihe* $z(x) := \sum_{k=0}^{\infty} x^k$ bilden, die also einfach die Verallgemeinerung der geometrischen Reihe ist. Durch die übliche Rechnung ergibt sich

$$(e-x)z(x) = \sum_{k=0}^{\infty} x^k - \sum_{k=0}^{\infty} x^{k+1} = e = z(x)(e-x),$$

d.h. $e - x$ ist invertierbar in X und es gilt

$$(e-x)^{-1} = \sum_{k=0}^{\infty} x^k \quad \text{sowie} \quad \left\|(e-x)^{-1}\right\| \leq \sum_{k=0}^{\infty} \|x\|^k = \frac{1}{1-\|x\|}.$$

Wir wollen daraus nun etwas allgemeiner etwas über die Menge $G(X)$ der invertierbaren Elemente in der Banachalgebra X lernen. Zunächst ist $G(X)$ eine Gruppe bzgl. der Multiplikation, wie sich leicht nachprüfen lässt (bzw. auch aus der Theorie der Ringe mit Einselement folgt). Wir werden aber auch noch topologische und analytische Eignschaften beisteuern.

Lemma: Die Gruppe $G(X)$ der invertierbaren Elemente der Banachalgebra X ist offen und die Abbildung $x \mapsto x^{-1}$ ist $\|\ \|$-stetig $G(X) \to G(X)$. Genauer: Für $x_0 \in G(X)$ und $r := 1/\left\|x_0^{-1}\right\|$ ist die offene Kugel $U_r(x_0)$ enthalten in $G(X)$ und für alle $x \in U_r(x_0)$ gilt

$$\left\|x^{-1} - x_0^{-1}\right\| \leq \frac{\|x - x_0\|}{r(r - \|x - x_0\|)}.$$

Beweis. Aus der allgemein gültigen Relation $x = x_0 - (x_0 - x) = x_0(e - x_0^{-1}(x_0 - x))$ sehen wir, dass wir für $x \in U_r(x_0)$ ein Produkt erhalten aus dem invertierbaren Element x_0 und dem Faktor $(e - x_0^{-1}(x_0 - x))$, der wegen $\left\|x_0^{-1}(x_0 - x)\right\| \leq \left\|x_0^{-1}\right\| \|x_0 - x\| = \|x_0 - x\|/r < 1$ gemäß Neumann-Reihe invertierbar ist. Also ist x invertierbar und es gilt

$$x^{-1} = (e - x_0^{-1}(x_0 - x))^{-1} x_0^{-1} = \sum_{k=0}^{\infty} \left(x_0^{-1}(x_0 - x)\right)^k x_0^{-1} = x_0^{-1} + \sum_{k=1}^{\infty} \left(x_0^{-1}(x_0 - x)\right)^k x_0^{-1},$$

woraus wir folgende Abschätzung mittels der geometrischen Reihe gewinnen:

$$\begin{aligned}\left\|x^{-1} - x_0^{-1}\right\| \leq \sum_{k=1}^{\infty} \|x_0 - x\|^k \left\|x_0^{-1}\right\|^{k+1} &= \|x_0 - x\| \left\|x_0^{-1}\right\|^2 \sum_{l=0}^{\infty} \left(\|x_0 - x\| \left\|x_0^{-1}\right\|\right)^l \\ &= \frac{\|x_0 - x\|}{r^2} \cdot \frac{1}{1 - \|x_0 - x\|/r} = \frac{\|x_0 - x\|}{r(r - \|x_0 - x\|)}.\end{aligned}$$

□

2 Lineare Operatoren und Dualräume

Wir betrachten nun lineare Abbildungen – oft sprechen wir stattdessen von Operatoren – zwischen normierten Räumen X und Y. Die Menge der *stetigen linearen Abbildungen (bzw. stetigen Operatoren)* $X \to Y$ bezeichnen wir in dieser VO mit $L(X, Y)$ und einem „gebräuchlichen Missbrauch“ folgend, werden wir die Norm in beiden Räumen X und Y nun meistens mit demselben Symbol $\| \ \|$ notieren. Ist $X = Y$ (und beide Male mit derselben Norm ausgestattet!), dann setzen wir $L(X) := L(X, X)$.

2.1. Theorem: Seien X und Y normierte Räume und $A \colon X \to Y$ linear. Dann sind folgende Aussagen äquivalent:

(i) A ist stetig bei 0,

(ii) A ist stetig,

(iii) A ist gleichmäßig stetig,

(iv) A ist *beschränkt*, d.h. es gibt ein $M \geq 0$, sodass $\|Ax\| \leq M\|x\|$ für alle $x \in X$ gilt.

Beweis. (iv) $\Rightarrow$ (iii) ist wegen $\|Ax - Ax_0\| = \|A(x - x_0)\| \leq M\|x - x_0\|$ sofort einzusehen und (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i) ist trivial. Es bleibt nur noch (i) $\Rightarrow$ (iv) zu zeigen, was wir durch Kontraposition erledigen: Ist A unbeschränkt, dann gibt es zu jedem $n \in \mathbb{N}$ ein x_n mit $\|Ax_n\| > n\|x_n\|$. Wegen $A0 = 0$ muss $x_n \neq 0$ gelten und wir können mit $y_n := x_n/(n\|x_n\|)$ zu $\|Ay_n\| \geq 1$ umformen. Wir erhalten daher $\|y_n\| = 1/n \to 0$ für $n \to \infty$ (d.h. $y_n \to 0$) und gleichzeitig $Ay_n \not\to 0 = A0$, d.h. A ist unstetig bei 0. □

Der Nulloperator ist immer stetig, also ist $L(X, Y) \neq \emptyset$ und Summen und skalare Vielfache stetiger Operatoren (wie üblich punktweise definiert) sind z.B. mittels Kriterium (i) oder (iv) im obigen Theorem auch sofort als stetig erkannt. Somit ist $L(X, Y)$ ein Vektorraum. Da beschränkte Mengen in metrischen/normierten Räumen gerade solche sind, die in (irgend)einer Kugel enthalten sind, besagt Eigenschaft (iv), dass *Bilder beschränkter Mengen* unter einem stetigen Operator *beschränkt* sind.

2.2. Definition: Sei $A \in L(X, Y)$, dann heißt das Infimum aller möglichen Schranken $M \geq 0$ in Eigenschaft (iv) von Theorem 2.1 die *Operatornorm* von A und wird mit $\|A\|$ bezeichnet. Insbesondere ist also $\|Ax\| \leq \|A\| \|x\|$.

Wenn wir für $x \neq 0$ die Ungleichung $\|Ax\| \leq M\|x\|$ umformen, sehen wir unmittelbar

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\|=1} \|Ax\| = \sup_{\|x\| \leq 1} \|Ax\|.$$

2.3. Bemerkung: (i) Wenn $D \subseteq X$ ein dichter Teilraum ist und ein Y Banachraum, dann kann $A \in L(D, Y)$ eindeutig als stetige lineare Abbildung auf X fortgesetzt werden.
(Hier folgt die Eindeutigkeit direkt aus der Dichtheit von D zusammen mit der geforderten Stetigkeit der Ausdehnung. Die Existenz einer Fortsetzung ergibt sich recht mühelos aus der gleichmäßigen Stetigkeit von A gemäß 2.1(iii): Zu $x \in X$ gibt es eine Folge (x_n) mit $x_n \in D$ und $x_n \to x$. Als konvergente Folge hat (x_n) die Cauchyfolgeneigenschaft. Letztere überträgt sich auf die Folge der gleichmäßig stetigen Bilder (Ax_n), also existiert $\widetilde{A}x := \lim Ax_n$ in Y. Die Unabhängigkeit des Wertes $\widetilde{A}x$ von der Wahl der approximierenden Folge (x_n) ergibt sich leicht mittels $\|A\|$ als gleichmäßige Schranke für Differenzen von solchen. Die Linearität von $\widetilde{A}$ is klar und für den Nachweis der Stetigkeit von $\widetilde{A}$ können wir einfach Beschränktheit mittels Stetigkeit der Norm auf X zeigen: $\|\widetilde{A}x\| = \lim \|Ax_n\| \leq \limsup \|A\|\|x_n\| = \|A\| \lim \|x_n\| = \|A\|\|x\|$.)

(ii) UE-Aufgabe: $A \in L(X, Y)$ besitzt eine auf $\mathrm{im}(A)$ definierte stetige Inverse, falls A *nach unten beschränkt* ist, d.h. es gibt ein $m > 0$, sodass $\|Ax\| \geq m\|x\|$ für alle $x \in X$.
Eine direkte Folgerung davon: A kann nicht stetig invertierbar sein, falls es eine Folge (x_n) in X gibt, sodass $\|x_n\| = 1$ und $Ax_n \to 0$.

2.4. Proposition: Die Operatornorm macht $L(X, Y)$ zum normierten Raum. Ist zudem Y vollständig, dann ist $L(X, Y)$ ein Banachraum (für jeden normierten Raum X).

Beweis. Von den Normeigenschaften ist höchstens bei der Dreiecksungleichung kurz zu argumentieren: Für $A, B \in L(X, Y)$ und $x \in X$ ist

$$\|(A + B)x\| = \|Ax + Bx\| \leq \|Ax\| + \|Bx\| \leq \|A\|\|x\| + \|B\|\|x\| = (\|A\| + \|B\|)\|x\|,$$

woraus sofort $\|A + B\| \leq \|A\| + \|B\|$ folgt.

Nun sei Y ein Banachraum und (A_n) eine Cauchyolge in $L(X, Y)$ bzgl. der Operatornorm. Für fixes beliebiges $x \in X$ ist wegen

$$\|A_n x - A_m x\| = \|(A_n - A_m)x\| \leq \|A_n - A_m\|\|x\|$$

$(A_n x)$ eine Cauchyfolge in Y, hat also einen Grenzwert $Ax := \lim A_n x \in Y$. Die so definierte Abbildung $X \to Y$, $x \mapsto Ax$ ist offensichtlich linear (weil jedes A_n und die Grenzwertbildung linear sind) und es ist noch die Stetigkeit von A sowie $\|A_n - A\| \to 0$ zu zeigen.

Sei $\varepsilon > 0$. Wähle gemäß Cauchyfolgeneigenschaft von (A_n) ein $n_0 \in \mathbb{N}$, sodass für $n, m \geq n_0$ immer $\|A_n - A_m\| < \varepsilon$ garantiert ist. Dann folgt für $n, m \geq n_0$ und jedes $x \in X$ wie oben

$$\|A_n x - A_m x\| \leq \|A_n - A_m\|\|x\| < \varepsilon\|x\|,$$

woraus durch $Ax = \lim_{m\to\infty} A_m x$ die Ungleichung

$$\forall n \geq n_0 \, \forall x \in X: \quad \|A_n x - Ax\| \leq \varepsilon \|x\|$$

wird. Insbesondere folgt gemäß Theorem 2.1(iv) die Stetigkeit von $A - A_{n_0}$ und somit auch die von $A = (A - A_{n_0}) + A_{n_0}$. Weiters besagt die obige Ungleichung natürlich $\|A_n - A\| \leq \varepsilon$ für alle $n \geq n_0$. Daher folgt $A_n \to A$. □

2.5. Proposition: Seien X, Y, Z normierte Räume, $A \in L(Y, Z)$ und $B \in L(X, Y)$. Dann ist $AB \in L(X, Z)$ und es gilt (für die entsprechenden Operatornormen) $\|AB\| \leq \|A\| \|B\|$.

Beweis. Als Verknüpfung linearer stetiger Abbildungen ist AB sebst linear und stetig. Für $x \in X$ gilt $\|ABx\| \leq \|A\| \|Bx\| \leq \|A\| \|B\| \|x\|$, woraus die Behauptung folgt. □

Obige Propositionen erlauben mit Rückblick auf 1.22 zwei sehr nützliche Folgerungen.

2.6. Korollar: Für einen Banachraum X ist $L(X)$ eine Banachalgebra mit der identischen Abbildung (notiert meistens als I oder 1) als Einselement.

2.7. Korollar: Ist X ein Banachraum und $A \in L(X)$ mit $\|A\| < 1$, dann ist $1 - A$ invertierbar und es gilt

$$(1 - A)^{-1} = \sum_{n=0}^{\infty} A^n, \quad \left\|(1 - A)^{-1}\right\| \leq \frac{1}{1 - \|A\|}.$$

Es wäre nun reine Routine auch die quantitativen Aussagen des Lemmas in 1.22 auf den Fall eines invertierbaren Operators zu übertragen. Die Gruppe der stetigen invertierbaren Operatoren auf dem Banachraum X wird in diesem Fall oft mit $GL(X)$ bezeichnet.

2.8. Beispiel (Fredholmsche Integralgleichung): Für vorgegebene Funktionen $y \in C([0, 1])$, $k \in C([0, 1] \times [0, 1])$ und $\lambda \in \mathbb{K}$ suchen wir eine Lösung $x \in C([0, 1])$ der Gleichung

$$x(s) - \lambda \int_0^1 k(s, t) x(t) \, dt = y(s) \quad (s \in [0, 1]).$$

Wir wollen dies als Operatorgleichung im Banachraum $(C([0, 1]), \| \; \|_\infty)$ interpretieren und studieren dazu zunächst die Abbildung $K\colon C([0, 1]) \to C([0, 1])$, gegeben durch

$$(Kx)(s) := \int_0^1 k(s, t) x(t) \, dt \quad (s \in [0, 1]).$$

Offensichtilch bildet K den Vektorraum $C([0,1])$ linear in sich selbst ab (vgl. Parameterintegral aus der Analysis). Die Funktion k wird *Integralkern* oder *Kernfunktion* des *Integraloperators* K genannt. Nun lässt sich die Integralgleichung abstrakt in der Form

$$x - \lambda K x = y \quad \text{bzw.} \quad (1-\lambda K)x = y$$

schreiben. Wir können leicht zeigen, dass K stetig bzg. der Norm $\| \ \|_\infty$ ist:

$$\begin{aligned}\|Kx\|_\infty &= \sup_{0\le s\le 1}\left|\int_0^1 k(s,t)x(t)\,dt\right| \le \sup_{0\le s\le 1}\int_0^1 |k(s,t)x(t)|\,dt \\ &\le \sup_{0\le s\le 1}\int_0^1 \sup_{0\le r\le 1}|k(s,r)x(r)|\,dt \le \sup_{0\le s,r\le 1}|k(s,r)|\sup_{0\le r\le 1}|x(r)| = \|k\|_\infty\|x\|_\infty.\end{aligned}$$

Insbesondere folgt auch $\|K\| \le \|k\|_\infty$ und somit $\|\lambda K\| \le |\lambda|\|k\|_\infty$. Ist z.B. $\|k\|_\infty \le M$ mit $M > 0$, dann erhalten wir für jedes $\lambda \in \mathbb{K}$ mit $|\lambda| < 1/M$ wegen $\|\lambda K\| < 1$ die Invertierbarkeit von $1 - \lambda K$ und dürfen mittels Neumann-Reihe schreiben

$$x = (1-\lambda K)^{-1}y = \sum_{n=0}^{\infty}\lambda^n K^n y = y + \sum_{n=1}^{\infty}\lambda^n K^n y,$$

d.h. es gibt eine eindeutige Lösung $x \in C([0,1])$ und diese hängt stetig von y ab.

Bemerkung: (i) Mit ein bisschen Rechnen (vgl. [Heu06, Wer18] könnten wir leicht sehen, dass sich die Potenzen K^n als Integraloperatoren mit stetigen Kernfunktionen k_n schreiben lassen und (für $|\lambda| < 1/M$) sogar die Funktionenreihe $\sum_{n=1}^\infty \lambda^n k_n$ gleichmäßig auf $[0,1] \times [0,1]$ gegen eine (stetige) Funktion h konvergiert, die als *(auf)lösender Kern* bezeichnet wird. Die Lösung x ist dann also selbst durch einen Integraloperator gegeben in der Form

$$x(s) = y(s) + \int_0^1 h(s,t)y(t)\,dt \quad (0 \le s \le 1).$$

(ii) Fredholmsche Integralgleichungen der obigen Form entstehen unter anderem als Umformulierungen von Rand- und Eigenwertproblemen für lineare Differentialoperatoren (insbesondere sogenannte Sturm-Liouville-Probleme; siehe z.B. [Heu06, §3]).

Wir wenden uns nun einigen Grundprinzipien der Funktionalanalysis zu, deren Beweise essentiell auf der *Baire-Eigenschaft vollständiger metrischer Räume* beruhen (vgl. [Hoe20, Theorem 8.5])

2.9. Die Sätze von der offenen Abbildung und von der stetigen Inversen: Wir erinnern zunächst daran, dass eine stetige bijektive Abbildung $f\colon T \to S$ zwischen topologischen Räumen genau dann ein Homömorphismus ist (also eine *stetige* Inverse f^{-1} hat), wenn f *offen* ist, d.h. die Bilder $f(O)$ offener Teilmengen $O \subseteq T$ stets offen in S sind.

Offene lineare Abbildungen zwischen normierten Räumen müssen stets surjektiv sein, weil ja z.B. das Bild der offenen Einheitskugel $U_1(0)$ eine offene Kugel um 0 enthält und somit durch Skalierungen der ganze Zielraum als Bild auftritt. Im Falle von Banachräumen gilt nun sogar eine Umkehrung davon.

Theorem (Satz von der offenen Abbildung): Seien X, Y Banachräume und $A\colon X \to Y$ linear, stetig und surjektiv. Dann ist A offen.

Beweis. Nachdem offene Kugeln Umgebungsbasen bilden und A linear ist, genügt es wegen der Stetigkeit von Skalierungen und Translationen zu zeigen, dass jedes Bild einer offenen Kugel um 0 in X eine offene Kugel um 0 in Y enthält.

1. Schritt: Sei $U := U_{2r}(0) \subseteq X$ mit $r > 0$ vorgelegt und $Q := U_r(0) \subseteq U_{2r}(0)$. Es gilt $X = \bigcup_{n\in\mathbb{N}} nQ$, daher wegen der Surjektivität von A auch $Y = \bigcup_{n\in\mathbb{N}} nA(Q) \subseteq \bigcup_{n\in\mathbb{N}} n\overline{A(Q)}$. Nach dem Satz von Baire gibt es ein $m \in \mathbb{N}$, sodass $m\overline{A(Q)}$ eine offene Kugel B enthält, also $\overline{A(Q)} \supseteq (1/m)B$. Weiters ist wegen $U \supseteq Q - Q$ auch $A(U) \supseteq A(Q) - A(Q)$ und daher

$$\begin{aligned}\overline{A(U)} &\supseteq \overline{A(Q) - A(Q)} \quad \text{[Stetigkeit der Differenz]}\\ &\supseteq \overline{A(Q)} - \overline{A(Q)} \supseteq \frac{1}{m}B - \frac{1}{m}B = \underbrace{\bigcup_{x\in\frac{1}{m}B}\left(x - \frac{1}{m}B\right)}_{\text{Vereinigung offener Mengen}} \ni 0,\end{aligned}$$

was zeigt, dass $\overline{A(U)}$ eine offene Kugel V um 0 enthalten muss.

2. Schritt: Sei $r_0 > 0$ beliebig und nun $U := U_{2r_0}(0)$. Wir wählen Zahlen $r_n > 0$ $(n \in \mathbb{N})$, sodass $\sum_{n=1}^{\infty} r_n < r_0$ gilt. Gemäß Schritt 1 können wir zu jedem $n \in \mathbb{N}_0$ eine offene Kugel V_n in Y um 0 mit Radius $\sigma_n > 0$ finden, sodass

$$\forall n \in \mathbb{N}_0 : V_n \subseteq \overline{A(U_{r_n}(0))},$$

wobei OBdA auch $\sigma_n \to 0$ gilt. Sobald die folgende Behauptung gezeigt ist, sind wir fertig.

Behauptung: $V_0 \subseteq A(U)$.

Sei $y \in V_0$. Wegen $y \in \overline{A(U_{r_0}(0))}$ gibt es also ein $x_0 \in U_{r_0}(0)$ mit $\|y - Ax_0\| < \sigma_1$, d.h. $y - Ax_0 \in V_1$. Daher folgt auch $y - Ax_0 \in \overline{A(U_{r_1}(0))}$ und wie vorhin finden wir nun ein $x_1 \in U_{r_1}(0)$ mit $y - Ax_0 - Ax_1 \in V_2$. Induktiv erhalten wir $x_n \in U_{r_n}(0)$ mit

$$(*) \qquad \|y - A(x_0 + x_1 + \ldots + x_n)\| < \sigma_{n+1} \quad (n = 0, 1, 2, \ldots).$$

Wegen $\|x_n\| < r_n$ ist $\sum_{n=0}^{\infty} \|x_n\| < r_0 + \sum_{n=1}^{\infty} r_n < 2r_0$, also existiert nach 1.6(ii) der Grenzwert $x := \sum_{n=0}^{\infty} x_n$. Es folgt $\|x\| \le \sum_{n=0}^{\infty} \|x_n\| < 2r_0$, also $x \in U_{2r_0}(0) = U$, daher mittels $(*)$ und der Stetigkeit von A auch $y = Ax$. Das heißt aber insgesamt $y \in A(U)$. $\square$

Im vorigen Beweis haben wir also im ersten Schritt die Vollständigkeit von Y verwendet, im zweiten dann jene von X. Nun folgt unmittelbar das nächste wichtige Prinzip.

Korollar (Satz von der stetigen Inversen): Seien X, Y Banachräume und $A\colon X \to Y$ linear, stetig und bijektiv. Dann ist A^{-1} stetig.

2.10. Korollar (Der Satz vom abgeschlossenen Graphen): Sei $A\colon X \to Y$ eine lineare Abbildung zwischen Banachräumen. Falls der *Graph* von A

$$\Gamma(A) := \{(x, Ax) \in X \times Y \mid x \in X\}$$

abgeschlossen in $X \times Y$ ist, dann ist A stetig.

Beweis. $\Gamma(A)$ ist als abgeschlossener Teilraum des Banachraumes[1] $X \times Y$ selbst ein Banachraum. Die Projektion $P\colon \Gamma(A) \to X$, $(x, Ax) \mapsto x$ ist linear, stetig und bijektiv, besitzt also eine stetige Inverse $P^{-1}\colon X \to \Gamma(A)$. Für jede Folge (x_n) in X mit $x_n \to x$ gilt daher $(x_n, Ax_n) = P^{-1}x_n \to P^{-1}x = (x, Ax)$, daher insbesondere auch $Ax_n \to Ax$. $\square$

2.11. Bemerkung: Graphenabgeschlossenheit von $A\colon X \to Y$ heißt mittels Folgen ausgedrückt, dass (i) $x_n \to x$ in X zusammen mit (ii) $\exists y \in Y\colon Ax_n \to y$ immer (iii) $y = Ax$ impliziert. Stetigkeit heißt gerade, dass allein aus (i) schon (ii) und (iii) folgt.

2.12. Prinzip der gleichmäßigen Beschränktheit und Banach-Steinhaus-Sätze: Hier geht es einerseits darum, aus punktweise gültigen Schranken für eine Familie von Operatoren auf globale Schranken zu schließen. Andererseits werden Kriterien für die punktweise Konvergenz von Operatorfolgen gegen einen stetigen Operator formuliert.

Theorem (Prinzip der gleichmäßigen Beschränktheit): Seien X ein Banachraum, Y ein normierter Raum, I eine Menge und $A_j \in L(X, Y)$ für jedes $j \in I$. Dann gilt:

$$\left(\forall x \in X\colon \sup_{j \in I} \|A_j x\| < \infty\right) \quad \Longrightarrow \quad \left(\sup_{j \in I} \|A_j\| < \infty\right)$$

Beweis. Für jedes $n \in \mathbb{N}$ setzen wir $X_n := \{x \in X \mid \sup_{j\in I} \|A_j x\| \leq n\}$, sodass laut Voraussetzung $X = \bigcup_{n\in\mathbb{N}} X_n$ gilt. Wegen der Stetigkeit von $x \mapsto \|A_j x\|$ für jedes $j \in I$ und $X_n = \bigcap_{j\in I}\{x \in X \mid \|A_j x\| \leq n\}$ ist jedes X_n abgeschlossen. Nach dem Satz von Baire existieren $m \in \mathbb{N}$, $x_0 \in X$ und $r > 0$, sodass $K_r(x_0) \subseteq X_m$; also gilt: $\|x - x_0\| \leq r \Rightarrow \forall j \in I\colon \|A_j x\| \leq m$.

Sei $z \in X$ beliebig mit $\|z\| \leq 1$. Dann ist $x := x_0 + rz \in K_r(x_0)$ und daher

$$\|A_j z\| = \left\| A_j \left(\frac{1}{r}(x - x_0)\right)\right\| \leq \frac{1}{r}\big(\|A_j x\| + \|A_j x_0\|\big) \leq \frac{m + m}{r} = \frac{2m}{r}.$$

Wir erhalten also $\|A_j\| = \sup_{\|z\|\leq 1} \|A_j z\| \leq 2m/r$ für alle $j \in I$. $\square$

[1] Normen auf dem Produktraum und die Vollständigkeit werden in den UE ausführlich besprochen.

Korollar A (Satz von Banach-Steinhaus): Seien X ein Banachraum, Y ein normierter Raum und $(A_n)_{n\in\mathbb{N}}$ eine Folge in $L(X,Y)$, sodass für jedes $x \in X$ der Grenzwert $Ax := \lim_{n\to\infty} A_n x \in Y$ existiert. Dann ist die Abbildung $A\colon X \to Y$, $x \mapsto Ax$ linear und stetig, die Folge der Normen $(\|A_n\|)$ ist beschränkt und es gilt $\|A\| \leq \liminf \|A_n\|$.

Beweis. Linearität von A ist klar. Außerdem ist für jedes $x \in X$ die Folge $(A_n x)$ als konvergente Folge beschränkt, also (A_n) punktweise beschränkt. Nach dem gerade bewiesenen Prinzip gibt es eine Konstante $M \geq 0$, sodass $\|A_n\| \leq M$ für alle $n \in \mathbb{N}$, d.h. $\|A_n x\| \leq M\|x\|$ für $n \in \mathbb{N}$ und $x \in X$. Mit $n \to \infty$ bei fixem x folgt daraus $\|Ax\| \leq M\|x\|$, also die Stetigkeit von A. Wegen $\|A_n x\| \leq \|A_n\|\|x\|$ und $A_n x \to Ax$ erhalten wir aber sogar $\|Ax\| \leq \liminf_{n\to\infty} \|A_n\|\|x\|$, also $\|A\| \leq \liminf \|A_n\|$. □

Korollar B (Satz von Banach-Steinhaus): Seien X, Y Banachräume und $(A_n)_{n\in\mathbb{N}}$ eine Folge in $L(X,Y)$, dann sind die folgenden beiden Aussagen äquivalent:

(i) (A_n) konvergiert punktweise gegen einen stetigen Operator $A \in L(X,Y)$,

(ii) die Folge der Normen $(\|A_n\|)$ ist beschränkt und es gibt eine dichte Teilmenge $D \subseteq X$, sodass für jedes $x \in D$ die Folge $(A_n x)_{n\in\mathbb{N}}$ konvergent ist.

Beweis. (i) $\Rightarrow$ (ii) folgt aus Korollar A.

(ii) $\Rightarrow$ (i): Sei $x \in X$ und $\varepsilon > 0$ beliebig. Wir wählen ein $z \in D$ mit $\|x - z\| < \varepsilon$ und setzen $M := \sup_{n\in\mathbb{N}} \|A_n\|$. Wegen der Konvergenz von $(A_n z)$ gibt es ein $n_0 \in \mathbb{N}$, sodass für $m, n \geq n_0$ stets $\|A_n z - A_m z\| < \varepsilon$ garantiert ist. Wir erhalten also für $m, n \geq n_0$

$$
\begin{aligned}
\|A_n x - A_m x\| &\leq \|A_n x - A_n z\| + \|A_n z - A_m z\| + \|A_m z - A_m x\| \\
&\leq M\|x - z\| + \varepsilon + M\|z - x\| < M\varepsilon + \varepsilon + M\varepsilon = (2M+1)\varepsilon.
\end{aligned}
$$

Daher ist $(A_n x)$ eine Cauchyfolge in Y und wir dürfen $Ax := \lim A_n x$ setzen. Somit ist (A_n) eine (auf ganz X) punktweise konvergente Folge von Operatoren und wir können wieder Korollar A anwenden. □

2.13. Der Dualraum eines normierten Raumes: Zu dem normierten Raum X ist $X' := L(X, \mathbb{K})$ der (stetige) *Dualraum*. Aus der Vollständigkeit des Grundkörpers $\mathbb{K}$ folgt mit Proposition 2.4, dass X' mit der durch die Operatornorm gegebenen Norm

$$
\|l\| = \sup_{x\neq 0} \frac{|l(x)|}{\|x\|} = \sup_{\|x\|=1} |l(x)| = \sup_{\|x\|\leq 1} |l(x)| \quad (l \in X')
$$

stets ein Banachraum ist. Im Vergleich zum rein algebraischen Dualraum X^* aller linearen Funktionale besteht also X' „nur“ aus den *stetigen* linearen Funktionalen $X \to \mathbb{K}$ und es gilt $X' \subseteq X^*$. Im Unendlichdimensionalen ist aber übrigens $X' \neq X^*$, d.h. es gibt dann immer unstetige lineare Funktionale auf X (allgemeiner gibt es in dem Fall für jeden normierten Raum $Y \neq \{0\}$ unstetige lineare Abbildungen $X \to Y$; vgl. z.B. [Meg98, Theorem 1.4.11]).

Für $\dim X < \infty$ wurde in der Linearen Algebra $\dim X = \dim X^*$ gezeigt und natürlich gilt hier auch $X' = X^*$, weil alle linearen Abbildungen auf einem endlichdimensionalen normierten Raum stetig sind (UE).

Wir werden einige Beispiele von Dualräumen besser durch geeignete Identifikation mit anders konstruierten Banachräumen verstehen. Dafür benötigen wir noch die folgenden Begriffe.

2.14. Definition: Seien X und Y normierte Räume.

(i) Ein stetiger linearer Operator $T \in L(X, Y)$ heißt *Isomorphismus*, falls T bijektiv und T^{-1} stetig ist.

(ii) Eine lineare Abbildung $T\colon X \to Y$ heißt *Isometrie*, falls $\|Tx\| = \|x\|$ für alle $x \in X$ gilt. Eine solche Abbildung ist automatisch stetig und injektiv.

(iii) $T\colon X \to Y$ heißt *isometrischer Isomorphismus*, falls T sowohl ein Isomorphismus als auch eine Isometrie ist. Wir schreiben in diesem Fall $X \cong Y$ (vermöge T).

2.15. Bemerkung: (i) Nach dem Satz von der stetigen Inversen ist für Banachräume X, Y bereits jede stetige bijektive lineare Abbildung $X \to Y$ ein Isomorphismus.

(ii) Sind X und Y isomorph und X vollständig, so muss auch Y ein Banachraum sein.

(iii) Eine surjektive Isometrie $T\colon X \to Y$ ist bereits ein isometrischer Isomorphismus, weil die Inverse einer Isometrie klarerweise beschränkt, also stetig, ist.

(iv) In einem unendlichdimensionalen normierten Raum Y kann ein echter abgeschlossener Teilraum $X \subseteq Y$, $X \neq Y$, trotzdem isomorph zu Y sein[2]. Ein prominentes Beispiel ist $X = c_0$ und $Y = c$ (wobei hier wiederum ein isometrischer Isomorphismus $c_0 \cong c$ unmöglich ist, wie sich durch Studium von Extremalpunkten der Einheitskugeln erweist; vgl. dazu [Schr00, §2.2, Aufgabe 12] oder [Kab18, §3.5, Aufgabe 7]).

2.16. Beispiel (Einige klassische Dualräume): 1) Wir betrachten $(l^p)'$ für $1 \leq p < \infty$. Wir wählen $q > 1$, sodass $\frac{1}{p} + \frac{1}{q} = 1$, wobei wir $q = \infty$ im Falle $p = 1$ setzen. Für $a = (a_n) \in l^q$ betrachten wir die Abbildung

$$Ta\colon l^p \to \mathbb{K}, (x_n) \mapsto \sum_{n=1}^{\infty} a_n x_n.$$

Aus der Hölder-Ungleichung folgt sofort, dass $|Ta(x)| \leq \|a\|_q \|x\|_p$ gilt, womit Ta als lineare Abbildung $l^p \to \mathbb{K}$ definiert und stetig ist; es folgt $\|Ta\| \leq \|a\|_q$ und die Linearität der Abbildung $a \mapsto Ta$ ist klar. Daher ist $T\colon l^q \to (l^p)'$ stetig und linear. Wir werden zeigen, dass T ein isometrischer Isomorphismus ist, also

$$(l^p)' \cong l^q.$$

[2] Im Falle von Hilberträumen wäre dies sogar mit isometrischer Isomorphie machbar.

Injektivität von T: Sei e_n die Folge mit Eintragung 1 an der n-ten Stelle und sonst 0. Wenn $a = (a_n) \in l^q$ mit $Ta = 0$ ist, dann folgt $a_n = Ta(e_n) = 0$ für jedes $n \in \mathbb{N}$, also $a = 0$.

Surjektivität von T: Sei $\mu \in (l^p)'$. Wir setzen $a_n := \mu(e_n)$, $a := (a_n)$ und werden $a \in l^q$ sowie $Ta = \mu$ zeigen.
Für den Fall $p > 1$ (somit $1 < q < \infty$): Es ist mit $b_n := 0$, falls $a_n = 0$, und $b_n := |a_n|^q/a_n$, falls $a_n \neq 0$, zunächst für jedes $N \in \mathbb{N}$

$$\sum_{n=1}^{N} |a_n|^q = \sum_{n=1}^{N} a_n b_n = \sum_{n=1}^{N} b_n \mu(e_n) = \mu\left(\sum_{n=1}^{N} b_n e_n\right) \leq \|\mu\| \left(\sum_{n=1}^{N} |b_n|^p\right)^{1/p}$$
$$= \|\mu\| \left(\sum_{n=1}^{N} |a_n|^{p(q-1)}\right)^{1/p} = \|\mu\| \left(\sum_{n=1}^{N} |a_n|^q\right)^{1/p},$$

daher wegen $1 - 1/p = 1/q$ nach Division einfach $\left(\sum_{n=1}^{N} |a_n|^q\right)^{1/q} \leq \|\mu\|$. Da N beliebig war, erhalten wir $a \in l^q$ mit $\|a\|_q \leq \|\mu\|$. Wegen $Ta(e_n) = a_n = \mu(e_n)$ stimmen die beiden stetigen linearen Funktionale Ta und μ auf dem (gemäß Beispiel 1.20(iv)) dichten Teilraum $\operatorname{span}\{e_n \mid n \in \mathbb{N}\}$ von l^p überein, müssen also gleich sein.
Für den Fall $p = 1$ (somit $q = \infty$): Wegen $|a_n| = |\mu(e_n)| \leq \|\mu\| \|e_n\|_\infty = \|\mu\|$ ist $a \in l^\infty$ mit $\|a\|_\infty \leq \|\mu\|$. Wiederum stimmen die Funktionale Ta und μ auf dem dichten Teilraum $\operatorname{span}\{e_n \mid n \in \mathbb{N}\}$ von l^p überein, müssen also gleich sein.

T ist Isometrie: Wir haben beim Surjektivitätsbeweis $\|a\|_q \leq \|\mu\|$ gezeigt, was wegen $\mu = Ta$ zusammen mit der früheren Beobachtung $\|Ta\| \leq \|a\|_q$ nun $\|Ta\| = \|a\|_q$ ergibt.

Wir haben nun sogar die Wahl und können die Stetigkeit der Inversen T^{-1} direkt aus der Isometrieeigenschaft ablesen oder aus dem Satz von der stetigen Inversen folgern, weil sowohl l^q als auch $(l^p)'$ Banachräume sind.

2) ${c_0}' \cong l^1$ folgt ähnlich wie in 1) mittels $T\colon l^1 \to {c_0}'$, wobei $Ta(x) := \sum_{n=1}^{\infty} a_n x_n$ für $a \in l^1$, $x \in c_0$ (Beweis als UE). Ebenso ergibt sich $c' \cong l^1$, was bemerkenswert ist, weil also ${c_0}' \cong c'$ gilt, während $c_0 \not\cong c$ ist.

3) Sei (Ω, μ) ein *σ-endlicher* Maßraum, d.h. es gibt eine Folge von messbaren Teilmengen $E_n \subseteq \Omega$ mit $\mu(E_n) < \infty$ und $\bigcup_{n \in \mathbb{N}} E_n = \Omega$. Beispiele dafür sind das Lebesgue-Maß auf beliebigen Intervallen $I \subseteq \mathbb{R}$ oder *Wahrscheinlichkeitsmaß*e (d.h. $\mu(\Omega) = 1$). Dann kann mit maßtheoretischen Versionen der Beweise aus 1) gezeigt werden (vgl. [Wer18, Satz II.2.4] oder [El18, Abschnitt VII.3] oder [Fried82, §4.14]), dass für $1 \leq p < \infty$ und $\frac{1}{p} + \frac{1}{q} = 1$ (wiederum $q = \infty$ für $p = 1$) die Abbildung $T\colon L^q(\Omega, \mu) \to L^p(\Omega, \mu)'$, gegeben durch

$$(Tg)(f) := \int_\Omega f g \, d\mu \qquad \forall g \in L^q(\Omega, \mu), \forall f \in L^p(\Omega, \mu),$$

ein isometrischer Isomorphismus ist, d.h. es gilt $L^p(\Omega, \mu)' \cong L^q(\Omega, \mu)$.

4) Warum wurde der spezielle Fall $p = \infty$ (daher $q = 1$ nötig für $\frac{1}{p} + \frac{1}{q} = 1$) in 1) bzw. 3) nicht erwähnt? Tatsächlich sind die Räume $(l^\infty)'$ bzw. $L^\infty([0,1])'$ jeweils „größer" als l^1 bzw. $L^1([0,1])$ in dem Sinne, dass wir durch Abbildungen analog zu 1) bzw. 3) jeweils Isometrien $l^1 \to (l^\infty)'$ bzw. $L^1([0,1]) \to L^\infty([0,1])'$ erhalten, die aber nicht surjektiv sind (siehe z.B. [Wer18, Satz III.1.11] konkret für l^∞). Für den Nachweis der Nichtsurjektivität ist der weiter unten behandelte Fortsetzungssatz von Hahn-Banach sehr praktisch. Außerdem ergibt sich in weiterer Folge (siehe Korollar 2.19) auch das Resultat, dass ein normierter Raum mit separablem Dualraum selbst separabel sein muss, was wegen der Separabilität von l^1 bzw. $L^1([0,1])$ und der Nichtseparabilität von l^∞ bzw. $L^\infty([0,1])$ also auch die Hoffnungen auf $l^1 \cong (l^\infty)'$ bzw. $L^1([0,1]) \cong L^\infty([0,1])'$ zerstört. (Und übrigens kann $L^1([0,1])$ gar nicht Dualraum eines normierten Raumes sein; siehe z.B. [Wer18, Korollar VIII.4.6].)

Bemerkung: Für allgemeine Maßräume lässt sich über $L^\infty(\Omega,\mu)'$ vs. $L^1(\Omega,\mu)$ nicht so viel sagen, weil schon einfache „Extremfälle" von $(\Omega,\ \mu)$ jeweils sehr verschiedene qualitative Eigenschaften hervorrufen; z.B. ist $L^\infty(\Omega,\mu)$ für eine endliche Menge Ω sicher endlichdimensional und insbesondere auch separabel, während für das Zählmaß auf einer überabzählbaren Menge Ω nichtmal $L^p(\Omega,\mu)$ mit $1 \leq p < \infty$ separabel ist.

5) Es sei K ein kompakter Hausdorff-Raum, dann lässt sich $C(K)'$ gemäß *Darstellungssatz von Riesz* ([Wer18, Theorem II.2.5] oder [Coh13, Section 7.2]) mit dem Raum $M(K)$ der *komplexen regulären Borelmaße* auf K, ausgestattet mit der *Variationsnorm*, identifizieren.

Zunächst zu den maßtheoretischen Grundbegriffen. Ein Element $\mu \in M(K)$ ist (zumindest) auf der σ-Algebra der *Borelmengen*, die von allen offenen Teilmengen von K erzeugt[3] wird, definiert und hat Werte in $\mathbb{C}$ (also ist $\pm\infty$ ausgeschlossen!); weiters ist $\mu(\emptyset) = 0$ und μ σ-additiv[4]. Man kann zeigen (vgl. [El18, Kapitel VII, §1, Aufgabe 1.7] oder [Coh13, Section 4.1]), dass für jede Borelmenge $A \subseteq K$

$$|\mu|(A) := \sup\left\{ \sum_{k=1}^{N} |\mu(A_k)| \,\middle|\, A_1,\dots,A_N \text{ disjunkte Borelmengen}, A = \bigcup_{k=1}^{N} A_k \right\} < \infty$$

ist und dadurch ein positives endliches Maß $|\mu|$ definiert wird; von diesem positiven endlichen Maß $|\mu|$ wird nun zusätzlich folgende Regularitätseigenschaft gefordert: Für jede Borelmenge A soll $|\mu|(A) = \sup\{|\mu|(C) \mid C \subseteq A, C \text{ kompakt}\} = \inf\{|\mu|(O) \mid A \subseteq O, O \text{ offen}\}$ gelten. Schließlich wird durch

$$\|\mu\| := |\mu|(K)$$

eine Norm auf $M(K)$ definiert, mit der $(M(K), \|\ \|)$ zum Banachraum wird (vgl. [Wer18, Beispiel I.1(j) und Definition I.2.14] oder [Coh13, Proposition 4.1.8]).

Der anfangs erwähnte Satz von Riesz besagt nun, dass $C(K)' \cong M(K)$ vermöge des isometrischen Isomorphismus $T\colon M(K) \to C(K)'$ gilt, wobei für ein Maß $\mu \in M(K)$ das lineare

[3] als Durchschnitt über alle die Topologie von K umfassenden σ-Algebren
[4] Für jede disjunkte Folge von Borelmengen A_k gilt $\mu(\bigcup_{k=1}^\infty A_k) = \sum_{k=1}^\infty \mu(A_k)$

Funktional $T\mu$ auf $C(K)$ durch Integration gegeben ist:

$$(T\mu)(f) := \int_K f(x)\, d\mu(x) \qquad \forall f \in C(K).$$

6) c_0 „ist kein Dualraum“, d.h. es gibt keinen normierten Raum X, sodass $c_0 \cong X'$ gilt. (Dies lässt sich aus einem Satz von Krein-Milman folgern, der in gewisser Weise besagt, dass Einheitskugeln in einem Dualraum eine reichhaltige Menge an sogenannten Extremalpunkten besitzen; siehe z.B. [Wer18, Theorem VIII.4.4 und Korollar VIII.4.6].) In dem Sinne ist auch c kein Dualraum, weil sich andernfalls relativ leicht mittels der Zerlegung $c = c_0 \oplus \operatorname{span}\{e\}$, wobei $e = (1, 1, 1, \ldots)$, zeigen ließe, dass $c_0 \cong U'$ für $U = \ker(e)$ sein muss – ein Widerspruch.

Wenn X, Y normierte Räume sind, $U \subseteq X$ ein Teilraum und $A_0\colon U \to Y$ eine stetige lineare Abbildung, so mag es zunächst nicht vermessen wirken, auf eine *stetige* Fortsetzung $A \in L(X, Y)$ mit $A|_U = A_0$ zu hoffen. Rein linear algebraisch, d.h. ohne Stetigkeitsbedingung, ist das ja immer möglich, indem wir einen Komplementärraum W zu U wählen, d.h. $X = U \oplus W$, und A auf W zum Beispiel gleich 0 setzen. Für die *stetige Fortsetzbarkeit* steckt jedoch tatsächlich die etwas tiefere Frage nach der Existenz von *abgeschlossenen* Komplementärräumen dahinter, wobei U selbst OBdA als abgeschlossen vorausgesetzt werden darf. Dann muss nach einem *abgeschlossenen* Teilraum $W \subseteq X$ gesucht werden, der $X = U \oplus W$ erfüllt. Aber das ist nicht immer möglich, z.B. erlaubt der abgeschlossene Teilraum c_0 von l^∞ keinen abgeschlossenen Komplementärraum (vgl. [Wer18, Satz IV.6.5]) und daher lässt sich auch die identische Abbildung $c_0 \to c_0$ nicht zu einem stetigen Operator $l^\infty \to c_0$ ausdehnen (vgl. die Diskussion auf Seite 181 in [Wer18]).

Erst vor dem Hintergrund des vorigen Absatzes ist die gute Nachricht der nun folgenden Hahn-Banach-Sätze über die Fortsetzbarkeit von linearen Funktionalen richtig zu würdigen.

2.17. Fortsetzungssätze von Hahn-Banach für lineare Funktionale: Wir definieren zunächst den Begriff einer *Halbnorm* auf einem Vektorraum X über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Das ist eine Abbildung $p\colon X \to \mathbb{R}$, sodass für alle $x, y \in X$ und $\forall \lambda \in \mathbb{K}$ gilt:

(i) $p(x) \geq 0$,

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

(iii) $p(x + y) \leq p(x) + p(y)$.

Satz von Hahn-Banach A: Sei X ein $\mathbb{K}$-Vektorraum mit Halbnorm p, $U \subseteq X$ ein Teilraum und $l\colon U \to \mathbb{K}$ linear mit der Eigenschaft $|l(u)| \leq p(u)$ für alle $u \in U$. Dann gibt es ein lineares Funktional $L\colon X \to \mathbb{K}$ mit den Eigenschaften

$$\forall u \in U\colon L(u) = l(u) \quad \text{und} \quad \forall x \in X\colon |L(x)| \leq p(x).$$

Beweis. 1) *Reduktion auf den reellen Fall:* Wir zeigen hier, wie das Resultat auf einen komplexen Vektorraum ausgedehnt wird, sobald es im Reellen bewiesen ist. Zur Vorbereitung bemerken wir, dass jedes $\mathbb{C}$-lineare Funktional μ auf X punktweise in den Real- und Imaginärteil zerlegt werden kann, sodass

$$\forall x \in X\colon \mu(x) = \mu_1(x) + i\mu_2(x)$$

mit $\mathbb{R}$-linearen Funktionalen μ_1 und μ_2. Es gilt dann

$$\mu_1(ix) + i\mu_2(ix) = \mu(ix) = i\mu(x) = i\mu_1(x) - \mu_2(x),$$

woraus $\mu_2(x) = -\mu_1(ix)$ folgt und daher

$$\mu(x) = \mu_1(x) - i\mu_1(ix).$$

Umgekehrt kann dadurch aus einem gegebenen $\mathbb{R}$-linearen Funktional μ_1 ein $\mathbb{C}$-lineares Funktional μ definiert werden.

Für das zu l gehörende $\mathbb{R}$-lineare Funktional l_1 auf U gilt nun die Abschätzung

$$\forall u \in U\colon |l_1(u)| \le |l(u)| \le p(u).$$

Aus dem reellen Resultat folgt nun, dass es ein $\mathbb{R}$-lineares Funktional L_1 auf X gibt, sodass

$$(*) \qquad \forall u \in U\colon L_1(u) = l_1(u) \quad \text{und} \quad \forall x \in X\colon |L_1(x)| \le p(x).$$

Wir definieren nun das $\mathbb{C}$-lineare Funktional L durch $L(x) := L_1(x) - iL_1(ix)$ für jedes $x \in X$ und erhalten sofort[5] die Fortsetzungseigenschaft

$$\forall u \in U\colon L(u) = L_1(u) - iL_1(iu) = l_1(u) - il_1(iu) = l(u).$$

Es bleibt noch, die Abschätzung $|L(x)| \le p(x)$ nachzuweisen: Bei beliebigem festen $x \in X$ verwenden wir eine Polardarstellung $L(x) = r \exp i\theta$ mit $r \ge 0$ und $0 \le \theta < 2\pi$. Dann ist $L(\exp(-i\theta)x) = \exp(-i\theta)L(x) = r \in \mathbb{R}$ und daher $L(\exp(-i\theta)x) = L_1(\exp(-i\theta)x)$. Nun können wir aus $(*)$ wie folgt schließen

$$|L(x)| = \left|e^{-i\theta}L(x)\right| = \left|L(e^{-i\theta}x)\right| = \left|L_1(e^{-i\theta}x)\right| \le p(e^{-i\theta}x) = p(x).$$

Wir nehmen also ab nun an, dass $\mathbb{K} = \mathbb{R}$ ist.

2) *„Abschütteln der Betragsstriche“:* Wir behaupten, dass für ein lineares Funktional h auf einem beliebigen Teilraum $H \subseteq X$ die Gültigkeit der Ungleichung $|h(y)| \le p(y)$ für alle $y \in H$ äquivalent ist zu

$$\forall z \in H\colon h(z) \le p(z).$$

[5] Beachte, dass U ein komplexer Teilraum ist, also insbesondere $iu \in U$ für jedes $u \in U$ gilt.

Das ist in einer Richtung trivial und umgekehrt ist nur $-h(z) = h(-z) \leq p(-z) = p(z)$ zusammen mit $h(z) \leq p(z)$ zu beachten.

Wir dürfen also nun das Statement des zu beweisenden Satzes so modifizieren, dass die Voraussetzung lautet

$$\forall u \in U\colon l(u) \leq p(u),$$

woraus die Existenz einer Fortsetzung L von l mit

$$\forall x \in X\colon L(x) \leq p(x)$$

zu schließen ist.

3) *Erweiterung um eine Dimension:* Im Falle $U = X$ ist nichts zu zeigen. Sei also $x_0 \in X \backslash U$. Im Teilraum $H := \operatorname{span}\{x_0\} \cup U$ hat jedes Element eine eindeutige Darstellung der Form $\lambda x_0 + u$ mit $\lambda \in \mathbb{R}$ und $u \in U$. Eine Fortsetzung h von l auf H muss notwendiger Weise die Gestalt

$$h(\lambda x_0 + u) = r_0 \lambda + l(u) \quad \forall \lambda \in \mathbb{R}, \forall u \in U$$

mit $r_0 \in \mathbb{R}$ haben und jedes solcher Art mit beliebigem $r_0 \in \mathbb{R}$ definierte h ist tatsächlich eine Fortsetzung von l auf H. Wir benötigen also ein $r_0 \in \mathbb{R}$ mit der Eigenschaft

$$(**) \qquad \forall \lambda \in \mathbb{R}, \forall u \in U\colon \quad r_0 \lambda + l(u) \leq p(\lambda x_0 + u).$$

Zunächst folgt daraus mit $\lambda = 1$ sofort die Relation

$$(\oplus) \qquad \forall u \in U\colon \quad r_0 \leq p(x_0 + u) - l(u)$$

und mit $\lambda = -1$ und $-u$ statt u die Ungleichung

$$(\ominus) \qquad \forall u \in U\colon \quad r_0 \geq -p(x_0 + u) - l(u).$$

Andererseits folgt aus $(\oplus)$ und $(\ominus)$ wiederum $(**)$, denn
(i) für $\lambda = 0$ entspricht $(**)$ gerade der (modifizierten) Annahme $l(u) \leq p(u)$ über l,
(ii) für $\lambda > 0$ mit u/λ statt u ergibt sich aus $(\oplus)$ sofort $\lambda r_0 \leq p(\lambda x_0 + u) - l(u)$,
(iii) für $\lambda < 0$ mit u/λ statt u lautet $(\ominus)$ wegen $|\lambda| = -\lambda$ zunächst $r_0 \geq (p(\lambda x_0 + u) - l(u))/\lambda$, was nach Multiplikation mit $\lambda < 0$ eben $\lambda r_0 \leq p(\lambda x_0 + u) - l(u)$ ergibt.

Die Aufgabe ist also, ein r_0 zu finden, das $(\oplus)$ und $(\ominus)$ erfüllt. Das ist genau dann möglich, falls

$$(***) \qquad \sup_{w \in U} (-p(x_0 + w) - l(w)) \leq \inf_{u \in U} (p(x_0 + u) - l(u))$$

gilt. Zum Beweis von $(***)$ seien $w, u \in U$ beliebig. Dann gilt

$$l(u) - l(w) = l(u - w) \leq p(u - w) = p((x_0 + u) - (x_0 + w)) \leq p(x_0 + u) + p(x_0 + w),$$

daher auch $-p(x_0 + w) - l(w) \leq p(x_0 + u) - l(u)$.

4) *„Den Rest erledigt Zorn“:* Wir betrachten die Menge $\mathcal{F}$ aller Paare (h, H_h), wobei h ein lineares Funktional auf dem Teilraum H_h mit $U \subseteq H_h \subseteq X$ ist und $h|_U = l$ sowie $h(x) \leq p(x)$ für alle $x \in H_h$ gilt. Wegen $(l, U) \in \mathcal{F}$ ist $\mathcal{F} \neq \emptyset$.

Wir führen auf $\mathcal{F}$ wie folgt eine (partielle) Ordnungsrelation ein

$$(g, H_g) \preceq (h, H_h) \quad :\Longleftrightarrow \quad H_g \subseteq H_h \text{ und } h|_{H_g} = g.$$

Wenn nun $\mathcal{K}$ eine Kette in $\mathcal{F}$ ist, dann erhalten wir durch $H := \bigcup_{(g,H_g)\in\mathcal{K}} H_g$ einen Teilraum von X (weil für $(f, H_f), (g, H_g) \in \mathcal{K}$ stets $H_f \subseteq H_g$ oder $H_g \subseteq H_f$ gilt) mit $U \subseteq H$. Wir definieren $h\colon H \to \mathbb{R}$ durch $h(x) := g(x)$, falls $x \in H_g$ mit $(g, H_g) \in \mathcal{K}$. Dies ist wegen der Ketteneigenschaft wohldefiniert, weil für $x \in H_f \cap H_g$ entweder $H_f \subseteq H_g$ mit $g|_{H_f} = f$ oder $H_g \subseteq H_f$ mit $f|_{H_g} = g$ gilt. Ganz ähnlich folgt die Linearität von h. Außerdem folgt $h(x) \leq p(x)$ für $x \in H$ laut Definition von $\mathcal{F}$ und h. Schließlich gilt laut Konstruktion $(g, H_g) \preceq (h, H)$ für alle $(g, H_g) \in \mathcal{K}$, weil natürlich $H_g \subseteq H$ und $h|_{H_g} = g$ ist.

Somit hat jede Kette in $\mathcal{F}$ eine obere Schranke, sodass nach dem Lemma von Zorn ein maximales Element $(L, H_L) \in \mathcal{F}$ existiert. Wäre $H_L \neq X$, dann könnten wir L gemäß 3) zumindest um eine Dimension unter Erhaltung der Halbnormungleichung ausdehnen, was der Maximalität von (L, H_L) widerspricht. Also muss $H_L = X$ sein und L hat bereits alle gewünschten Eigenschaften. $\square$

Satz von Hahn-Banach B: Sei X ein normierter Raum, $U \subseteq X$ ein Teilraum und $l\colon U \to \mathbb{K}$ linear und stetig. Dann gibt es ein stetiges lineares Funktional $L \in X'$ mit den Eigenschaften

$$\forall u \in U\colon L(u) = l(u) \quad \text{und} \quad \|L\| = \|l\|.$$

Beweis. Durch $p(x) := \|l\|\,\|x\|$ für $x \in X$ wird eine Halbnorm p definiert und es gilt nach Konstruktion $|l(u)| \leq p(u)$ für $u \in U$. Gemäß obigem Satz A gibt es eine Fortsetzung L von l mit $|L(x)| \leq p(x) = \|l\|\,\|x\|$ für alle $x \in X$. Daher ist L stetig und $\|L\| \leq \|l\|$. Für jedes $u \in U$ ist $|l(u)| = |L(u)| \leq \|L\|\,\|u\|$, woraus mittels Supremum über $u \in U$ mit $\|u\| \leq 1$ auch die umgekehrte Ungleichung $\|l\| \leq \|L\|$ folgt, insgesamt also $\|L\| = \|l\|$. $\square$

2.18. Unmittelbare Folgerungen aus dem Fortsetzungssatz von Hahn-Banach: Sei X ein normierter Raum und U ein Teilraum.

(i) Ist $x_0 \neq 0$, dann gibt es $l \in X'$ mit $\|l\| = 1$ und $l(x_0) = \|x_0\|$. Insbesondere ist X' *punktetrennend*, d.h. zu $x_1, x_2 \in X$ mit $x_1 \neq x_2$ gibt es $l \in X'$ mit $l(x_1) \neq l(x_2)$.

Beweis: Der zweite Teil folgt aus dem ersten mit $x_0 := x_1 - x_2$. Definiere auf dem eindimensionalen Teilraum $\operatorname{span}\{x_0\}$ das Funktional $\lambda x_0 \mapsto \lambda \|x_0\|$ und setze es gemäß Satz von Hahn-Banach normgleich auf X fort. $\square$

(ii) Für jedes $x \in X$ gilt $\|x\| = \sup\limits_{l\in X', \|l\|\leq 1} |l(x)|$. [Erinnere: $\|l\| = \sup_{\|x\|\leq 1} |l(x)|$.]

Beweis: Wegen $|l(x)| \leq \|l\|\|x\|$ gilt bereits „$\geq$" in der behaupteten Gleichung. Für $x = 0$ gilt die Behauptung sowieso, also dürfen wir für den Teil „$\leq$" gleich $x \neq 0$ annehmen. Dann gibt es nach (i) ein $l \in X'$ mit $\|l\| = 1$ und $l(x) = \|x\|$ und wir sind fertig. □

(iii) Ist U abgeschlossen und $x_0 \in X \setminus U$, dann gibt es $l \in X'$ mit $l|_U = 0$ und $l(x_0) \neq 0$.

Beweis: Wegen der Abgeschlossenheit von U ist $\delta := \inf\{\|x_0 - z\| \mid z \in U\} > 0$, weil wir andernfalls eine Folge (z_n) in U mit $z_n \to x_0$ konstruieren könnten und somit $x_0 \in U$ wäre.

Es sei $W := \operatorname{span}\{x_0\} \cup U$ und $h\colon W \to \mathbb{K}$ definiert durch $h(\lambda x_0 + u) := \delta\lambda$, wobei $\lambda x_0 + u$ die eindeutige Darstellung für Elemente in W mit $u \in U$ und $\lambda \in \mathbb{K}$ ist. Es ist $h(u) = 0$ für $u \in U$ und für $\lambda \neq 0$, $u \in U$ ist (wegen $z := -u/\lambda \in U$ und Definition von $\delta > 0$ oben)

$$|h(\lambda x_0 + u)| = \delta|\lambda| \leq \left\|x_0 + \frac{1}{\lambda}u\right\| |\lambda| = \|\lambda x_0 + u\|,$$

daher ist h stetig. Nun setzen wir h nach Hahn-Banach zu $l \in X'$ fort, dann gilt $l(u) = h(u) = 0$ für $u \in U$ und $l(x_0) = h(x_0) = \delta \neq 0$. □

(iv) U ist dicht in $X \iff (\forall l \in X'\colon l|_U = 0 \;\Rightarrow\; l = 0)$.

Beweis: $\boxed{\Rightarrow}$ Dies gilt, weil jede stetige Abbildung $X \to \mathbb{K}$ durch die Werte auf der dichten Teilmenge U eindeutig festgelegt ist.
$\boxed{\Leftarrow}$ Wenn $\overline{U} \neq X$, dann gibt es nach (iii) ein $l \in X'$ mit $l|_{\overline{U}} = 0$, aber $l \neq 0$. □

Wir ziehen eine Folgerung, die zum Beispiel belegt, dass nicht $(l^\infty)' \cong l^1$ sein kann.

2.19. Korollar: Sei X ein normierter Raum und X' separabel. Dann ist X selbst separabel

Beweis. Da sich für metrische Räume die Separabilität auf Teilmengen vererbt ist auch $S := \{l \in X' \mid \|l\| = 1\}$ separabel. Es gibt also eine abzählbare Teilmenge $\{l_1, l_2, \ldots\} \subseteq S$, die dicht in S ist. Wegen $\|l_n\| = 1$ gibt es für jedes $n \in \mathbb{N}$ ein $x_n \in X$ mit $\|x_n\| = 1$ und $|l_n(x_n)| \geq 1/2$. Wir setzen $U := \operatorname{span}\{x_1, x_2, \ldots\}$ und behaupten, dass $\overline{U} = X$ gilt, also U ein dichter Teilraum von X ist, woraus gemäß Lemma 1.19 die Separabilität von X folgt. Wir wollen 2.18(iv) anwenden. Sei indirekt angenommen es gäbe ein $l \in X'$ mit $l|_U = 0$ und $l \neq 0$. Dann ist OBdA $\|l\| = 1$ (Skalierung), also $l \in S$. Es gibt also ein $m \in \mathbb{N}$ mit $\|l - l_m\| \leq 1/4$. Dann ergibt sich wegen $l(x_m) = 0$ aber

$$\frac{1}{2} \leq |l_m(x_m)| = |l_m(x_m) - l(x_m)| = |(l_m - l)(x_m)| \leq \|l_m - l\| \|x_m\| \leq \frac{1}{4},$$

also ein Widerspruch. □

2.20. Bemerkung: Für Anwendungen auf Optimierungsaufgaben von entscheidender Bedeutung sind oftmals geometrische Versionen des Satzes von Hahn-Banach, sogenannte Trennungssätze. Eine Version für einen reellen normierten Raum X erlaubt zum Beispiel für zwei disjunkte konvexe Teilmengen V_1 und V_2 von X, von denen eine offen ist, ein $l \in X'$ zu finden, sodass

$$l(x_1) < l(x_2) \quad \forall x_1 \in V_1, \forall x_2 \in V_2$$

gilt. Eine Skizze dazu im $\mathbb{R}^2$ oder $\mathbb{R}^3$ zeigt: Mit passendem $\alpha \in \mathbb{R}$ erhalten wir eine affine Hyperebene in der Form $\{x \mid l(x) = \alpha\}$, sodass V_1 auf einer Seite davon liegt und V_2 auf der anderen. (Vgl. [Wer18, Abschnitt III.2], [Heu06, §42] [Cons16, Theorems 3.17, 3.19].)

2.21. Definition: Sei X ein normierter Raum mit Dualraum X'.

(i) Eine Folge (x_n) in X heißt *schwach konvergent* gegen $x \in X$, falls gilt

$$\forall l \in X' \colon \quad l(x_n) \to l(x) \quad (n \to \infty).$$

(ii) Eine Folge (l_n) in X' heißt *schwach* konvergent* gegen $l \in X'$, falls gilt

$$\forall x \in X \colon \quad l_n(x) \to l(x) \quad (n \to \infty).$$

2.22. Bemerkung: (i) Wegen $|l(x_n) - l(x)| \leq \|l\| \|x - x_n\|$ bzw. $|l_n(x) - l(x)| \leq \|l_n - l\| \|x\|$ sind normkonvergente Folgen immer auch schwach bzw. schwach* konvergent.

(ii) Gemäß 2.18(i) sind schwache Limiten eindeutig. Für die schwach* Konvergenz folgt das sogar direkt aus der Definition, weil es sich einfach nur um die punktweise Konvergenz von Abbildungen $X \to \mathbb{K}$ handelt.

2.23. Lemma (Satz von Banach-Alaoglu): Ist (l_n) eine beschränkte Folge im Dualraum X' eines separablen Banachraumes, dann hat sie eine schwach* konvergente Teilfolge.

Beweis. Sei $W := \operatorname{span}\{x_1, x_2, \ldots\}$ dicht in X. Die Zahlenfolge $(l_n(x_1))$ ist beschränkt und hat daher eine konvergente Teilfolge $(l_{n_1}(x_1))$. Ebenso gibt es zu $(l_{n_1}(x_2))$ wiederum eine konvergente (Teil)Teilfolge $(l_{n_2}(x_2))$ usw. Wir betrachten die „zugehörige Diagonalfolge“ (l_{n_n}) in X', die laut Konstruktion in jedem Punkt x_k ($k \in \mathbb{N}$) konvergiert und somit auch für jedes $x \in W$. Nach dem Satz von Banach-Steinhaus, Version B, ist (l_{n_n}) also punktweise konvergent auf X, d.h. schwach* konvergent. □

2.24. Die kanonische Isometrie $\iota \colon X \to X''$: Wir betrachten die kanonische Injektion ι von einem normierten Raum X in seinen Bidualraum $X'' := (X')'$, d.h. für jedes $x \in X$ ist $\iota(x) \in X'' = L(X', \mathbb{K})$ gegeben durch $l \mapsto l(x)$, also

$$\iota(x)(l) = l(x) \quad (x \in X, l \in X').$$

Die Stetigkeit von $\iota(x)$ ist wegen $|\iota(x)(l)| = |l(x)| \leq \|l\| \|x\|$ klar. Ebenso ist die Linearität von $x \mapsto \iota(x)$ routinierte Offensichtlichkeit. Gemäß 2.18(ii) gilt $\|x\| = \sup_{\|l\| \leq 1} |l(x)| = \sup_{\|l\| \leq 1} |\iota(x)(l)| = \|\iota(x)\|$, weshalb das folgende Resultat bewiesen ist.

Satz: Die kanonische Injektion ist eine Isometrie, d.h. es gilt $\|\iota(x)\| = \|x\|$.

Wir erhalten somit durch $\mathrm{im}(\iota)$ eine isometrisch isomorphe „Kopie" von X in X'', die wir durch den Abschluss in X'' vervollständigen können, falls X es nicht ist. Mit anderen Worten haben wir damit für normierte Räume eine „konkretere" Konstruktion für die Vervollständigung zur Verfügung.

Nachdem X'' stets ein Banachraum ist, kann ι für einen nicht vollständigen normierten Raum X niemals ein isometrischer Isomorphismus sein. Aber unter den Banachräumen wird dies zum Kriterium für eine besonders interessante Klasse.

2.25. Definition: Ein Banachraum X heißt *reflexiv*, wenn die kanonische Isometrie auch surjektiv ist.

2.26. Bemerkung: Für einen reflexiven Banachraum X gilt somit insbesondere $X \cong X''$. Allerdings darf die Gültigkeit dieser „Gleichung" alleine nicht mit der Definition der Reflexivität verwechselt werden, wo ja spezifischer verlangt wird, dass ι ein isometrischer Isomorphismus ist.

Aus Korollar 2.19 angewandt auf X' statt X lernen wir, dass für einen reflexiven separablen Raum auch sein Dualraum separabel sein muss ($X \cong X''$ ist dual zu X'), also gilt insgesamt die folgende Aussage.

2.27. Proposition: Ein reflexiver Banachraum X ist genau dann separabel, wenn sein Dualraum X' es ist.

2.28. Beispiel: 1) Endlichdimensionale normierte Räume sind natürlich reflexiv.

2) Wir werden später sehen, dass jeder Hilbertraum reflexiv ist.

3) Für $1 < p < \infty$ ist l^p reflexiv: Wir bezeichnen mit $R\colon l^q \to (l^p)'$ und $S\colon l^p \to (l^q)'$ die in 2.16.1) konstruierten isometrischen Isomorphismen, wobei $\frac{1}{p} + \frac{1}{q} = 1$ ist. Insbesondere tritt jedes $f \in (l^p)'$ in der Form $f = Ry$ mit einem eindeutigen $y \in l^q$ auf. Wir berechnen also für $x \in l^p$ nun $\iota(x) \in (l^p)''$ durch

$$\iota(x)(f) = f(x) = (Ry)(x) = \sum_{n=1}^{\infty} y_n x_n = (Sx)(y) = (Sx)(R^{-1}f),$$

daher ist $\iota(x) = (Sx) \circ R^{-1}$. Wenn nun $\zeta \in (l^p)'' = L((l^p)', \mathbb{K})$ vorgegeben ist, dann ist $\zeta \circ R$ eine stetige lineare Abbildung $l^q \to \mathbb{K}$, d.h. $\zeta \circ R \in (l^q)'$ und es gibt ein (eindeutiges) $x \in l^p$ mit $\zeta \circ R = Sx$. Dann ist aber $\zeta = (Sx) \circ R^{-1} = \iota(x)$. Somit ist ι surjektiv.

4) Für $1 < p < \infty$ ist $L^p(\Omega, \mu)$ reflexiv. Das kann im Prinzip so ähnlich wie in 3) gezeigt werden (siehe z.B. [Fried82, Corollary 4.14.5]). Alternativ gibt es Beweise über die Eigenschaft der sogenannten gleichmäßigen Konvexität (siehe z.B. [Kab18, Schr00]).

5) *Einige wichtige Banachräume, die nicht reflexiv sind:* Da ein reflexiver Raum insbesondere Dualraum eines normierten Raumes sein muss, wissen wir schon aus 2.16.6), dass

weder c_0 noch c reflexiv sein können. Weiters haben wir oben gesehen, dass für einen reflexiven separablen Raum auch sein Dualraum separabel sein muss. Somit kann weder l^1 noch $L^1([0,1])$ reflexiv sein, weil sie separable Räume sind, während ihre Dualräume l^∞ und $L^\infty([0,1])$ es nicht sind. Ebenso kann der separable Banachraum $C([0,1])$ nicht reflexiv sein, weil sein Dualraum $M([0,1])$ nicht separabel ist (z.B. haben in der überabzählbaren Teilmenge von Dirac-Maßen $\{\delta_{t_0} \mid t_0 \in [0,1]\} \subseteq M([0,1])$ je zwei verschiedene Elemente Abstand 2 voneinander). Schließlich ist weder l^∞ noch $L^\infty([0,1])$ reflexiv, wie wir mittels 2.30 einsehen werden (weil l^1 bzw. $L^1([0,1])$ nicht reflexiv sind).

2.29. Proposition: Abgeschlossene Teilräume reflexiver Räume sind reflexiv.

Beweis. Sei X reflexiv und U ein abgeschlossener Teilraum von X. Wir müssen zeigen, dass die kanonische Isometrie $\iota_U : U \to U''$ surjektiv ist. Sei $\mu \in U''$ vorgegeben. Wir gewinnen daraus eine lineare Abbildung $\widetilde{\mu} : X' \to \mathbb{K}$, $l \mapsto \mu(l|_U)$, die stetig ist, denn

$$|\widetilde{\mu}(l)| = |\mu(l|_U)| \le \|\mu\| \, \|l|_U\| \le \|\mu\| \, \|l\|.$$

Also ist $\widetilde{\mu} \in X''$ und wegen der Reflexivität von X gibt es ein kanonisches Urbild $x \in X$ von $\widetilde{\mu}$ unter $\iota : X \to X''$, d.h.

$$\forall l \in X' : \quad l(x) = \mu(l|_U).$$

Es muss $x \in U$ sein, denn andernfalls gäbe es wegen der Abgeschlossenheit von U nach Hahn-Banach, genauer 2.18(iii), ein $l \in X'$ mit $l|_U = 0$ und $l(x) \neq 0$, was im Widerspruch zu obiger Gleichung ist.

Schließlich zeigen wir noch $\iota_U(x) = \mu$: Für beliebig gegebenes $l_0 \in U'$ sei $l \in X'$ eine Hahn-Banach-Fortsetzung von l_0, dann gilt

$$\mu(l_0) = \mu(l|_U) = l(x) = l_0(x) = \iota_U(x)(l_0).$$

□

2.30. Theorem: Sei X ein Banachraum, dann gilt: X reflexiv $\iff$ X' reflexiv.

Beweis. In diesem Beweis bezeichnen wir mit ι nur die kanonische Isometrie $X \to X''$.

$\boxed{\Rightarrow}$ Sei $L \in X'''$. Dann ist die lineare Abbildung $l := L \circ \iota$ stetig $X \to \mathbb{K}$, also gilt $l \in X'$. Wir behaupten, dass L gerade das Bild von l unter der kanonischen Isometrie $X' \to X'''$ ist, d.h. dass $L(\mu) = \mu(l)$ gilt für alle $\mu \in X''$. Das folgt aber nun leicht aus der Reflexivität von X, weil jedes $\mu \in X''$ eindeutig als $\iota(y)$ mit $y \in X$ auftritt und somit

$$L(\mu) = L(\iota(y)) = (L \circ \iota)(y) = l(y) = \iota(y)(l) = \mu(l)$$

ergibt.

$\boxed{\Leftarrow}$ Vorbemerkung: Die Vollständigkeit von X überträgt sich auf das isometrisch isomorphe Bild $\iota(X) \subseteq X''$, weshalb $\iota(X)$ sicher ein abgeschlossener Teilraum von X'' ist.

Coole Beweisvariante: Wir dürfen aus dem Beweisteil $\boxed{\Rightarrow}$ bereits schließen, dass X'' reflexiv ist. Nach Proposition 2.29 ist also $\iota(X)$ reflexiv. Nun folgt die Reflexivität von X sofort, falls wir uns auf die Tatsache berufen, dass Isomorphie die Reflexivität überträgt (siehe z.B. [Meg98, Proposition 1.11.8]). (Nach [Meg98, Proposition 1.11.22] genügt sogar eine stetige lineare Surjektion $E \to F$, um aus der Reflexivität von E auf jene von F zu schließen.)

Eine auch nicht uncoole Beweisvariante: Nehmen wir indirekt an, dass $\iota(X) \neq X''$ gilt. Dann gibt es nach Hahn-Banach ein $L \in X'''$, $L \neq 0$, das aber auf dem abgeschlossenen Teilraum $\iota(X)$ verschwindet. Wegen der Reflexivität von X' gibt es ein eindeutiges kanonisches Urbild $l \in X'$ zu L, d.h.

$$\forall \mu \in X'' : \quad L(\mu) = \mu(l).$$

Für die speziellen Elemente $\mu = \iota(x) \in X''$ mit $x \in X$ bedeutet dies aber

$$\forall x \in X : l(x) = \iota(x)(l) = L(\iota(x)) = 0.$$

Daraus folgt $l = 0$ und somit weiter $L = 0$, was einen Widerspruch erzeugt. □

2.31. Theorem: In einem reflexiven Banachraum besitzt jede beschränkte Folge eine schwach konvergente Teilfolge.

Beweis. Sei X ein reflexiver Banachraum und die Folge (x_n) in X beschränkt. Der abgeschlossene Teilraum $U := \overline{\operatorname{span}\{x_1, x_2, \ldots\}}$ ist nach 2.29 reflexiv und offensichtlich separabel. Wegen $U \cong U''$ ist U'' ebenfalls separabel und weiters ist nach 2.19 auch U' separabel. Wir können (x_n) vermöge $\iota \colon U \to U''$ als beschränkte Folge in $U'' = (U')'$ auffassen und den Satz von Banach-Alaoglu 2.23 anwenden. Demnach gibt es eine Teilfolge (x_{n_k}) und ein $x \in U$, sodass $l_0(x_{n_k}) = \iota(x_{n_k})(l_0) \to \iota(x)(l_0) = l_0(x)$ gilt für alle $l_0 \in U'$.
Sei nun $l \in X'$ beliebig und $l_0 = l|_U \in U'$, dann folgt

$$l(x_{n_k}) = l_0(x_{n_k}) \to l_0(x) = l(x).$$

Weil $l \in X'$ beliebig war, folgt daraus die schwache Konvergenz von (x_{n_k}) gegen x. □

2.32. Adjungierte Operatoren: Hier seien stets X, Y normierte Räume.

Definition: Sei $A \in L(X, Y)$, dann ist der zu A *adjungierte Operator* $A' \colon Y' \to X'$ gegeben durch $y' \mapsto y' \circ A$ für jedes $y' \in Y'$, d.h. es ist $(A'y')(x) = y'(Ax)$ für alle $x \in X$.

(Linearität von $x \mapsto A'y'(x)$ und von $y' \mapsto A'y'$ ist klar und aus $|y'(Ax)| \leq \|y'\| \|A\| \|x\|$ folgt auch jeweils die Stetigkeit.)

Beispiel: Für $1 \leq p < \infty$ betrachten wir den Linksshift $L \colon l^p \to l^p$ mit

$$L(x_1, x_2, x_3, \ldots) := (x_2, x_3, \ldots) \quad \forall x = (x_n) \in l^p$$

und bestimmen die Adjungierte L', die wir in diesem Fall mittels $(l^p)' \cong l^q$ (mit $1/q = 1 - 1/p$) als Operator $l^q \to l^q$ auffassen können. Wir berechnen also für $y = (y_n) \in l^q$ und $x = (x_n) \in l^p$ einfach [mit dem expliziten isometrischen Isomorphismus $T\colon l^q \to (l^p)'$ aus 2.16.1)]

$$\begin{aligned}(L'y)(x) = Ty(Lx) = \sum_{n=1}^{\infty} y_n x_{n+1} = 0x_1 + \sum_{k=2}^{\infty} y_{k-1} x_k \\ = \sum_{k=1}^{\infty} x_k z_k = Tz(x) \quad \text{mit } z_k := \begin{cases} 0, & k = 1, \\ y_{k-1}, & k \geq 2. \end{cases}\end{aligned}$$

Wir erhalten somit $L' = R$ mit dem Rechtsshift $R\colon l^q \to l^q$,

$$R(y_1, y_2, \ldots) := (0, y_1, y_2, \ldots) \quad \forall y = (y_n) \in l^q.$$

Bemerkung: (i) Die Abbildung $A \mapsto A'$ ist eine Isometrie $L(X,Y) \to L(Y', X')$, wobei die Ungleichung $\|A'\| \leq \|A\|$ unmittelbar aus der Definition folgt und für den Nachweis der Gleichheit dann am besten die Formel aus 2.18(ii) herangezogen wird.

(ii) Durch Routinerechnung ergibt sich $(AB)' = B' \cdot A'$ für $A \in L(Y,Z)$, $B \in L(X,Y)$.

Satz: Sei $A \in L(X,Y)$, dann ist der Abschluss des Bildes von A in Y gleich dem *Annihilator* von $\ker(A')$, d.h.

$$\overline{\operatorname{im}(A)} = \ker(A')_\perp := \{y \in Y \mid \forall y' \in \ker(A')\colon y'(y) = 0\}.$$

Beweis. Der Teilraum $U := \overline{\operatorname{im}(A)}$ ist abgeschlossen in Y und $\ker(A')_\perp$ ist als Durchschnitt über die abgeschlossenen Mengen $\ker(y')$ für $y' \in \ker(A')$ ebenfalls abgeschlossen.

Wir zeigen zuerst $U \supseteq \ker(A')_\perp$, indem wir $Y \setminus U \subseteq Y \setminus \ker(A')_\perp$ nachweisen: Zu $y \in Y \setminus U$ gibt es nach Hahn-Banach ein $y' \in Y'$ mit $y'|_U = 0$ und $y'(y) \neq 0$. Daher gilt für alle $x \in X$ dann $0 = y'(Ax) = (A'y')(x)$, d.h. $A'y' = 0$ in X', daher $y' \in \ker(A')$. Wegen $y'(y) \neq 0$ ist also $y \notin \ker(A')_\perp$.

Nun zeigen wir noch $U \subseteq \ker(A')_\perp$: Es genügt, $\operatorname{im}(A) \subseteq \ker(A')_\perp$ zu zeigen, weil wie oben bemerkt $\ker(A')_\perp$ abgeschlossen ist. Sei nun $y \in \operatorname{im}(A)$ beliebig. Es gibt ein $x \in X$ mit $y = Ax$ und daher folgt für jedes $y' \in \ker(A')$ einfach $0 = (A'y')(x) = y'(Ax) = y'(y)$, also $y \in \ker(A')_\perp$. □

Aus dem obigen Satz können wir als unmittelbare Folgerung ein häufig praktikables Kriterium für die Lösbarkeit von inhomogenen Operatorgleichungen $Ax = y$ gewinnen, weil sich Kerne von Operatoren in der Regel besser (und teilweise sogar explizit) bestimmen lassen.

Korollar: Sei $A \in L(X,Y)$ mit abgeschlossenem Bild, d.h. $\operatorname{im}(A)$ ist ein abgeschlossener Teilraum von Y. Dann sind die folgenden beiden Aussagen für jedes $y \in Y$ äquivalent:

(i) Die Gleichung $Ax = y$ ist lösbar für (mindestens) ein $x \in X$,

(ii) $\forall y' \in Y'\colon \quad A'y' = 0 \quad \Rightarrow \quad y'(y) = 0.$

Wir verallgemeinern nun einige Aspekte der Theorie der Eigenwerte für lineare Abbildungen auf endlichdimensionalen Vektorräumen auf die Situation von

stetigen Endomorphismen $A \in L(X)$ auf einem *Banachraum* X

(insbesondere haben wir weiterhin als Grundkörper $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$).

2.33. Resolvente und Resolventenabbildung:

Definition: Sei X ein Banachraum und $A \in L(X)$. Die *Resolventenmenge* von A ist

$$\varrho(A) := \{\lambda \in \mathbb{K} \mid \exists (\lambda - A)^{-1} \in L(X)\}$$

und $\varrho(A) \to L(X), \lambda \mapsto (\lambda - A)^{-1} =: R_\lambda(A)$ heißt *Resolventenabbildung*.

Beachte: Wegen des Satzes von der stetigen Inversen gilt für eine Zahl $\lambda \in \mathbb{K}$ bereits genau dann $\lambda \in \varrho(A)$, wenn $\lambda - A$ bijektiv ist.

Weil $L(X)$ eine Banachalgebra ist, erhalten wir aus dem Lemma in 1.22 und der Beobachtung $\|(\lambda - A) - (\mu - A)\| = |\lambda - \mu|$, dass für $\lambda \in \varrho(A)$ stets auch die Menge aller $\mu \in \mathbb{K}$ mit $|\lambda - \mu| < 1/\|R_\lambda(A)\|$ zu $\varrho(A)$ gehört. Somit ist das folgende Resultat bewiesen.

Lemma: $\varrho(A)$ ist eine offene Teilmenge von $\mathbb{K}$.

Bemerkung: Mit Hilfe der Neumann-Reihe lässt sich auch ganz leicht zeigen, dass die Resolventenabbildung analytisch ist in dem Sinne, dass sich die Resolvente lokal immer durch eine Potenzreihe bzgl. λ mit Koeffizienten in $L(X)$ entwickeln lässt.

Während für endlichdimensionales X und $A \in L(X)$ jedes $\lambda \in \mathbb{K} \setminus \varrho(A)$ zwingend ein Eigenwert von A sein muss, ist das im Unendlichdimensionalen sogar sehr häufig nicht der Fall. Ein ganz einfaches Beispiel erhalten wir durch den Rechtsshift $R(x_1, x_2, \ldots) = (0, x_1, x_2, \ldots)$ auf l^p mit $\lambda = 0$: Hier ist $(0 - R)x = (0, -x_1, -x_2, \ldots)$ offensichtlich nicht bijektiv, daher sicher $0 \in \mathbb{K} \setminus \varrho(R)$, aber wegen $\ker(0 - R) = \{0\}$ gibt es auch keinen Eigenvektor $x \neq 0$ zu 0, weil das gerade $Rx = 0$ verlangen würde.

2.34. Spektrum: Sei X ein Banachraum und $A \in L(X)$, dann heißt $\sigma(A) := \mathbb{K} \setminus \varrho(A)$ das *Spektrum* von A.

Satz: $\sigma(A)$ ist kompakt und es gilt $|\lambda| \leq \|A\|$ für jedes $\lambda \in \sigma(A)$.

Beweis. Als Komplement einer offenen Menge ist $\sigma(A)$ natürlich abgeschlossen. Für $\lambda \in \mathbb{K}$ mit $|\lambda| > \|A\|$ gilt $\|A/\lambda\| < 1$, daher ist nach 2.7 der Operator $1 - A/\lambda$ (stetig) invertierbar und somit auch $\lambda - A = \lambda(1 - A/\lambda)$. Es folgt also wegen $\sigma(A) \subseteq \{\lambda \in \mathbb{K} \mid |\lambda| \leq \|A\|\}$ die Beschränktheit von $\sigma(A)$. Nach dem Satz von Heine-Borel ist $\sigma(A)$ kompakt. $\square$

Bemerkung: Im Falle $\mathbb{K} = \mathbb{C}$ ist $\sigma(A) \neq \emptyset$ garantiert. Beweise dafür nützen typischer Weise die Analytizität der Resolventenabbildung und den Satz von Liouville aus der Komplexen Analysis (vgl. alle Funktionalanalysisbücher in der Literaturliste). Im Falle $\mathbb{K} = \mathbb{R}$ wissen wir bereits aus der endlichdimensionalen Linearen Algebra, dass $\sigma(A) = \emptyset$ passieren kann.

Wegen des Satzes von der stetigen Inversen ist $\lambda \in \sigma(A)$ gleichwertig damit, dass $\lambda - A$ nicht bijektiv ist. Wir erhalten die folgende **disjunkte Zerlegung des Spektrums**

$$\sigma(A) = \sigma_p(A) \cup \sigma_c(A) \cup \sigma_r(A)$$

mit dem *Punktspektrum* bzw. der Menge der *Eigenwerte*

$$\sigma_p(A) := \{\lambda \in \mathbb{K} \mid \lambda - A \text{ nicht injektiv}\},$$

dem *kontinuierlichem Spektrum*

$$\sigma_c(A) := \{\lambda \in \mathbb{K} \mid \lambda - A \text{ injektiv, nicht surjektiv und } \operatorname{im}(\lambda - A) \text{ dicht}\}$$

und dem *Residualspektrum*

$$\sigma_r(A) := \{\lambda \in \mathbb{K} \mid \lambda - A \text{ injektiv und } \operatorname{im}(\lambda - A) \text{ nicht dicht}\}.$$

2.35. Kompakte Operatoren: Es seien X, Y normierte Räume.

Definition: Eine lineare Abbildung $A\colon X \to Y$ heißt *kompakt*, falls $A(K_1(0))$ (Bild der abgeschlossenen Einheitskugel) kompakten Abschluss in Y hat, also relativkompakt ist.

Nachdem kompakte Mengen stets beschränkt sind, ist jeder kompakte Operator auch stetig. Die Menge aller kompakten Operatoren $X \to Y$ bildet einen Teilraum $K(X, Y)$ von $L(X, Y)$. Im Falle $X = Y$ schreiben wir auch $K(X)$.

Äquivalent zur Kompaktheit von $A\colon X \to Y$ sind klarerweise auch folgende Bedingungen:

(i) A bildet beschränkte Mengen auf relativkompakte Mengen ab,

(ii) ist (x_n) eine beschränkte Folge in X, dann besitzt die Bildfolge (Ax_n) in Y eine konvergente Teilfolge.

Es lässt sich mühelos z.B. mittels Folgen zeigen, dass für $A \in L(Y, Z)$, $B \in L(X, Y)$ stets die Kompaktheit von AB folgt, falls A oder B kompakt ist. Insbesondere ist demnach $K(X)$ immer ein Ideal in $L(X)$.

Lemma: Sei nun Y ein Banachraum. Falls $A \in L(X, Y)$ der Limes (bzgl. Operatornorm) einer Folge (A_n) von Operatoren mit endlichdimensionem Bild ist, d.h. für jedes $n \in \mathbb{N}$ ist $\dim \operatorname{im}(A_n) < \infty$, dann ist A kompakt.

Beweis. Wir werden zeigen, dass $A(K_1(0))$ totalbeschränkt (oder präkompakt) ist, dann ist der Abschluss $\overline{A(K_1(0))}$ kompakt (weil vollständig in Y!) gemäß einer bekannten Charakterisierung kompakter Mengen in metrischen Räumen (siehe z.B. [Hoe20, 8.8]).

Sei $\varepsilon > 0$ vorgegeben. Wähle $N \in \mathbb{N}$, sodass $\|A - A_N\| < \varepsilon/3$. Weil $A_N(K_1(0))$ im endlichdimensionalen normierten Raum $\mathrm{im}(A_N)$ totalbeschränkt ist, gibt es $x_1, \ldots, x_m \in K_1(0) \subseteq X$, sodass die Menge $M_\varepsilon := \{A_N x_1, \ldots A_N x_m\}$ folgende Eigenschaft hat: Zu jedem $x \in K_1(0)$, gibt es ein x_l, sodass $\|A_N x - A_N x_l\| < \varepsilon/3$ gilt. Somit erhalten wir insgesamt

$$\begin{aligned}\|Ax - Ax_l\| &\le \|Ax - A_N x\| + \|A_N x - A_N x_l\| + \|A_N x_l - Ax_l\| \\ &\le \|A - A_N\|\|x\| + \frac{\varepsilon}{3} + \|A_N - A\|\|x_l\| < \frac{\varepsilon}{3}\cdot 1 + \frac{\varepsilon}{3} + \frac{\varepsilon}{3}\cdot 1 = \varepsilon.\end{aligned}$$

Daher ist $A(K_1(0))$ durch ε-Kugeln mit den Mittelpunkten $Ax_1, \ldots, Ax_m$ überdeckt. Weil $\varepsilon > 0$ beliebig war, ist somit $A(K_1(0))$ als totalbeschränkt erkannt. □

Beispiele: 1) Operatoren mit endlichdimensionalem Bild sind kompakt.

2) Die Identität $X \to X$ ist genau dann kompakt, wenn X endlichdimensional ist.

3) Sei $a \in c_0$ fest, $1 \le p \le \infty$ und $A \in L(l^p)$ gegeben durch $Ax = (a_n x_n)$ für $x = (x_n) \in l^p$. Dann ist A kompakt, weil bzgl. Operatornorm approximierbar durch die Folge (A_k) von Operatoren $A_k \in L(l^p)$ mit $A_k x = (a_1 x_1, \ldots, a_k x_k, 0, 0, \ldots)$, für die offensichtlich $\mathrm{im}(A_k) \subseteq \mathrm{span}\{e_1, \ldots, e_k\}$ gilt: Für $x \in l^p$ mit $\|x\|_p \le 1$ gilt nämlich

$$\|Ax - A_k x\|_p = \|(0, \ldots, 0, a_{k+1}x_{k+1}, a_{k+2}x_{k+2}, \ldots)\|_p \le \sup_{n>k} |a_n| \cdot \|x\|_p \le \sup_{n>k} |a_n|,$$

daher ist $\|A - A_k\| = \sup_{\|x\|_p \le 1} \|Ax - A_k x\|_p \le \sup_{n>k} |a_n| \to 0$ für $k \to \infty$.

Bemerkung: Viele Integraloperatoren auf $C([0,1])$ oder auf L^p-Räumen sind kompakt. Benötigt werden für den Nachweis jeweils praktikable hinreichende Bedingungen für die Kompaktheit von Teilmengen in diesen Funktionenräumen. Für $L^p(\mathbb{R})$ mit $1 \le p < \infty$ vgl. etwa [Wer18, Satz II.3.5] und für $C([0,1])$ beruht dies häufig auf dem **Satz von Arzelá-Ascoli** (vgl. [Wer18, Satz II.3.4] oder [Hoe20, 8.10]), demgemäß z.B. $M \subseteq C([0,1])$ relativkompakt ist, falls M beschränkt ist und *gleichgradig stetig*, d.h.

$$\forall \varepsilon > 0\, \exists \delta > 0\, \forall f \in M\colon |s - t| < \delta \Rightarrow |f(s) - f(t)| < \varepsilon.$$

Eine Charakterisierung der relativkompakten Mengen in l^p findet sich übrigens in [Wer18, Aufgabe II.5.24]. Damit könnte alternativ auch das obige Beispiel 3) geklärt werden.

Proposition: Seien X, Y Banachräume, $A\colon X \to Y$ linear und folgende Aussage vorgelegt:

$(*)$ Ist (x_n) in X schwach konvergent gegen x,

dann ist (Ax_n) in Y normkonvergent gegen Ax.

Dann gilt:

(i) Jedes $A \in K(X, Y)$ erfüllt $(*)$.

(ii) Ist X reflexiv und gilt $(*)$ für $A \in L(X, Y)$, dann ist A kompakt.

Die Bedingung $(*)$ wird auch als *Vollstetigkeit* bezeichnet. Beweis als UE (mit Hinweisen).

2.36. Einige Lücken in unserem Stoff in Kauf nehmen: Wir beschreiben hier einige allgemein sehr brauchbare Resultate, deren Beweise wir aber im Rahmen der VO nicht ausführen. Außerdem werden diese auch in unserer Skizze der Rieszschen Theorie kompakter Operatoren weiter unten eine Rolle spielen.

(i) Lemma von Riesz (wurde in UE behandelt): Wenn U ein echter und abgeschlossener Teilraum des normierten Raumes E ist, dann gibt es zu jedem $\eta \in \mathbb{R}$ mit $0 < \eta < 1$ ein $x_\eta \in E$ mit $\|x_\eta\| = 1$, sodass

$$\forall u \in U: \quad \|u - x_\eta\| \geq \eta.$$

(ii) Satz von Schauder (siehe [Wer18, Satz III.4.4]): Seien X, Y Banachräume, dann gilt

$$A \in L(X, Y) \text{ kompakt} \Leftrightarrow A' \in L(Y', X') \text{ kompakt}.$$

(Der Beweis zeigt eine weitere Anwendung des oben erwähnten Satzes von Arzelá-Ascoli.)

(iii) Sei X ein Banachraum und $A \in L(X)$, dann gilt $\sigma(A) = \sigma(A')$.
(Für einen Beweis siehe [Wer18, Satz VI.1.2].)

(iv) Sei X ein normierter Raum mit abgeschlossenen Teilräumen U, V, sodass $X = U \oplus V$ gilt. Die direkte Summenzerlegung heißt *topologisch*, falls X isomorph ist zum Banachraum $U \times V$ mit der Norm $\|(u, v)\|_1 = \|u\| + \|v\|$ (bzw. einer der auf $U \times V$ äquivalenten Normen $\|(u, v)\|_p$ wie in den UE besprochen). Wenn X ein *Banachraum* ist, dann gelten folgende Aussagen ([Wer18, Satz IV.6.3]):
(a) $X = U \oplus V$ ist topologisch,
(b) es existiert eine stetige Projektion P auf U, d.h. $P \in L(X)$ mit $P^2 = P$ und $\operatorname{im} P = U$,
(c) es ist X/U isomorph zu V.

Bemerkung: Falls U ein endlichdimensionaler Teilraum eines normierten Raumes X ist, dann gibt es immer eine stetige Projektion P auf U ([Wer18, Satz IV.6.2]) und $\ker P$ ist dann ein abgeschlossener Komplementärraum ([Wer18, Lemma IV.6.1]).

(v) Verwandt zum Satz, den wir in 2.32 bewiesen haben, ist der sogenannte Satz vom abgeschlossenen Bild ([Wer18, Satz IV.5.1]): Seien X, Y Banachräume und $A \in L(X, Y)$, dann sind folgende Aussagen äquivalent:
(a) $\operatorname{im} A$ ist abgeschlossen in Y,
(b) $\operatorname{im} A = \ker(A')_\perp := \{y \in Y \mid \forall y' \in \ker(A') \colon y'(y) = 0\}$,
(c) $\operatorname{im}(A')$ ist abgeschlossen in X',
(d) $\operatorname{im}(A') = \ker(A)^\perp := \{x' \in X' \mid \forall x \in \ker(A) \colon x'(x) = 0\}$.

(vi) Sei X ein normierter Raum, U ein abgeschlossener Teilraum und $U^\perp := \{x' \in X' \mid \forall u \in U \colon x'(u) = 0\}$. Dann gilt $U' \cong X'/U^\perp$ und $(X/U)' \cong U^\perp$. (Vgl. [Wer18, Satz III.1.10].)

Ab nun sei X stets ein Banachraum. Hier ist nun ein erstes Resultat über das Spektrum $\sigma(A)$ eines kompakten Operators A auf dem Banachraum X. Viele weitere wichtige Eigenschaften von $\sigma(A)$ folgen unten.

2.37. Proposition: Ist $A \in K(X)$ und X unendlichdimensional, dann gilt $0 \in \sigma(A)$.

[Aber 0 ist nicht notwendig ein Eigenwert von A.]

Beweis. Wenn $0 \in \varrho(A)$ gilt, dann ist A stetig invertierbar. Wie in 2.35 (vor dem Lemma) festgestellt ist dann AA^{-1} kompakt, also die Identität auf X kompakt, was nach Beispiel 2) in 2.35 $\dim X < \infty$ nach sich zieht. □

Der obige Satz betraf den Spektralwert 0 für einen kompakten Operator $A \in K(X)$ auf einem unendlichdimensionalen Banachraum. Wenn wir nun mögliche Spektralwerte $\lambda \neq 0$ untersuchen wollen, so bemerken wir zunächst, dass wir uns wegen $\lambda - A = \lambda(1 - A/\lambda)$ stets qualitativ auf den Fall $1 - A$ für A kompakt zurückziehen können.

2.38. Theorem (Satz von Riesz-Schauder A): Sei $A \in K(X)$, dann gilt:

(i) $\ker(1 - A)$ ist endlichdimensional,

(ii) $\operatorname{im}(1 - A)$ ist abgeschlossen und $X/\operatorname{im}(1 - A)$ ist endlichdimensional.

Beweis. (i): Setze $U := \ker(1 - A)$ und sei (x_n) eine beschränkte Folge im abgeschlossenen Teilraum U von X. Wegen der Kompaktheit von A konvergiert eine geeignete Teilfolge (Ax_{n_k}) in X. Wegen $(1-A)x_n = 0$ folgt insbesondere $x_{n_k} = Ax_{n_k}$ und somit die Konvergenz von (x_{n_k}). Wegen der Abgeschlossenheit von U liegt der Grenzwert von (x_{n_k}) in U. Wir haben also gezeigt, dass Kugeln in U relativkompakt sind, woraus $\dim U < \infty$ folgt.

(ii): Wir setzen $F := 1 - A$. Es ist $\ker F$ abgeschlossen sowie $X/\ker F$ gemäß Proposition 1.21 vollständig. Über der kanonischen Surjektion $q\colon X \to X/\ker F$ faktorisiert $F\colon X \to \operatorname{im} F$ wie üblich zur bijektiven linearen Abbildung $\widetilde{F}\colon X/\ker F \to \operatorname{im} F$ mit $\widetilde{F} \circ q = F$. Aufgrund der Definition der Quotientennorm (bzw. der zugehörigen Quotiententopologie) ist $\widetilde{F}$ stetig. Wir werden zeigen, dass auch $\widetilde{F}^{-1}$ stetig ist[6], somit $\widetilde{F}$ ein Isomorphismus und daher auch $\operatorname{im} F$ vollständig ist, also $\operatorname{im} F$ abgeschlossen in X.

Unstetigkeit von $\widetilde{F}^{-1}$ ist äquivalent (vgl. UE) zur Existenz einer Folge $(x_n + \ker F)_{n\in\mathbb{N}}$ mit $\|x_n + \ker F\|' = 1$ für alle $n \in \mathbb{N}$, aber $Fx_n = \widetilde{F}(x_n + \ker F) \to 0$ für $n \to \infty$. OBdA ist $\|x_n\| \leq 2$ (Definition von $\|\ \|'$ als Infimum) und wegen der Kompaktheit von A ist eine Teilfolge (Ax_{n_k}) konvergent, daher auch $x_{n_k} = Ax_{n_k} + Fx_{n_k}$. Sei $x = \lim x_{n_k}$, dann folgt $Fx = \lim Fx_{n_k} = 0$, d.h. $x \in \ker F$ und daher $\|x + \ker F\|' = 0$ im Widerspruch zu $\|x + \ker F\|' = \lim \|x_{n_k} + \ker F\|' = 1$. Also ist $\widetilde{F}^{-1}$ stetig und somit $\operatorname{im} F$ abgeschlossen.

Es bleibt zu zeigen, dass $X/\operatorname{im} F$ endlichdimensional ist. Zunächst gilt $(\operatorname{im} F)^\perp = \ker F'$, weil $F'(x') = 0$ für $x' \in X'$ gleichbedeutend mit $x'(Fx) = 0$ für alle $x \in X$ ist. Für den

[6] Den Satz von der stetigen Inversen können wir nicht anwenden, weil wir noch nicht „wissen“, dass $\operatorname{im} F$ vollständig ist.

abgeschlossenen Teilraum $\operatorname{im} F \subseteq X$ gilt laut oben erwähnter Lücke (vi), dass $(X/\operatorname{im} F)' \cong (\operatorname{im} F)^\perp = \ker F'$ ist. Laut Lücke (ii) ist auch A' kompakt und daher ist der erste Beweisteil auf $F' = 1 - A'$ anwendbar, womit $\dim \ker F' < \infty$ folgt. Daher ist auch der dazu isometrisch isomorphe Raum $(X/\operatorname{im} F)'$ endlichdimensional und insbesondere reflexiv. Gemäß Theorem 2.30 ist dann auch $X/\operatorname{im} F$ reflexiv, also gilt speziell $X/\operatorname{im} F \cong (X/\operatorname{im} F)'' = ((X/\operatorname{im} F)')'$, wobei der letzte Raum als Dualraum eines endlichdimensionalen Raumes selbst endlichdimensional ist. Somit ist $X/\operatorname{im} F$ endlichdimensional. □

Nun folgen einige Konstruktionen, die an die Jordan-Zerlegung aus der endlichdimensionalen Linearen Algebra erinnern.

2.39. Lemma: Sei $A \in K(X)$ und für $m \in \mathbb{N}$ jeweils $N_m := \ker(1 - A)^m$, $N_0 := \{0\}$ und $R_m := \operatorname{im}(1 - A)^m$, $R_0 := X$. Dann sind alle N_m endlichdimensional, alle R_m abgeschlossen mit $\dim X/R_m < \infty$ und ergeben die Inklusionsketten $N_0 \subseteq N_1 \subseteq N_2 \subseteq \cdots$ und $R_0 \supseteq R_1 \supseteq R_2 \supseteq \cdots$. Weiters gelten folgende Aussagen:

(a) Es gibt minimale $p, q \in \mathbb{N}_0$ mit $N_p = N_{p+1}$ bzw. $R_q = R_{q+1}$,

(b) Für $k \in \mathbb{N}$ ist $N_{p+k} = N_p$ und $R_{q+k} = R_q$,

(c) $N_p \cap R_p = \{0\}$,

(d) $N_q + R_q = X$,

(e) $p = q$.

Beweis. Es ist $(1-A)^m = \sum_{k=0}^{m} \binom{m}{k}(-1)^k A^k = 1-AB$ mit $B \in L(X)$, also AB ein kompakter Operator, sodass wir für $1-AB$ analog die Resultate des obigen Riesz-Schauder-Satzes anwenden dürfen. Daher ist jedes R_m abgeschlossen und N_m sowie X/R_m endlichdimensional. Wir setzen nun zur Abkürzung $F := 1 - A$.

Die Endlichkeit der aufsteigenden bzw. absteigenden Ketten von Teilräumen N_m bzw. R_m, also (a), wird indirekt bewiesen, wobei mittels sukzessiver Anwendung des Lemmas von Riesz (siehe Lücke (i)) ein Widerspruch zur Kompaktheit von A hergestellt wird. Genau so eine Technik führen wir im Beweis von Korollar 2.46 weiter unten im Detail aus, weshalb wir das hier mal ruhig überspringen können.

(b): $N_{p+k} \supseteq N_p$ und $R_{q+k} \subseteq R_q$ ist klar, bleibt also nur $N_{p+k} \subseteq N_p$ und $R_{q+k} \supseteq R_q$ zu zeigen. Für $x \in N_{p+k} = \ker F^{p+k}$ ist $F^{k-1}x \in N_{p+1} = N_p$, also $x \in N_{p+k-1}$ und, falls $k-1 > 0$, folgt nach weiteren $k-1$ Schritten schließlich $N_{p+k} \subseteq N_p$. Analog für $R_{q+k} \supseteq R_q$.

(c): Ist $x \in N_p \cap R_p$, dann also $F^p x = 0$ und es gibt ein $y \in X$ mit $x = F^p y$. Es folgt $F^{2p} y = 0$, sodass $y \in N_{2p} = N_p$ gilt. Daher ist letztlich $0 = F^p y = x$.

(d): Sei $x \in X$ beliebig. Es ist $F^q x \in R_q = R_{2q}$ und daher gibt es ein $y \in X$ mit $F^q x = F^{2q} y$. Daraus folgt $F^q(x - F^q y) = 0$, d.h. $x - F^q y \in N_q$. Wegen $F^q y \in R_q$ erreichen wir somit immer eine Darstellung $x = (x - F^q y) + F^q y$ passend zu $N_q + R_q$.

(e): Angenommen $p > q$, dann folgt aus (b) zunächst $R_p = R_q$. Wegen der Minimalität von p gemäß (a) folgt gleichzeitig, dass es ein $x \in N_p \setminus N_q$ gibt. Mittels (d) schreiben wir $x = y+z$ mit gewissen $y \in N_q$ und $z \in R_q = R_p$. Dann ergibt sich $z = x - y \in N_p + N_q \subseteq N_p$, also zusammen $z \in N_p \cap R_p$, was nach (c) nur $z = 0$ bedeuten kann. Demnach wäre $x = y \in N_q$, ein Widerspruch.
Angenommen $p < q$, dann ist $N_q = N_p$ nach (b) und wegen (a) gibt es ein $x \in R_p \setminus R_q$. Wieder haben wir nach (d) eine Darstellung $x = y + z$ mit $y \in N_q = N_p$ und $z \in R_q$. Nun folgt $y = x - z \in R_p + R_q \subseteq R_p$, also $y \in N_p \cap R_p = \{0\}$ gemäß (c). Also ist $x = z \in R_q$, ein Widerspruch. □

Wir ziehen eine fast unmittelbare Folgerung aus Lemma 2.39.

2.40. Korollar: Sei $A \in K(X)$, dann gibt es einen endlichdimensionalen Teilraum N und einen abgeschlossenen Teilraum R von X, sodass $X = N \oplus R$ topologisch ist, $(1-A)(N) \subseteq N$, $(1-A)(R) \subseteq R$ und $(1-A)|_R$ ein Isomorphismus $R \to R$ ist.

Beweis. Wir wählen p sowie $N := N_p$ und $R := R_p$ gemäß Lemma 2.39. Dann ist R abgeschlossen, N endlichdimensional und $X = N \oplus R$ (wegen 2.39(c)(d)), was nach dem Bericht aus Lücke (iv) automatisch ergibt, dass die direkte Summe auch topologisch ist.

Wir setzen $F := 1 - A$ und erhalten ebenso aus Lemma 2.39 nach Konstruktion dort $F(N) = F(N_p) = F(\ker F^p) \subseteq \ker F^{p-1} = N_{p-1} \subseteq N_p = N$ sowie $F(R) = F(R_p) = F(\operatorname{im} F^p) = \operatorname{im} F^{p+1} = R_{p+1} = R_p = R$. Aus Letzterem folgt bereits die Surjektivität von $F|_R$ als Abbildung $R \to R$. Sei $x \in R$ und $Fx = 0$, dann gibt es wegen $R = R_p$ zunächst ein $z \in X$ mit $x = F^p z$, was wegen $0 = Fx = F^{p+1} z$ sogar $z \in N_{p+1} = N_p = \ker F^p$ impliziert, d.h. $x = F^p z = 0$. Daher ist $F|_R$ als Abbildung $R \to R$ aufgefasst bijektiv, nach dem Satz über die stetige Inverse also ein Isomorphismus des Banachraumes R. □

2.41. Theorem (Satz von Riesz-Schauder B): Es sei $A \in K(X)$, dann gilt

$$\dim(X/\operatorname{im}(1-A)) = \dim(\ker(1-A)) = \dim(X'/\operatorname{im}(1-A')) = \dim(\ker(1-A')).$$

Beweis. Wir setzen wieder $F := 1 - A$, dann wissen wir bereits aus dem A-Teil des Satzes, dass $\operatorname{im} F$ abgeschlossen ist und $\ker F$ sowie $X/\operatorname{im} F$ endlichdimensional sind. Außerdem ergeben die Beobachtungen gegen Ende von dessen Beweis schon $\dim(\ker F') = \dim(X/\operatorname{im} F)' = \dim(X/\operatorname{im} F)$, was schon mal die Gleichheit der äußersten Terme in der behaupteten Gleichungskette ergibt.

Wenden wir (vi) „aus der Lücke“ mit $U = \ker F$ an, so folgt $(\ker F)' \cong X'/(\ker F)^\perp$ und wegen der Abgeschlossenheit von $\operatorname{im} F$ ergibt sich mit (v)(d) ebendort auch noch $(\ker F)^\perp = \operatorname{im}(F')$, sodass unmittelbar $\dim(X'/\operatorname{im} F') = \dim(\ker F)'$ folgt. Da $\ker F$ endlichdimensional ist, haben wir zuätzlich $\dim(\ker F) = \dim(\ker F)'$, also gilt das mittlere Gleichheitszeichen der Behauptung, womit schließlich nur noch Folgendes zu zeigen bleibt:

$$(\Delta) \qquad \dim(X/\operatorname{im} F) = \dim(\ker F).$$

Wir wenden nun Korollar 2.40 an und erhalten mit $F = 1 - A$ eine topologische direkte Summenzerlegung $X = N \oplus R$ mit endlichdimensionalem N, $F(N) \subseteq N$, $F(R) \subseteq R$ und $F|_R$ als Isomorphismus $R \to R$. Wir setzen $\overline{F} := F|_N$, sodass $\overline{F} \in L(N)$. Nachdem N endlichdimensional ist, verhilft uns die Lineare Algebra zu

$$(*) \qquad \dim(N/\operatorname{im}\overline{F}) = \dim(N) - \dim(\operatorname{im}\overline{F}) = \dim(\ker\overline{F}).$$

Wir zeigen nun, dass $\ker\overline{F} = \ker F$ gilt, wobei $\ker\overline{F} \subseteq \ker F$ klar ist. Sei also $x \in \ker F$ zerlegt gemäß $X = N \oplus R$ in der Form $x = x_1 + x_2$ mit eindeutigen $x_1 \in N$ und $x_2 \in R$. Wegen $0 = Fx = Fx_1 + Fx_2$ ist $Fx_2 = -Fx_1 \in N \cap R = \{0\}$, also $Fx_2 = 0$ und $\overline{F}x_1 = Fx_1 = 0$. Nach Letzterem ist $x_1 \in \ker\overline{F}$ und die Isomorphie $F|_R \colon R \to R$ erzwingt $x_2 = 0$, somit insgesamt $x = x_1 + x_2 = x_1 \in \ker\overline{F}$. Wir dürfen also nun in $(*)$ einsetzen und erhalten als Zwischenergebnis

$$\dim(N/\operatorname{im}\overline{F}) = \dim(\ker F).$$

Im letzten Schritt zeigen wir nun, dass die beiden endlichdimensionalen Vektorräume $N/\operatorname{im}\overline{F}$ und $X/\operatorname{im}F$ isomorph sind, womit dann der Beweis für (Δ) erbracht sein wird.

Vorweg bemerken wir, dass natürlich $\operatorname{im}\overline{F} \subseteq \operatorname{im}F$ gilt und daher für $x, y \in N$ mit $x + \operatorname{im}\overline{F} = y + \operatorname{im}\overline{F}$ in $X/\operatorname{im}\overline{F}$ auch stets $x + \operatorname{im}F = y + \operatorname{im}F$ in $X/\operatorname{im}F$ ist. Mit anderen Worten: Wir erhalten durch $x + \operatorname{im}\overline{F} \mapsto x + \operatorname{im}F$ eine wohldefinierte lineare Abbildung $Q \colon N/\operatorname{im}\overline{F} \to X/\operatorname{im}F$.

Q ist surjektiv: Sei $x + \operatorname{im}F \in X/\operatorname{im}F$ beliebig und $x = x_1 + x_2$ die eindeutige Zerlegung mit $x_1 \in N$ und $x_2 \in R$. Dann ist $x - x_1 = x_2 \in R = \operatorname{im}(F|_R) \subseteq \operatorname{im}F$ und daher $Q(x_1 + \operatorname{im}\overline{F}) = x_1 + \operatorname{im}F = x + \operatorname{im}F$.

Q ist injektiv: Sei $Q(x + \operatorname{im}\overline{F}) = 0 + \operatorname{im}F$ mit $x \in N$, was wegen $Q(x + \operatorname{im}\overline{F}) = x + \operatorname{im}F$ einfach $x \in N \cap \operatorname{im}F$ bedeutet. Es gibt also ein $z \in X$ mit $Fz = x$. Mit der Zerlegung $z = z_1 + z_2$, $z_1 \in N$, $z_2 \in R$ folgt $Fz_2 = Fz - Fz_1 = x - Fz_1 \in N$, weil ja $F(N) \subseteq N$ gilt. Aber natürlich ist ebenso $Fz_2 \in F(R) \subseteq R$, also $Fz_2 \in N \cap R = \{0\}$, somit $z_2 = 0$, weil $F|_R$ ein Isomorphismus ist. Wir erhalten $x = Fz_1 \in F(N) = \operatorname{im}\overline{F}$, d.h. $x + \operatorname{im}\overline{F}$ ist die Nullklasse in $N/\operatorname{im}\overline{F}$. □

2.42. Bemerkung: Ein Operator $B \in L(X)$ mit Eigenschaften analog zu jenen für $1 - A$ aus dem Riesz-Schauder-Satz A, d.h. $\ker B$ endlichdimensional, $\operatorname{im}B$ abgeschlossen und $X/\operatorname{im}B$ endlichdimensional heißt übrigens *Fredholmoperator*. Der sogenannte *Index* von B ist dann gegeben durch

$$\operatorname{ind}(B) := \dim(\ker B) - \dim(X/\operatorname{im}B).$$

Also besagen die Riesz-Schauder-Sätze A und B zusammen gerade, dass $1-A$ für kompaktes A stets ein Fredholmoperator mit $\operatorname{ind}(1 - A) = 0$ ist.

2.43. Korollar (Fredholmsche Alternative): Sei $A \in K(X)$ und $\lambda \in \mathbb{K}$ mit $\lambda \neq 0$, dann tritt genau einer der beiden folgenden Fälle ein:

(a) Die homogene Gleichung $\lambda x - Ax = 0$ hat nur die triviale Lösung $x = 0$ und dann ist für jedes $y \in X$ die inhomogene Gleichung $\lambda x - Ax = y$ eindeutig lösbar für $x \in X$.

(b) Es ist $0 < n = \dim(\ker(\lambda - A)) < \infty$, d.h. die homogene Gleichung $\lambda x - Ax = 0$ hat n linear unabhängige Lösungen, ebenso folgt $n = \dim(\ker(\lambda - A'))$, und dann ist die inhomogene Gleichung $\lambda x - Ax = y$ für $x \in X$ lösbar genau für rechte Seiten y aus $\ker(\lambda - A')_\perp = \{z \in X \mid \forall y' \in \ker(\lambda - A') \colon y'(z) = 0\}$.

Beweis. Folgt einfach mittels $\lambda - A = \lambda(1 - A/\lambda)$ aus den obigen Sätzen A und B von Riesz-Schauder zusammen mit dem Satz 2.32. □

2.44. Bemerkung: In obigem Korollar gibt also jeweils die „Größe der Kerne“ Aufschluss über die Lösbarkeit der inhomogenen Gleichung. Der obige Satz entsprang historisch aus umfangreichen Untersuchungen von Fredholm zu Integralgleichungen um 1900 (vgl. z.B. die stets sehr lesenswerten historischen Notizen in [Heu06, Wer18]).

2.45. Korollar: Sei $A \in K(X)$, dann gilt $\sigma(A) \setminus \{0\} \subseteq \sigma_p(A)$, d.h. jedes $\lambda \in \sigma(A)$ mit $\lambda \neq 0$ ist ein Eigenwert von A. Weiters ist für ein solches λ der Eigenraum $\ker(\lambda - A)$ immer endlichdimensional und es gibt eine topologische direkte Summenzerlegung $X = N_\lambda \oplus R_\lambda$ mit endlichdimensionalem $N_\lambda \supseteq \ker(\lambda - A)$, sowohl N_λ als auch R_λ sind invariant[7] unter A und $(\lambda - A)|_{R_\lambda}$ ist ein Isomorphismus $R_\lambda \to R_\lambda$.

Beweis. Mittels $\lambda - A = \lambda(1 - A/\lambda)$ folgt dies direkt aus den Riesz-Schauder-Sätzen, Korollar 2.40 und der Fredholm-Alternative. □

Nach Konstruktion ist übrigens $(\lambda - A)|_{N_\lambda}$ wegen 2.39(b) nilpotent (auf N_λ).

2.46. Korollar: Sei $A \in K(X)$, dann gilt:

(i) die Menge $\sigma(A) \setminus \{0\}$ ist höchstens abzählbar, [$\sigma_p(A)$ kann auch leer sein.]

(ii) $\sigma(A)$ hat keinen Häufungspunkt ungleich 0. [Und auch 0 muss keiner sein.]

Beweis. Wir zeigen (i) und (ii) auf einen Schlag, indem wir folgende Behauptung beweisen:

$$\text{Für jedes } \varepsilon > 0 \text{ ist die Menge } C_\varepsilon := \{\lambda \in \sigma(A) \mid |\lambda| \geq \varepsilon\} \text{ endlich.}$$

Sei $\varepsilon > 0$ und C_ε unendlich. Dann gibt es eine Folge (λ_n) in C_ε mit $\lambda_n \neq \lambda_m$ für $n \neq m$. Wir wissen bereits aus Korollar 2.45, dass C_ε nur aus Eigenwerten bestehen kann. Daher gibt es für jedes $n \in \mathbb{N}$ auch einen Eigenvektor $x_n \neq 0$ zum Eigenwert λ_n. Die Menge $\{x_n \mid n \in \mathbb{N}\}$ ist linear unabhängig, weil es sich um Eigenvektoren zu paarweise verschiedenen Eigenwerten handelt (und für jede endliche Auswahl von paarweise verschiedenen Vektoren das analoge Argument aus der Linearen Algebra ausreicht).

[7] d.h. $A(N_\lambda) \subseteq N_\lambda$ und $A(R_\lambda) \subseteq R_\lambda$

Wir setzen $U_n := \operatorname{span}\{x_1, \ldots, x_n\}$ und erhalten eine Folge von endlichdimensionalen (daher abgeschlossenen) Teilräumen mit $U_1 \subsetneq U_2 \subsetneq \cdots$ und $A(U_n) \subseteq U_n$. Mittels Lemma von Riesz (Lücke (i)) können wir für jedes $n \in \mathbb{N}$ ein $z_n \in U_n$ finden, sodass

$$\|z_n\| = 1 \quad \text{und} \quad \forall u \in U_{n-1} \colon \|z_n - u\| \geq \frac{1}{2}.$$

Für $n > m$ ist sicherlich $Az_m \in U_m \subseteq U_{n-1}$. Weiters ergibt sich auch $\lambda_n z_n - Az_n \in U_{n-1}$, weil $z_n = y + \alpha x_n$ mit $y \in U_{n-1}$, $\alpha \in \mathbb{K}$ gilt und somit $\lambda_n z_n - Az_n = (\lambda_n y + \lambda_n \alpha x_n) - (Ay + \lambda_n \alpha x_n) = \lambda_n y - Ay \in U_{n-1}$. Daher ist auch $w := Az_m + \lambda_n z_n - Az_n \in U_{n-1}$ und weiter $\lambda_n z_n - w = \lambda_n(z_n - w')$ mit $w' = w/\lambda_n \in U_{n-1}$. Schließlich erhalten wir damit einen Widerspruch zur Kompaktheit von A: Für $m, n \in \mathbb{N}$ mit $n > m$ gilt $\|z_n\| = 1$ und

$$\|Az_n - Az_m\| = \|\lambda_n z_n - \underbrace{(Az_m + \lambda_n z_n - Az_n)}_{w}\| = \|\lambda_n(z_n - w')\| = |\lambda_n| \|z_n - w'\| \geq \frac{\varepsilon}{2}.$$

$\square$

Also enthält das Spektrum eines kompakten Operators auf einem unendlichdimensionalen Banachraum zumindest 0 (ohne, dass dies ein Eigenwert sein muss) und besteht abgesehen davon nur aus Eigenwerten, von denen es aber höchstens abzählbar viele gibt (eventuell sogar gar keine) und die sich, wenn überhaupt, nur bei 0 häufen können.

3 Hilberträume

Wir haben es weiterhin bei allen hier betrachteten Vektorräumen nur mit dem Skalarenkörper $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ zu tun. Typischer Weise werden wir sogar gleich die Formulierungen für den komplexen Fall geben, wobei dann im Reellen z.B. die komplexe Konjugation einfach zu ignorieren ist (weil $z \mapsto \overline{z}$ auf $\mathbb{R}$ eingeschränkt wie die Identität wirkt). Wie wir aus der Linearen Algebra bereits wissen, gibt es aber bei der Existenz von Eigenwerten natürlich deutliche Vorzüge für $\mathbb{K} = \mathbb{C}$. Weiters kennen wir aus der Linearen Algebra den Begriff des *euklidischen* oder *unitären* Raumes, d.h. eines $\mathbb{K}$-Vektorraumes X mit einem *Skalarprodukt* $\langle .,. \rangle \colon X \times X \to \mathbb{K}$, das folgende Eigenschaften hat:

(a) Für jedes $y \in X$ ist $x \mapsto \langle x, y \rangle$ linear $X \to \mathbb{K}$,
(b) $\forall x, y \in X$: $\langle y, x \rangle = \overline{\langle x, y \rangle}$, (Hermitizität bzw. Symmetrie)
(c) $\forall x \in X$: $\langle x, x \rangle \geq 0$ und $(\langle x, x \rangle = 0 \Leftrightarrow x = 0)$. ($\langle .,. \rangle$ ist positiv definit)

Es folgt, dass für fixes $x \in X$ die Abbildung $y \mapsto \langle x, y \rangle$, $X \to \mathbb{K}$, konjugiert linear ist, d.h. $\langle x, \lambda_1 y_1 + \lambda_2 y_2 \rangle = \overline{\lambda_1} \langle x, y_1 \rangle + \overline{\lambda_2} \langle x, y_2 \rangle$, was im Falle $\mathbb{K} = \mathbb{R}$ einfach linear heißt und im Falle $\mathbb{K} = \mathbb{C}$ oft auch für die erste Komponente in $\langle .,. \rangle$ statt in der zweiten verlangt wird (insbesondere z.B. in der Quantenphysik). Im Rellen erhalten wir durch $\langle .,. \rangle$ also eine *bilineare* Abbildung $X \times X \to \mathbb{K}$, im Komplexen eine *sesquilineare*.

Wir erinnern an die in der Linearen Algebra bewiesene *Cauchy-Schwarz-Ungleichung:*

$$\forall x, y \in X: \quad |\langle x, y \rangle|^2 \leq \langle x, x \rangle \langle y, y \rangle,$$

wobei Gleichheit genau dann eintritt, wenn x, y linear abhängig sind. Daraus wurde dann weiters gezeigt, dass durch

$$\forall x \in X: \quad \|x\| := \sqrt{\langle x, x \rangle}$$

eine *Norm* auf X definiert wird, sodass die Cauchy-Schwarz-Ungleichung auch so geschrieben werden kann:

$$\forall x, y \in X: \quad |\langle x, y \rangle| \leq \|x\| \|y\|.$$

Wenn nichts anderes gesagt wird, betrachten wir einen Vektorraum mit Skalarprodukt immer wie oben gegeben automatisch als normierten Raum.

3.1. Definition: Sei X ein Vektorraum mit Skalarprodukt. Dann heißt der daraus wie oben definierte normierte Raum *Prähilbertraum.* Falls X mit dieser Norm ein Banachraum ist, heißt er *Hilbertraum.*

3.2. Lemma: Sei X ein Prähilbertraum, dann gilt:

(i) Das Skalarprodukt ist stetig $X \times X \to \mathbb{K}$,

(ii) ist U ein dichter Teilraum und $x \in X$ mit $\langle x, u \rangle = 0$ für alle $u \in U$, dann folgt $x = 0$.

Beweis. (i): Für $x_0, x, y_0, y \in X$ erhalten wir nämlich

$$|\langle x, y\rangle - \langle x_0, y_0\rangle| = |\langle x - x_0, y\rangle + \langle x_0, y - y_0\rangle| \leq \|x - x_0\| \|y\| + \|x_0\| \|y - y_0\|.$$

(ii): Gemäß (i) ist insbesondere die (konjugiert lineare) Abbildung $l_x \colon y \mapsto \langle x, y\rangle$ stetig $X \to \mathbb{K}$ und verschwindet auf einer dichten Teilmenge, muss also insgesamt die Nullabbildung sein. Daher ist speziell $0 = l_x(x) = \langle x, x\rangle$, also $x = 0$. □

3.3. Bemerkung: (i) In der Linearen Algebra (VO und UE) sind wir auch ein kleines Stück der Frage nachgegangen, welche Normen denn überhaupt von Skalarprodukten stammen. Wir hatten gesehen, dass die Gültigkeit der *Parallelogrammgleichung*

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2 \quad \forall x, y \in X$$

sowie der sogenannten *Polarisierungsformeln*

$$4\langle x, y\rangle = \|x + y\|^2 - \|x - y\|^2 \quad (\text{für } \mathbb{K} = \mathbb{R})$$
$$\text{bzw.} \quad 4\langle x, y\rangle = \|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2 \quad (\text{für } \mathbb{K} = \mathbb{C})$$

notwendig ist und bereits bemerkt, dass umgekehrt ein Skalarprodukt mittels Polarisierungsformeln aus einer Norm „rekonstruiert" werden kann, falls diese die Parallelogrammgleichung erfüllt. (Für einen Beweis siehe [Wer18, Satz V.1.7]).

(ii) Teilräume von Prähilberträumen sind Prähilberträume. Weiters ist die Vervollständigung eines Prähilbertraumes (als normierter Raum) ein Hilbertraum, weil auch die Norm auf der Vervollständigung aus Stetigkeitsgründen die Parallelogrammgleichung erfüllt.

3.4. Beispiel: 1) Auf $\mathbb{C}^n$ haben wir das *Standardskalarprodukt* gegeben durch

$$\langle x, y\rangle = \sum_{k=1}^{n} x_k \overline{y_k} \quad \text{für } x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in \mathbb{C}^n.$$

2) Auf l^2 entsteht die Norm $\|.\|_2$ genau aus dem *Standardskalarprodukt*

$$\langle x, y\rangle = \sum_{k=1}^{\infty} x_k \overline{y_k} \quad \text{für } x = (x_k), y = (y_k) \in l^2.$$

Somit ist l^2 ein Hilbertraum.

3) Sei S eine nichtleere Menge und bezeichne $l^2(S)$ die Menge aller Funktionen $f\colon S \to \mathbb{C}$, sodass $f(s) \neq 0$ höchstens für s aus einer abzählbaren Teilmenge $S_f = \{s_1, s_2, \dots\} \subseteq S$ gilt und $\sum_{j=1}^{\infty} |f(s_j)|^2$ konvergent ist. Wir schreiben[1] für diese absolut konvergente Summe im Weiteren einfach $\sum_{s\in S} |f(s)|^2$.

Für $f, g \in l^2(S)$ setzen wir $\langle f, g\rangle := \sum_{s\in S} f(s)\overline{g(s)}$, was sich leicht als Skalarprodukt erkennen lässt, und gewinnen daraus die Norm $\|f\| := \sqrt{\langle f, f\rangle}$, deren Vollständigkeit genau wie im Falle der l^p-Räume gezeigt werden kann.

Für $S = \mathbb{N}$ erhalten wir wieder genau l^2. Für überabzählbares S ist $l^2(S)$ ein Beispiel eines nichtseparablen Hilbertraumes: Zwei Vektoren aus der Menge $\{\delta_t \mid t \in S\}$, wobei $\delta_t(s) = 0$ für $s \neq t$ und $\delta_t(t) = 1$ sei, haben für $t_1 \neq t_2$ den Abstand $\|\delta_{t_1} - \delta_{t_2}\| = \sqrt{2}$.

4) Für einen Maßraum (Ω, μ) ist auf $L^2(\Omega, \mu)$ durch

$$\langle f, g\rangle := \int_{\Omega} f(\omega)\overline{g(\omega)}\, d\mu(\omega) \quad (f, g \in L^2(\Omega, \mu))$$

ein Skalarprodukt definiert, das gerade die Norm $\|.\|_2$ generiert. Somit ist $L^2(\Omega, \mu)$ ein Hilbertraum. Im Falle $\Omega = S$ mit dem Zählmaß μ ergibt sich daraus als Spezialfall $l^2(S)$.

Die folgenden Begriffe sind allesamt schon aus der Linearen Algebra bekannt.

3.5. Definition: Zwei Vektoren x und y in einem Prähilbertraum X heißen *orthogonal*, falls $\langle x, y\rangle = 0$ gilt. In diesem Fall schreiben wir auch $x \perp y$. Für Teilmengen $A, B \subseteq X$ soll $A \perp B$ einfach bedeuten, dass $x \perp y$ für alle $x \in A$ und $y \in B$ gilt. Schließlich heißt

$$A^\perp := \{x \in X \mid \forall a \in A\colon \langle x, a\rangle = 0\}$$

das *orthogonale Komplement* von A.

3.6. Einfache Eigenschaften der Orthogonalitätsbegriffe: Sei X ein Prähilbertraum. Die ersten beiden der folgenden Punkte sind Wiederholungen aus der Linearen Algebra.

(i) Satz von Pythagoras: Sind $x_1, \dots, x_n \in X$ paarweise orthogonal, dann gilt

$$\left\| \sum_{j=1}^{n} x_j \right\|^2 = \sum_{j=1}^{n} \|x_j\|^2.$$

(ii) Für Teilmengen $A, B \subseteq X$ gilt: $A \cap A^\perp = \{0\}$, falls $0 \in A$, und $A \cap A^\perp = \emptyset$ sonst; $A^\perp$ ist ein Teilraum von X; $A \subseteq B$ impliziert $B^\perp \subseteq A^\perp$; $\operatorname{span} A \subseteq (A^\perp)^\perp$.

[1] Dies lässt sich rechtfertigen, weil sogenannte *summierbare Familien* immer nur höchstens abzählbar viele Elemente ungleich Null haben können (vgl. [Schr00, Bemerkung 4 zu Definition 1.2.14]).

(iii) $A^\perp$ ist ein abgeschlossener Teilraum (als Durchschnitt der Kerne der stetigen linearen Abbildungen $x \mapsto \langle x, a\rangle$ für $a \in A$). Daher folgt mit (ii) auch gleich $\overline{\operatorname{span} A} \subseteq (A^\perp)^\perp$, wobei wir später sogar Gleichheit zeigen werden.

(iv) $A^\perp = (\overline{\operatorname{span} A})^\perp$ [$\supseteq$ folgt aus $A \subseteq \overline{\operatorname{span} A}$ und (ii); $\subseteq$ folgt daraus, dass $a \mapsto \langle x, a\rangle$ konjugiert linear und stetig ist.]

3.7. Approximationsaufgaben: Es sei hier H ein Hilbertraum.

Proposition: Ist K eine nichtleere, konvexe und abgeschlossene Teilmenge von H, so gibt es für jedes $x \in H$ ein eindeutiges $y \in K$ mit der Eigenschaft

$$\|x - y\| = \inf\{\|x - z\| \mid z \in K\}.$$

Beweis. Wir setzen $\gamma := \inf\{\|x - z\| \mid z \in K\}$ und wählen eine Folge (y_n) in K mit $\|x - y_n\| \to \gamma$ für $n \to \infty$. Seien $n, m \in \mathbb{N}$ und $u_m := x - y_m$, $v_n := x - y_n$. Dann ist

$$u_m + v_n = 2\left(x - \frac{y_m + y_n}{2}\right), \quad u_m - v_n = y_n - y_m$$

und die Parallelogrammregel ergibt

$$\begin{aligned}\|y_n - y_m\|^2 = \|u_m - v_n\|^2 &= 2\|u_m\|^2 + 2\|v_n\|^2 - \|u_m + v_n\|^2 \\ &= 2\|x - y_m\|^2 + 2\|x - y_n\|^2 - 4\left\|x - \frac{y_m + y_n}{2}\right\|^2.\end{aligned}$$

Wegen der Konvexität von K ist $(y_m + y_n)/2 \in K$, also $\|x - (y_m + y_n)/2\| \geq \gamma$ und daher

$$\|y_n - y_m\|^2 \leq 2\|x - y_m\|^2 + 2\|x - y_n\|^2 - 4\gamma^2.$$

Wegen $\|x - y_m\|^2 \to \gamma^2$ für $m \to \infty$ und $\|x - y_n\|^2 \to \gamma^2$ für $n \to \infty$ ist also (y_n) eine Cauchyfolge in K und besitzt somit einen Grenzwert $y \in K$, für den $\|x - y\| = \gamma$ gilt. Damit ist die Existenz von y gezeigt und nur noch die Eindeutigkeit bleibt zu behandeln.

Ist $\tilde{y} \in K$ eine weitere Lösung der Minimierungsaufgabe, d.h. $\|x - \tilde{y}\| = \gamma = \|x - y\|$, dann können wir die Vektorfolge (y_n) in K mit $y_{2k+1} := y$ und $y_{2k} := \tilde{y}$ betrachten, für die trivialer Weise $\lim \|x - y_n\| = \gamma$ gilt. Wie oben folgt $\|y_n - y_m\|^2 \leq 2\|x - y_m\|^2 + 2\|x - y_n\|^2 - 4\gamma^2$, woraus sich für $n = 2k + 1$ und $m = 2k$ ergibt, dass $y = \tilde{y}$ sein muss. □

Korollar: Ist U ein abgeschlossener Teilraum von H, so gibt es für jedes $x \in H$ ein eindeutiges $y \in U$ mit der Eigenschaft

$$\|x - y\| = \inf\{\|x - u\| \mid u \in U\}$$

und es gilt $x - y \perp U$.

Beweis. Die eindeutige Existenz von y folgt aus obiger Proposition, wir müssen also noch $\langle x-y, u\rangle = 0$ für jedes $u \in U$ zeigen. Da dies für $u = 0$ trivial ist, betrachten wir nur noch $u \in U \setminus \{0\}$. Für beliebiges $\lambda \in \mathbb{K}$ gilt wegen $y + \lambda u \in U$ dann

$$\begin{aligned}\|x-y\|^2 \leq \|x-y-\lambda u\|^2 = \|x-y\|^2 - \lambda\langle u, x-y\rangle - \overline{\lambda}\langle x-y, u\rangle + |\lambda|^2\|u\|^2 \\ = \|x-y\|^2 - 2\operatorname{Re}(\lambda\langle u, x-y\rangle) + |\lambda|^2\|u\|^2,\end{aligned}$$

was für $\lambda := \langle x-y, u\rangle/\|u\|^2$ sofort zur Ungleichung

$$\|x-y\|^2 \leq \|x-y\|^2 - 2\frac{|\langle x-y, u\rangle|^2}{\|u\|^2} + \frac{|\langle x-y, u\rangle|^2}{\|u\|^2} = \|x-y\|^2 - \frac{|\langle x-y, u\rangle|^2}{\|u\|^2}$$

führt. Diese kann aber nur gelten, falls $\langle x-y, u\rangle = 0$ ist. □

3.8. Zerlegungssatz: Ist U ein abgeschlossener Teilraum des Hilbertraumes H, dann gilt die *orthogonale direkte Summenzerlegung* $H = U \oplus U^\perp$.

Insbesondere hat also in einem Hilbertraum jeder abgeschlossene Teilraum einen abgeschlossenen Komplementärraum.

Beweis. Die Eigenschaft $U \cap U^\perp = \{0\}$ ist klar, daher bleibt nur $H = U + U^\perp$ zu zeigen. Für $x \in H$ sei $y \in U$ die eindeutige Bestapproximation an x aus U gemäß Korollar 3.7, dann ist insbesondere $x - y \in U^\perp$ und die Darstellung $x = y + (x-y)$ beendet den Beweis. □

3.9. Korollar: Ist U ein Teilraum des Hilbertraumes H, dann gilt $\overline{U} = (U^\perp)^\perp$.

Beweis. Die Relation $\overline{U} \subseteq (U^\perp)^\perp$ folgt bereits aus 3.6(iii). Nach 3.6(iv) ist $U^\perp = \overline{U}^\perp$ und der Zerlegungssatz angewandt auf $\overline{U}$ ergibt nun $H = \overline{U} \oplus U^\perp$. Angenommen, es wäre $x \in (U^\perp)^\perp \setminus \overline{U}$. Dann gibt es gemäß der Zerlegung eine Darstellung $x = u + v$ mit eindeutigen $u \in \overline{U} \subseteq (U^\perp)^\perp$ und $v \in U^\perp$, wobei wegen $x \notin \overline{U}$ jedenfalls $v \neq 0$ ist. Andererseits gilt $v = x - u \in (U^\perp)^\perp + (U^\perp)^\perp = (U^\perp)^\perp$, daher folgt $v \in U^\perp \cap (U^\perp)^\perp = \{0\}$ nach 3.6(ii), was ein Widerspruch zu $v \neq 0$ ist. □

3.10. Theorem (Orthogonalprojektion): Ist $U \neq \{0\}$ ein abgeschlossener Teilraum des Hilbertraumes H, dann existiert genau ein $P \in L(H)$ mit $\|P\| = 1$, $\operatorname{im} P = U$, $\ker P = U^\perp$, $P^2 = P$ und $\langle Px, y\rangle = \langle x, Py\rangle$ für alle $x, y \in H$.

Die lineare Abbildung P heißt *Orthogonalprojektion* auf U und wird oft auch mit P_U bezeichnet. Im Falle $U = H$ ist $P_U = 1$, im Falle $U \neq H$ ist $1 - P_U$ die Orthogonalprojektion auf $U^\perp$ mit $\|1 - P_U\| = 1$. Für $U = \{0\}$ setzen wir $P_U = 0$. (Bemerkung aus der Linearen Algebra: Allgemein heißen lineare Abbildungen $A\colon H \to H$ mit $A^2 = A$ Projektoren, aber nicht jeder Projektor ist eine Orthogonalprojektion.)

Beweis. Gemäß Zerlegungssatz ist $H = U \oplus U^\perp$, d.h. für jedes $x \in H$ gibt es eindeutige Vektoren $u \in U$ und $u_\perp \in U^\perp$ mit $x = u + u_\perp$. Wir verwenden diese Notation nun durchgehend im weiteren Beweisverlauf.

Eindeutigkeit: Sei P eine lineare Abbildung mit den genannten Eigenschaften. Es folgt $Pu_\perp = 0$ und $Px = Pu$. Wegen $P(1-P)x = Px - P^2x = Px - Px = 0$ ist $\operatorname{im}(1-P) \subseteq \ker P = U^\perp$. Für beliebiges $y \in H$ sei $y = v + v_\perp$ mit $v \in U$, $v_\perp \in U^\perp$, dann erhalten wir

$$\begin{aligned}\langle u - Pu, y\rangle = \langle u - Pu, v + v_\perp\rangle &= \langle u, v\rangle - \langle Pu, v\rangle + \langle u, v_\perp\rangle - \langle Pu, v_\perp\rangle \\ &= \langle u, v\rangle - \langle u, Pv\rangle = \langle u, (1-P)v\rangle = 0,\end{aligned}$$

daher $u = Pu$, weil y beliebig war. Somit ist P aber sowohl auf U als auch auf $U^\perp$ festgelegt.

Existenz: Wir definieren $Px := u$ als lineare Abbildung und erhalten wegen $\|Px\|^2 = \|u\|^2 \leq \|u\|^2 + \|u_\perp\|^2 = \|x\|^2$ sofort $\|Px\| \leq \|x\|$, also die Stetigkeit mit $\|P\| \leq 1$. Wegen $U \neq \{0\}$ gibt es $u \in U$ mit $\|u\| = 1$ und daher $\|Pu\| = \|u\| = 1$, also gilt $\|P\| = 1$.

Laut Definition von P ist $\operatorname{im} P \subseteq U$, $\ker P \subseteq U^\perp$ und außerdem $P^2x = Pu = u = Px$, d.h. $P^2 = P$. Ist $x \in U$, dann gilt wegen $u = x$ auch $x = Px$, also $U \subseteq \operatorname{im} P$, womit schon mal $\operatorname{im} P = U$ feststeht. Ist $x \in U^\perp$, dann folgt wegen $u_\perp$ ebenfalls aus der Definition $Px = 0$, also $U^\perp \subseteq \ker P$, d.h. insgesamt $\ker P = U^\perp$.

Zerlegen wir wieder $y \in H$ in der Form $y = v + v_\perp$ mit $v \in U$, $v_\perp \in U^\perp$, dann erhalten wir

$$\langle Px, y\rangle = \langle u, v + v_\perp\rangle = \langle u, v\rangle = \langle u, Py\rangle = \langle u + u_\perp, Py\rangle = \langle x, Py\rangle.$$

□

3.11. Theorem (Satz von Fréchet-Riesz): Sei H ein Hilbertraum und für jedes $y \in H$ das Element $l_y \in H'$ definiert durch $l_y(x) := \langle x, y\rangle$. Dann ist die Abbildung $F\colon H \to H'$, $y \mapsto l_y$ konjugiert linear, isometrisch und bijektiv. Insbesondere gibt es zu jedem $l \in H'$ genau ein $z \in H$, sodass

$$l(x) = l_z(x) = \langle x, z\rangle \quad \forall x \in H$$

und $\|l\| = \|z\|$ gilt.

Beweis. Es ist klar, dass F konjugiert linear ist und weiters mittels Cauchy-Schwarz-Ungleichung $\|F(y)\| = \sup_{\|x\|\leq 1} |\langle x, y\rangle| \leq \sup_{\|x\|\leq 1} \|x\|\|y\| = \|y\|$ folgt. Für $y \neq 0$ ist $F(y)(y/\|y\|) = \langle y, y\rangle/\|y\| = \|y\|$, daher ist $\|F(y)\| = \|y\|$; der Fall $y = 0$ ist trivial, also ist F isometrisch, weshalb auch die Injektivität klar ist.

Für den Beweis der Surjektivität sei also $l \in H'$ gegeben, wobei wir OBdA $\|l\| = 1$ annehmen dürfen. Wir setzen $U := \ker l$ und erhalten aus dem Zerlegungssatz $H = U \oplus U^\perp$.

Behauptung: $\dim U^\perp = 1$.

Zunächst wiederholen wir aus der Linearen Algebra, dass $\dim(H/U) = 1$ gilt. Wir betrachten die lineare Abbildung $f\colon H/U \to U^\perp$ mit $f(x + U) := u_\perp$, wobei $x = u + u_\perp$

die eindeutige Zerlegung mit $u \in U$ und $u_\perp \in U^\perp$ sei. Ist $f(x) = 0$, dann heißt dies $u_\perp = 0$, was wiederum $x = u \in U$ bedeutet, also $x + U = 0 + U$. Somit ist f injektiv. Für beliebiges $v \in U^\perp$ ist $f(v + U) = v$, weil $v = 0 + v$ die eindeutige Zerlegung gemäß $U \oplus U^\perp$ ist. Also ist f auch surjektiv und insgesamt ein Vektorraumisomorphismus. Daher ist $\dim U^\perp = \dim(H/U) = 1$.

Nach der Behauptung gibt es ein eindeutiges $z \in U^\perp$ mit $U^\perp = \operatorname{span}\{z\}$ und $l(z) = 1$. Jedes $x \in H = U \oplus U^\perp$ zerlegen wir nun in der Form $x = u + \lambda z$ mit eindeutigen $u \in U$ und $\lambda \in \mathbb{K}$. Es folgt $l(x) = l(u + \lambda z) = l(u) + \lambda l(z) = \lambda$ und $l_z(x) = \langle x, z\rangle = \langle u + \lambda z, z\rangle = \lambda \|z\|^2$, d.h. $F(z) = l_z = \|z\|^2 l$; aber wegen $\|z\| = \|F(z)\| = \|z\|^2 \|l\| = \|z\|^2$ muss auch $\|z\| = 1$ gelten, also insgesamt $l = l_z$. □

3.12. Korollar: Jeder Hilbertraum ist reflexiv.

Beweis. Sei H ein Hilbertraum und $F\colon H \to H'$ die konjugiert lineare, bijektive und isometrische Abbildung aus dem Satz von Fréchet-Riesz. Dann wird durch

$$\langle F(x), F(y)\rangle' := \langle y, x\rangle \quad \forall x, y \in H$$

ein Skalarprodukt auf H' definiert (die Reihenfolge der Faktoren ist auf der rechten Seite absichtlich umgedreht!), das wegen der Isometrie von F genau zur Operatornorm auf H' passt. Daher ist H' ein Hilbertraum und es gibt wiederum nach Fréchet-Riesz die entsprechende konjugiert lineare, bijektive und isometrische Abbildung $F'\colon H' \to H''$. Wir beachten, dass $F' \circ F\colon H \to H''$ als Verknüpfung konjugiert linearer Abbildungen *linear* ist; bijektiv und isometrisch ist $F' \circ F$ natürlich auch, also ein isometrischer Isomorphismus von H mit H''. Wir müssen noch zeigen, dass $F' \circ F$ gleich der kanonischen Isometrie $\iota\colon H \to H''$ ist.

Seien $x \in H$ und $l \in H'$ beliebig, dann ist bekanntlich $\iota(x)(l) = l(x)$. Schreiben wir $l = F(y)$ mit einem eindeutigen Urbild $y \in H$, so erhalten wir

$$\iota(x)(l) = l(x) = F(y)(x) = \langle x, y\rangle = \langle F(y), F(x)\rangle' = F'(F(x))(F(y)) = (F' \circ F)(x)(l).$$

Weil $l \in H'$ beliebig war, folgt $\iota(x) = (F' \circ F)(x)$. Weil auch $x \in H$ beliebig war, heißt dies gerade $\iota = F' \circ F$. □

3.13. Korollar: Sei H ein Hilbertraum, dann gelten folgende Aussagen:

(i) Eine Folge (x_n) in H konvergiert genau dann schwach gegen $x \in H$, falls

$$\forall y \in H\colon \quad \langle x_n, y\rangle \to \langle x, y\rangle \quad (n \to \infty).$$

(ii) Jede beschränkte Folge in H besitzt eine schwach konvergente Teilfolge.

Beweis. (i): Laut Definition heißt schwache Konvergenz von (x_n) gegen x gerade $l(x_n) \to l(x)$ für alle $l \in H'$, was wir mittels Satz von Fréchet-Riesz umformulieren können in $\langle x_n, y\rangle \to \langle x, y\rangle$ für alle $y \in H$.

(ii): Im vorigen Korollar wurde bewiesen, dass H reflexiv ist. Daher folgt diese Behauptung nun direkt aus Theorem 2.31. □

Wie in der Linearen Algebra gezeigt wurde, gibt es in endlichdimensionalen euklidischen oder unitären Räumen stets Basen, die zugleich Orthonormalsysteme sind, d.h. alle Basisvektoren sind paarweise orthogonal und auf Länge 1 normiert. Solche Orthonormalbasen waren z.B. für die Berechnung der eindeutigen Koeffizienten eines Vektors in seiner Basisentwicklung besonders vorteilhaft. Es zeigt sich, dass wir viel von diesem letzten Aspekt gewissermaßen „ins Unendlichdimensionale hinüber retten" können, wenn wir unendliche Summen als Darstellungen bzgl. eines Orthonormalsystems zulassen und somit den Anspruch aufgeben, gleichzeitig auch Vektorraumbasen zu haben. (Letztere haben ja die Eigenschaft, dass für jeden Vektor immer nur endliche Linearkombinationen benötigt werden.)

3.14. Definition: Eine Teilmenge $S \subseteq H$ eines Prähilbertraumes H heißt *Orthonormalsystem*, falls für alle $e, f \in S$ gilt

$$\|e\| = 1, \quad \langle e, f\rangle = 0, \text{ falls } e \neq f.$$

Ein Orthonormalsystem S heißt *vollständig* oder *maximal*, falls für jedes Orthonormalsystem $\widetilde{S} \subseteq H$ mit $S \subseteq \widetilde{S}$ automatisch $S = \widetilde{S}$ gilt.

Orthonormalbasen endlichdimensionaler euklidischer oder unitärer Räume sind stets vollständige Orthonormalsysteme. In der Literatur wird allgemeiner für vollständige Orthonormalsysteme oft die Bezeichnung *Orthonormalbasis* verwendet, auch wenn diese im Unendlichdimensionalen keine Basen sein können. Die Begründung liegt einfach darin, dass damit trotzdem viele Berechnungen der Funktionalanalysis in Hilberträumen fast analog zur Linearen Algebra verlaufen.

3.15. Beispiel: 1) Die Menge $S = \{e_n \mid n \in \mathbb{N}\}$ ist ein vollständiges Orthonormalsystem in l^2: Orthonormalität ist klar. Ist $\widetilde{S} \supseteq S$ ein Orthonormalsystem und $x = (x_n) \in \widetilde{S} \setminus S$, dann folgt $0 = \langle x, e_n\rangle = x_n$ für alle $n \in \mathbb{N}$, d.h. $x = 0$. Also muss $\widetilde{S} = S$ gelten.

2) In $l^2(S)$, wobei S eine beliebige nichtleere Menge ist (vgl. Beispiel 3.4.3)), ist $D := \{\delta_t \mid t \in S\}$ ein vollständiges Orthonormalsystem: Orthonormalität ist klar. Ist $\widetilde{D} \supseteq D$ ein Orthonormalsystem und $f \in \widetilde{D} \setminus D$, dann folgt $0 = \langle f, \delta_t\rangle = f(t)$ für alle $t \in S$, d.h. $f = 0$.

3) Die aus der Analysis bekannten Integralformeln namens „Orthogonalitätsrelationen für trigonometrische Funktionen" ergeben, dass die Menge $S_\mathbb{R} := \{1/\sqrt{2\pi}\} \cup \{c_n \mid n \in \mathbb{N}\} \cup \{s_n \mid n \in \mathbb{N}\}$ mit den Funktionen $c_n(x) := \cos(nx)/\sqrt{\pi}$, $s_n(x) := \sin(nx)/\sqrt{\pi}$ $(x \in [0, 2\pi]$, $n \in \mathbb{N})$ ein Orthomormalsystem im reellen $L^2_\mathbb{R}([0, 2\pi])$ ist; im komplexen $L^2([0, 2\pi])$ ist

$S := \{e_k \mid k \in \mathbb{Z}\}$ mit $e_k(x) := \exp(ikx)/\sqrt{2\pi}$ ein Orthonormalsystem. Die Vollständigkeit von $S_{\mathbb{R}}$ folgt aus jener von S, für die sich mehrere Beweise in der Literatur finden lassen ([Cons16, Section 4.3], [Conw10, Chapter I, Theorem 5.7], [Kab18, Theorem 6.7], [Wer18, Beispiel (b) und (c), Seiten 255-256]).

Das aus der Linearen Algebra bekannte Gram-Schmidt-Verfahren zur Erzeugung eines Orthonormalsystems aus endlich vielen linear unabhängigen Vektoren lässt sich induktiv zumindest auf den Fall abzählbarer Mengen fortsetzen.

3.16. Proposition (Gram-Schmidt-Orthonormalisierung): Sei $X = \{x_1, x_2, \ldots\}$ eine endliche oder abzählbar unendliche linear unabhängige Menge im Prähilbertraum H. Dann lässt sich ein Orthonormalsystem $S = \{e_1, e_2, \ldots\}$ konstruieren, sodass $\operatorname{span}\{e_1, \ldots, e_n\} = \operatorname{span}\{x_1, \ldots, x_n\}$ für alle n gilt.

Beweis. Wir starten mit $e_1 := x_1/\|x_1\|$. Sei bereits ein Orthonormalsystem $\{e_1, \ldots, e_{n-1}\}$ definiert mit $\operatorname{span}\{e_1, \ldots, e_k\} = \operatorname{span}\{x_1, \ldots, x_k\}$ für $k = 1, \ldots, n-1$. Wir setzen

$$v_n := x_n - \sum_{k=1}^{n-1} \langle x_n, e_k \rangle e_k \in \operatorname{span}\{x_1, \ldots, x_n\}$$

und bemerken, dass wegen der linearen Unabhängigkeit von $x_1, \ldots, x_n$ sicher $v_n \neq 0$ gilt. Außerdem erhalten wir für $1 \leq l \leq n-1$

$$\langle v_n, e_l \rangle = \langle x_n, e_l \rangle - \sum_{k=1}^{n-1} \langle x_n, e_k \rangle \underbrace{\langle e_k, e_l \rangle}_{\delta_{kl}} = \langle x_n, e_l \rangle - \langle x_n, e_l \rangle = 0.$$

Daher können wir durch $e_n := v_n/\|v_n\|$ geeignet orthonormal ergänzen und erhalten offensichtlich $\operatorname{span}\{e_1, \ldots, e_n\} \subseteq \operatorname{span}\{x_1, \ldots, x_n\}$. Aus der Induktionsvoraussetzung und $x_n = v_n + \sum_{k=1}^{n-1} \langle x_n, e_k \rangle e_k$ folgt aber ebenso $\operatorname{span}\{x_1, \ldots, x_n\} \subseteq \operatorname{span}\{e_1, \ldots, e_n\}$. $\square$

Einige Klassen von sogenannten *speziellen Funktionen* (Legendre-Polynome, Laguerre- und Hermite-Funktionen usw) entstehen durch Anwendung des Gram-Schmidt-Verfahrens auf linear unabhängige Systeme von einfachen Kombinationen von Polynomen und Exponentialfunktionen (vgl. [Mus14, Section 10.6] oder [Heu06, Aufgaben zu §20]).

3.17. Reihen mit abzählbaren Orthonormalsystemen: Es sei H ein Prähilbertraum und $\{e_n \mid n \in \mathbb{N}\}$ ein Orthonormalsystem in H.

(i) $\forall x \in H\colon \quad \sum_{n=1}^{\infty} |\langle x, e_n \rangle|^2 \leq \|x\|^2$ (Bessel-Ungleichung).

Beweis. Für $m \in \mathbb{N}$ setze $y_m := x - \sum_{n=1}^m \langle x, e_n \rangle e_n$, dann ist $y_m \perp e_n$ für $n = 1, \ldots, m$ und daher auch (beachte $\|e_n\| = 1$ und den Satz von Pythagoras)

$$\sum_{n=1}^m |\langle x, e_n \rangle|^2 \leq \sum_{n=1}^m |\langle x, e_n \rangle|^2 + \|y_m\|^2 = \left\| \sum_{n=1}^m \langle x, e_n \rangle e_n \right\|^2 + \|y_m\|^2$$
$$= \left\| \sum_{n=1}^m \langle x, e_n \rangle e_n + y_m \right\|^2 = \|x\|^2.$$

Nachdem dies für alle $m \in \mathbb{N}$ gilt, ist der Beweis erbracht. □

(ii) $\forall x, y \in H\colon \quad \sum_{n=1}^\infty |\langle x, e_n \rangle \langle e_n, y \rangle| < \infty.$

Das folgt direkt aus der Cauchy-Schwarz-Ungleichung für Reihen zusammen mit (i).

3.18. Erweiterung auf beliebige Orthonormalsysteme: Es sei H ein Prähilbertraum und $S \subseteq H$ ein Orthonormalsystem.

(i) Für jedes $x \in H$ ist die Menge $S_x := \{e \in S \mid \langle x, e \rangle \neq 0\}$ höchstens abzählbar.

Beweis. Sei $x \in H$. Zunächst bemerken wir, dass für jedes $\varepsilon > 0$ die Menge $R_\varepsilon := \{e \in S \mid |\langle x, e \rangle| \geq \varepsilon\}$ endlich ist, denn für $e_1, \ldots, e_m \in S$ mit $|\langle x, e_n \rangle| \geq \varepsilon$ für $n = 1, \ldots, m$ folgt nach der Bessel-Ungleichung

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

also haben wir für m die Beschränkung $m \leq \|x\|^2 / \varepsilon^2$. Wenn wir nun der Reihe nach $\varepsilon = 1, 1/2, 1/3, \ldots$ setzen, dann erhalten wir

$$S_x = \bigcup_{n=1}^\infty R_{1/n}$$

als abzählbare Vereinigung von endlichen Mengen. □

(ii) *Mini-Exkurs über unbedingte Konvergenz von Reihen in einem normierten Raum V:* Sei J eine beliebige nichtleere Indexmenge und $(v_j)_{j \in J}$ eine Familie in V mit der Eigenschaft, dass $J_0 := \{j \in J \mid v_j \neq 0\}$ höchstens abzählbar ist. Falls es $v \in V$ gibt, sodass für jede gewählte Aufzählung $J_0 = \{j_n \mid n \in \mathbb{N}\}$ die Reihe $\sum_{n=1}^\infty v_{j_n}$ gegen v konvergiert, dann heißt $\sum_{j \in J} v_j$ *unbedingt konvergent* gegen v in V und wir schreiben

$$v = \sum_{j \in J} v_j.$$

Bemerkung: Wir können das auch als Konvergenz eines Netzes (w_F) auffassen, indem wir als Indexmenge alle endlichen Teilmengen F von J, gerichtet mit der Mengeninklusion, nehmen und $w_F := \sum_{j \in F} v_j$. (Vgl. den Begriff der *summierbaren Familie* z.B. [Schr00, Definition 1.2.14 und Bemerkung 4 dazu].) Für endlichdimensionales V ist obiger Konvergenzbegriff äquivalent zur *absoluten Konvergenz (bzw. Summierbarkeit)*. Falls V ein Banachraum ist, impliziert zwar die absolute Konvergenz stets die unbedingte (vgl. 1.6(ii)), aber nicht umgekehrt (siehe z.B. Hinweise auf den sogenannten Satz von Dvoretzky-Rogers in [Schr00, Ausblick zu Definition 1.2.14] oder [Wer18, Bemerkung nach Definition V.4.6]).

(iii) $\forall x \in H\colon \quad \sum_{e \in S} |\langle x, e\rangle|^2 \leq \|x\|^2$ (allgemeine Bessel-Ungleichung).

Das folgt direkt aus (i) und 3.17(i) mit dem Konzept aus (ii) angewandt auf den normierten Raum $\mathbb{K}$.

Wir nehmen ab nun immer an, dass H auch *vollständig* ist, also ein Hilbertraum.

3.19. Proposition: Sei S ein Orthonormalsystem im Hilbertraum H, dann gilt:

(i) Für jedes $x \in H$ ist $\sum_{e \in S} \langle x, e\rangle e$ unbedingt konvergent in H,

(ii) $P\colon H \to H$ mit $Px := \sum_{e \in S} \langle x, e\rangle e$ ist die Orthogonalprojektion auf $U := \overline{\operatorname{span} S}$.

Beweis. Für beliebiges fixes $x \in H$ sei $\{e \in S \mid \langle x, e\rangle \neq 0\} = \{e_n \mid n \in \mathbb{N}\}$ eine Aufzählung.

(i): Zunächst folgt aus dem Satz von Pythagoras für beliebige $M, N \in \mathbb{N}$ mit $M \leq N$

$$\left\| \sum_{n=M}^{N} \langle x, e_n\rangle e_n \right\|^2 = \sum_{n=M}^{N} |\langle x, e_n\rangle|^2$$

und somit aus der Bessel-Ungleichung, dass $\sum_{n=1}^{\infty} \langle x, e_n\rangle e_n$ eine Cauchy-Reihe ist. Daher existiert $y := \sum_{n=1}^{\infty} \langle x, e_n\rangle e_n \in H$. Gemäß 3.17(ii) ist für jedes $z \in H$ die Reihe $\sum_{n=1}^{\infty} \langle x, e_n\rangle \langle e_n, z\rangle$ absolut konvergent in $\mathbb{K}$ und somit auch unbedingt konvergent. Wegen der Stetigkeit des Skalarproduktes (im ersten Argument) gilt

$$\langle y, z\rangle = \sum_{n=1}^{\infty} \langle x, e_n\rangle \langle e_n, z\rangle,$$

wobei die rechte Seite für jede Umordnung der Reihensummation gegen denselben Wert konvergiert. Wären wir also oben mit einer anderen Reihenfolge der Aufzählung gestartet und hätten als Grenzwert stattdessen $y' \in H$ erhalten, so muss dennoch $\langle y', z\rangle = \langle y, z\rangle$ für alle $z \in H$ gelten. Das heißt aber $y' - y \in H^\perp = \{0\}$, also $y' = y$ und die unbedingte Konvergenz ist nachgewiesen.

(ii): Wir betrachten die orthogonale Zerlegung $H = U \oplus U^\perp$ und müssen zeigen, dass für $x = u + u_\perp$ mit $u \in U$ und $u_\perp \in U^\perp$ stets $Px = u$ gilt. Es ist klar, dass $Px \in U$ ist und es bleibt zu zeigen, dass $x - Px \in U^\perp$ gilt. Wir beachten, dass $U^\perp = S^\perp$ gemäß 3.6(iv) gilt, daher müssen wir nur $\langle x - Px, e\rangle = 0$ für alle $e \in S$ nachweisen.

Mit obiger Aufzählung dürfen wir $Px = \sum_{n=1}^\infty \langle x, e_n\rangle e_n$ schreiben. Falls $e \in S$ mit $\langle x, e\rangle = 0$ folgt wegen $e \notin \{e_n \mid n \in \mathbb{N}\}$ auch $\langle Px, e\rangle = 0$, also $\langle x - Px, e\rangle = 0$. Falls $e \in S$ mit $\langle x, e\rangle \neq 0$, dann gibt es ein $m \in \mathbb{N}$ mit $e = e_m$ und es folgt

$$\langle x - Px, e\rangle = \langle x - Px, e_m\rangle = \langle x, e_m\rangle - \sum_{n=1}^\infty \langle x, e_n\rangle\langle e_n, e_m\rangle = \langle x, e_m\rangle - \langle x, e_m\rangle = 0.$$

□

3.20. Vollständige Orthonormalsysteme in Hilberträumen: Es sei H ein Hilbertraum und S ein Orthonormalsystem in H.

Proposition: Es gibt ein vollständiges Orthonormalsystem S' von H mit $S \subseteq S'$.

Beweis. Das ist eine Routineanwendung des Lemmas von Zorn, indem wir die Menge M aller Orthonormalsysteme R in H mit der Eigenschaft $R \supseteq S$ mittels Mengeninklusion partiell ordnen und sehen, dass wir für jede Kette durch Vereinigung eine obere Schranke gewinnen. □

Insbesondere kann man im Falle[2] $H \neq \{0\}$ mit einem beliebigen Vektor $e \neq 0$ und $\|e\| = 1$ und dem Orthonormalsystem $S = \{e\} \neq \emptyset$ beginnen und erhält als Resultat die Garantie der Existenz eines vollständigen Orthonormalsystems $S' \neq \emptyset$ von H.

Theorem: Die folgenden Aussagen sind äquivalent:

(i) Das Orthonormalsystem S ist vollständig,

(ii) $\forall x \in H\colon \quad x \in S^\perp \Rightarrow x = 0$,

(iii) $H = \overline{\operatorname{span} S}$, d.h. $\operatorname{span} S$ ist dicht in H,

(iv) $\forall x \in H\colon \quad x = \sum_{e \in S} \langle x, e\rangle e$,

(v) $\forall x, y \in H\colon \quad \langle x, y\rangle = \sum_{e \in S} \langle x, e\rangle\langle e, y\rangle$,

(vi) $\forall x \in H\colon \quad \|x\|^2 = \sum_{e \in S} |\langle x, e\rangle|^2$ (Parseval-Gleichung).

[2]Für $H = \{0\}$ ist $S = \emptyset$ ein maximales Orthonormalsystem.

Beweis. (i) $\Rightarrow$ (ii): Für $x \in S^\perp$ mit $x \neq 0$ wäre OBdA $\|x\| = 1$ und somit $S \cup \{x\}$ eine echte Erweiterung des Orthonormalsystems S.

(ii) $\Rightarrow$ (iii): Nach Korollar 3.9 und 3.6(iv) ist $\overline{\operatorname{span} S} = (S^\perp)^\perp$ und (ii) besagt $S^\perp = \{0\}$, also $(S^\perp)^\perp = H$.

(iii) $\Rightarrow$ (iv): Nach Proposition 3.19(ii) ist $\sum_{e\in S}\langle x, e\rangle e$ die Projektion von x auf $\overline{\operatorname{span} S} = H$, also gleich x.

(iv) $\Rightarrow$ (v): Es ist $y = \sum_{f\in S}\langle y, f\rangle f$ und wegen der Stetigkeit des Skalarprodukts und der Eigenschaft 3.17(ii) folgt mittels $\langle e, f\rangle = \delta_{ef}$ für $e, f \in S$ einfach

$$\langle x, y\rangle = \sum_{e,f\in S}\langle x, e\rangle\overline{\langle y, f\rangle}\langle e, f\rangle = \sum_{e\in S}\langle x, e\rangle\langle e, y\rangle.$$

(v) $\Rightarrow$ (vi): Wir müssen in (v) nur $y = x$ setzen.

(vi) $\Rightarrow$ (i): Wenn S nicht maximal ist, dann gibt es ein $x \in S^\perp$ mit $\|x\| = 1$. Nun folgt $1 = \|x\|^2 = \sum_{e\in S} |\langle x, e\rangle|^2 = 0$, ein Widerspruch. $\square$

Hauptbeispiel der Fourierreihen: Wie bereits in einem Beispiel weiter oben bemerkt, erhalten wir mit den Funktionen $e_n\colon [0, 2\pi] \to \mathbb{C}$, $e_n(x) = \exp(inx)/\sqrt{2\pi}$ ein vollständiges Orthonormalsystem $S = \{e_n \mid n \in \mathbb{Z}\}$ im Hilbertraum $L^2([0, 2\pi])$, dessen Elemente auch geeignet mit 2π-periodischen messbaren Funktionen $f\colon \mathbb{R} \to \mathbb{C}$ identifiziert werden können, für die $|f|^2$ über ein beliebiges Intervall der Periodenlänge integrierbar ist. Die Reihe aus (iv) im obigen Theorem entspricht dann genau der *Fourierreihenentwicklung* $f = \sum_{n\in\mathbb{Z}}\langle f, e_n\rangle e_n$, wobei die Konvergenz eben im Sinne der $\|.\|_2$-Norm garantiert ist, was in klassischer Sprechweise *Konvergenz im quadratischen Mittel* heißt. Die *Fourierkoeffizienten* ergeben sich aus $\langle f, e_n\rangle$ mit den klassischen Integralformeln. (Verschiedene Konventionen bzgl. Periodenlänge und/oder Faktoren im Skalarprodukt sind zu beachten und oft auch lästig.)

Manchmal wird der Begriff *Fourierreihe* auch in allgemeinen Hilberträumen für die Entwicklung in (iv) gebraucht. Außerdem gibt sie auch eine gewisse Rechtfertigung dafür, vollständige Orthonormalsysteme manchmal als *Orthonormalbasen* zu bezeichnen.

Korollar: Folgende Aussagen sind äquivalent:

(i) H ist separabel,

(ii) alle Orthonormalsysteme von H sind höchstens abzählbar,

(iii) es gibt ein endliches oder abzählbares vollständiges Orthonormalsystem.

Beweis. (i) $\Rightarrow$ (ii): Sei S ein Orthonormalsystem von H, dann gilt für alle $e, f \in S$ mit $e \neq f$, dass $\|e - f\|^2 = \langle e - f, e - f\rangle = \langle e, e\rangle + \langle f, f\rangle = 2$. Wäre S überabzählbar, so könnte H nicht separabel sein.

(ii) $\Rightarrow$ (iii): Es gibt nach obiger Proposition zumindest ein vollständiges Orthonormalsystem S, das nach (ii) höchstens abzählbar ist.

(iii) $\Rightarrow$ (i): Dies folgt direkt aus (iii) im obigen Theorem zusammen mit Lemma 1.19. $\square$

Bemerkung: Alle vollständigen Orthonormalsysteme von H haben dieselbe Kardinalität. Für endlichdimensionale euklidische oder unitäre Räume wurde das in der Linearen Algebra bewiesen. Im Falle unendlicher Dimension kann das aber auch leicht mit Hilfe des mengentheoretischen Satzes von Schröder-Bernstein ([Cie97, Theorem 5.1.2 und Remark following Corollary 5.1.4]) bewiesen werden (siehe [Wer18, Lemma V.4.11]).

Für n-dimensionale euklidische oder unitäre Räume wurde in der Linearen Algebra gezeigt, dass diese – vermöge Wahl einer Orthonormalbasis $e_1, \ldots, e_n$ und Koordinatenisomorphismus – isometrisch isomorph zu $\mathbb{K}^n$ mit dem Standardskalarprodukt sind. Dieses Resultat lässt sich mit Hilfe der Räume $l^2(S)$ aus Beispiel 3.4.3) sogar ins Unendlichdimensionale übertragen.

3.21. Hilbertraummodelle bis auf isometrische Isomorphie: Sei H ein Hilbertraum.

Theorem: Sei S ein vollständiges Orthonormalsystem in H, dann gilt $H \cong l^2(S)$.

Beweis. Für $x \in H$ ist die Funktion $\Phi(x)\colon e \mapsto \langle x, e\rangle$ gemäß 3.18(i) und Theorem 3.20 ein Element in $l^2(S)$ und nach der Parseval-Gleichung ist $\|\Phi(x)\|^2 = \|x\|^2$, also ist die Abbildung $\Phi\colon H \to l^2(S)$ isometrisch. Linearität von Φ ist klar, also ist nur noch die Surjektivität zu zeigen.

Sei $f \in l^2(S)$ beliebig vorgegeben. Laut Definition von $l^2(S)$ ist $S_f := \{e \in S \mid f(e) \neq 0\}$ höchstens abzählbar. Außerdem wissen wir, dass $\sum_{e\in S} |f(e)|^2$ konvergent ist. Nun kann mit derselben Argumentation wie im Beweis von Proposition 3.19(i) (mit $f(e)$ an Stelle von $\langle x, e\rangle$) nachgewiesen werden, dass die Reihe $\sum_{e\in S} f(e)e$ in H unbedingt konvergiert; dabei ist nur noch die Zusatzüberlegung nötig, dass für $y := \sum_{n=1}^{\infty} f(e_n)e_n$ jeweils $\langle y, e_m\rangle = \sum_{n=1}^{\infty} f(e_n)\langle e_n, e_m\rangle = f(e_m)$ folgt.

Nun setzen wir schließlich $x := \sum_{e\in S} f(e)e \in H$ und betrachten $\Phi(x) \in l^2(S)$. Für jedes $s \in S$ ist laut Konstruktion

$$\Phi(x)(s) = \langle x, s\rangle = \sum_{e\in S} f(e)\langle e, s\rangle = f(s),$$

d.h. wir erhalten $\Phi(x) = f$. $\square$

Wir machen das Resultat für den separablen Spezialfall nur noch einmal besonders deutlich.

Korollar: Ist H ein unendlichdimensionaler separabler Hilbertraum, dann gilt $H \cong l^2$.

Als weiteren Spezialfall enthält dies eine Version eines klassischen Meilensteins der Fourieranalysis sowie Ursprungs für viele Entwicklungen der Funktionalanalysis.

Korollar (Satz von Riesz-Fischer): Es gilt $L^2([0, 2\pi]) \cong l^2$.

4 Lineare Operatoren auf Hilberträumen

4.1. Adjungierte Operatoren im Sinne der Hilberträume: Zu einer stetigen linearen Abbildung $A\colon H_1 \to H_2$ zwischen Hilberträumen gibt es gemäß 2.32 den adjungierten Operator $A'\colon H_2' \to H_1'$, der einfach durch $y_2' \mapsto y_2' \circ A$ gegeben ist. Mittels der konjugiert linearen isometrischen Bijektionen $F_j\colon H_j \to H_j'$ aus dem Satz von Fréchet-Riesz lässt sich dies aber „auf die Hilberträume selbst zurückspielen“, d.h. wir definieren $A^* := F_1^{-1} \circ A' \circ F_2$ und erhalten eine stetige lineare Abbildung $A^*\colon H_2 \to H_1$, die wir als *Adjungierte (im Sinne der Hilberträume)* von A bezeichnen.

Lemma: $A^* \in L(H_2, H_1)$ ist genau dann adjungiert (im Sinne der Hilberträume) zu $A \in L(H_1, H_2)$, wenn gilt:

$$\forall x \in H_1, \forall y \in H_2\colon \quad \langle Ax, y\rangle_{H_2} = \langle x, A^* y\rangle_{H_1}.$$

Beweis. Einerseits gilt für $A^* := F_1^{-1} \circ A' \circ F_2$ und beliebige $x \in H_1$ und $y \in H_2$ einfach laut Definition

$$\langle x, A^* y\rangle_{H_1} = F_1(A^* y)(x) = A'(F_2(y))(x) = F_2(y)(Ax) = \langle Ax, y\rangle_{H_2}.$$

Sei andererseits $B \in L(H_2, H_1)$ ein Operator mit der Eigenschaft

$$\forall x \in H_1, \forall y \in H_2\colon \quad \langle Ax, y\rangle_{H_2} = \langle x, By\rangle_{H_1}.$$

Dann folgt zunächst

$$\forall x \in H_1\colon \quad F_1(By)(x) = \langle x, By\rangle_{H_1} = \langle Ax, y\rangle_{H_2} = F_2(y)(Ax) = A'(F_2(y))(x),$$

d.h. $F_1(By) = A'(F_2(y))$. Und weil $y \in H_2$ beliebig war, erhalten wir $F_1 \circ B = A' \circ F_2$, also letztlich $B = F_1^{-1} \circ A' \circ F_2 = A^*$. □

Unmittelbare Eigenschaften: Seien H_1, H_2, H_3 Hilberträume und $B, C \in L(H_1, H_2)$, $A \in L(H_2, H_3)$, $\lambda, \mu \in \mathbb{K}$, dann gelten folgende Aussagen:

(i) $(\lambda B + \mu C)^* = \overline{\lambda} B^* + \overline{\mu} C^*$,

(ii) $(AB)^* = B^* A^*$,

(iii) $B^{**} := (B^*)^* = B$,

(iv) $\|B^*\| = \|B\|$,

(v) $\|BB^*\| = \|B^*B\| = \|B\|^2$,

(vi) $\ker B = (\operatorname{im} B^*)^\perp$,

(vii) $\overline{\operatorname{im} B} = (\ker B^*)^\perp$ (vgl. den Satz in 2.32) bzw. $\ker B^* = (\operatorname{im} B)^\perp$.

Hier folgen (i)-(iii) direkt aus der Definition und (iv) aus der Eigenschaft $\|B'\| = \|B\|$ zusammen mit der Isometrie von F_1 und F_2. Beweise für (v)-(vii) als UE.

Beispiel: Wir betrachten den Links- und den Rechtsshift $L, R \in L(l^2)$: Es ist bekanntlich $L(x_1, x_2, x_3, \ldots) := (x_2, x_3, \ldots)$ und $R(x_1, x_2, \ldots) := (0, x_1, x_2, \ldots)$ für $(x_1, x_2, \ldots) \in l^2$ und wir rechnen für beliebige $x, y \in H$ nach, dass

$$\langle Lx, y\rangle = x_2\overline{y_1} + x_3\overline{y_2} + \ldots = x_1\overline{0} + x_2\overline{y_1} + x_3\overline{y_2} + \ldots = \langle x, Ry\rangle$$

gilt. Also ist $L^* = R$ und $R^* = L^{**} = L$.

Bemerkung und kleiner Ausblick auf Anwendungen: Wie schon mal in der Linearen Algebra VO für endlichdimensionale euklidische Räume angedeutet, ist der Begriff des adjungierten Operators sowie ein Konstrukt der Form B^*B usw. ein sehr wichtiges Hilfsmittel vor allem bei der „Lösung unlösbarer Gleichungen“ mit Konzepten wie *verallgemeinerten* oder *Moore-Penrose-Inversen* in unendlichdimensionalen Hilberträumen. Einen raschen Überblick dazu inklusive Hinweisen für Anwendungen auf Bildrekonstruktion oder Tomographie gibt z.B. [Mus14, Section 10.5].

4.2. Unitäre Operatoren: Ein Operator $U \in L(H_1, H_2)$ zwischen Hilberträumen H_1 und H_2 heißt *unitär*, falls

$$U^*U = \mathrm{id}_{H_1} \quad \text{und} \quad UU^* = \mathrm{id}_{H_2}$$

gilt. Äquivalent dazu ist also, dass U (stetig) invertierbar ist und $U^{-1} = U^*$ gilt.

Bemerkung: Anders als im Falle gleicher endlicher Dimension von H_1 und H_2 sind wirklich *beide* Gleichungen $U^*U = \mathrm{id}_{H_1}$ und $UU^* = \mathrm{id}_{H_2}$ zu prüfen, wie das Beispiel von Links- und Rechtsshift $L, R \in L(l^2)$ aus 4.1 lehrt: Es ergibt sich zwar $LL^*(x_1, x_2, \ldots) = L(0, x_1, x_2) = (x_1, x_2, \ldots)$, also $LL^* = \mathrm{id}$, aber $L^*L(x_1, x_2, \ldots) = R(x_2, \ldots) = (0, x_2, \ldots)$, also $L^*L \neq \mathrm{id}$.

Unitarität von $U \in L(H_1, H_2)$ ist weiters durch die folgende Eigenschaft charakterisiert (UE):

$$U \text{ ist surjektiv und } \forall x, y \in H_1 \colon \langle Ux, Uy\rangle_{H_2} = \langle x, y\rangle_{H_1}.$$

Insbesondere erhalten wir mit $x = y$, dass $\|Ux\|^2 = \|x\|^2$ gilt, also *jede unitäre Abbildung eine Isometrie* ist. (Mittels Polarisierungsformeln lässt sich übrigens auch zeigen, dass surjektive Isometrien zwischen Hilberträumen stets unitär sind.)

4.3. Normale Operatoren: Sei H ein Hilbertraum. Ein Operator $A \in L(H)$ heißt *normal*, falls $AA^* = A^*A$ gilt. Wir erkennen, dass jeder unitäre Operator normal ist.

Eine charakterisierende Eigenschaft für Normalität von $A \in L(H)$ ist (Beweis in UE)

$$\forall x, y \in H\colon \quad \langle Ax, Ay \rangle = \langle A^*x, A^*y \rangle.$$

Speziell erhalten wir daraus für $x = y$ die Gleichung $\|Ax\|^2 = \|A^*x\|^2$ und somit die folgende Aussage.

Proposition: Sei $A \in L(H)$ normal, dann gilt $\|A^*x\| = \|Ax\|$ für alle $x \in H$. Insbesondere ist in diesem Fall also $\ker A = \ker A^*$.

4.4. Selbstadjungierte Operatoren: Sei H ein Hilbertraum. Ein Operator $A \in L(H)$ heißt *selbstadjungiert*, falls $A^* = A$ gilt. Jeder selbstadjungerte Operator ist normal.

Ein selbstadjungierter Operator $A \in L(H)$ erfüllt die folgende Symmetriebedingung:

$$\text{(S)} \qquad \forall x, y \in H\colon \quad \langle Ax, y \rangle = \langle x, Ay \rangle.$$

Es lässt sich aber nun sogar zeigen, dass allein aus (S) schon die *Stetigkeit* und Selbstadjungiertheit von A folgt.

Theorem (Satz von Hellinger-Toeplitz): Sei H ein Hilbertraum und $A\colon H \to H$ linear mit der Eigenschaft (S). Dann ist A stetig und selbstadjungiert.

Beweis. Stetigkeit: Wir verwenden den Satz vom abgeschlossenen Graphen und nehmen an, es sei (x_n) eine Folge in H mit $x_n \to x$ und $Ax_n \to y$ in H. Wir müssen $y = Ax$ zeigen. Für jedes $z \in H$ gilt wegen (S)

$$\langle y, z \rangle = \lim \langle Ax_n, z \rangle = \lim \langle x_n, Az \rangle = \langle x, Az \rangle = \langle Ax, z \rangle,$$

daher folgt $y = Ax$. Also ist $A \in L(H)$.

Selbstadjungiertheit: Gemäß Lemma 4.1 folgt nun aus (S) auch $A = A^*$. □

Beispiel: Auf $H = L^2([0,1])$ betrachten wir für $h \in C([0,1]) \subseteq L^\infty([0,1])$ den Multiplikationsoperator $M_h f := hf$. Wenn h *reellwertig* ist, dann ist M_h selbstadjungiert, weil sich aus $\langle M_h f, g \rangle = \int_0^1 (hf)\overline{g} = \int_0^1 f\overline{(hg)} = \langle f, M_h g \rangle$ gerade die Eigenschaft (S) ergibt.

Proposition: Ist $A \in L(H)$ selbstadjungiert, dann gilt

$$\|A\| = \sup_{\|x\| \le 1} |\langle Ax, x \rangle|.$$

Speziell folgt also in dem Fall aus $\langle Ax, x \rangle = 0$ für alle $x \in H$, dass $A = 0$ sein muss.

Beweis. Mittels Cauchy-Schwarz-Ungleichung folgt $|\langle Ax, x\rangle| \leq \|Ax\|\|x\| \leq \|A\|\|x\|^2$, d.h. '$\geq$' in der behaupteten Gleichung. Wir setzen $M := \sup_{\|x\|\leq 1} |\langle Ax, x\rangle|$ und müssen also noch $\|A\| \leq M$ zeigen.

Zunächst bemerken wir, dass für $\lambda > 0$ und $x \in H$ die Parallelogrammgleichung direkt die Relation

$$(*) \qquad \left\|\lambda x + \frac{1}{\lambda}Ax\right\|^2 + \left\|\lambda x - \frac{1}{\lambda}Ax\right\|^2 = 2\lambda^2\|x\|^2 + \frac{2}{\lambda^2}\|Ax\|^2$$

liefert. Weiters bemerken wir, dass wir wegen $\langle A0, 0\rangle = 0$ und $|\langle Ax, x\rangle| = |\langle \frac{Ax}{\|x\|}, \frac{x}{\|x\|}\rangle|\|x\|^2 \leq |\langle \frac{Ax}{\|x\|}, \frac{x}{\|x\|}\rangle|$ für $0 < \|x\| \leq 1$ die Größe M auch wie folgt berechnen können:

$$(**) \qquad M = \sup_{y\neq 0} |\langle \frac{Ay}{\|y\|}, \frac{y}{\|y\|}\rangle|.$$

Außerdem erhalten wir mit einer einfachen Anwendung der Bilinearität des Skalarproduktes und und der Symmetriebedingung (S) auch noch die folgende Relation

$$(***) \quad \langle A\left(\lambda x + \frac{1}{\lambda}Ax\right), \lambda x + \frac{1}{\lambda}Ax\rangle - \langle A\left(\lambda x - \frac{1}{\lambda}Ax\right), \lambda x - \frac{1}{\lambda}Ax\rangle$$
$$= \lambda^2\langle Ax, x\rangle + 2\langle A^2x, x\rangle + \frac{1}{\lambda^2}\langle A^2x, Ax\rangle - \lambda^2\langle Ax, x\rangle + 2\langle A^2x, x\rangle - \frac{1}{\lambda^2}\langle A^2x, Ax\rangle$$
$$= 4\langle Ax, Ax\rangle = 4\|Ax\|^2.$$

Nun setzen wir in der folgenden Abschätzung der Reihe nach $(***)$, $(**)$ und $(*)$ ein, um

$$4\|Ax\|^2 = \langle A\left(\lambda x + \frac{1}{\lambda}Ax\right), \lambda x + \frac{1}{\lambda}Ax\rangle - \langle A\left(\lambda x - \frac{1}{\lambda}Ax\right), \lambda x - \frac{1}{\lambda}Ax\rangle$$
$$\leq M\left\|\lambda x + \frac{1}{\lambda}Ax\right\|^2 + M\left\|\lambda x - \frac{1}{\lambda}Ax\right\|^2 = 2M\lambda^2\|x\|^2 + 2M\frac{1}{\lambda^2}\|Ax\|^2$$

zu erhalten.

Im Falle $Ax \neq 0$ (daher auch $x \neq 0$ und $M > 0$) wird mit $\lambda := \|Ax\|/(\|x\|\sqrt{M})$ aus obiger Ungleichung $4\|Ax\|^2 \leq 2\|Ax\|^2 + 2M^2\|x\|^2$, daher auch $\|Ax\| \leq M\|x\|$.
Im Falle $Ax = 0$ ist $\|Ax\| \leq M\|x\|$ sowieso richtig.

Zusammenfassend gilt also $\|Ax\| \leq M\|x\|$ für alle $x \in H$ und daher auch $\|A\| \leq M$. $\square$

Bemerkung: Sei H ein *komplexer* Hilbertraum, also $\mathbb{K} = \mathbb{C}$, und $A \in L(H)$ selbstadjungiert, dann folgt wegen $\overline{\langle Ax, x\rangle} = \langle x, Ax\rangle = \langle Ax, x\rangle$, dass $\langle Ax, x\rangle \in \mathbb{R}$ für jedes $x \in H$ gilt. Es lässt sich sogar recht elementar zeigen, dass $A \in L(H)$ genau dann selbstadjungiert ist, wenn $\langle Ax, x\rangle \in \mathbb{R}$ für jedes $x \in H$ gilt. (Beweis wie in einer entsprechenden Aufgabe der Linearen Algebra UE im SoSe 19 bzw. auch in [Wer18, Satz V.5.6].)

4.5. Theorem: Sei H ein Hilbertraum und $P \in L(H)$ ein Projektor, d.h. $P^2 = P$. Dann sind folgende Aussagen äquivalent:

(i) P ist eine Orthogonalprojektion (nämlich *die* Orthogonalprojektion auf $\operatorname{im} P$),

(ii) P ist selbstadjungiert.

Beweis. (i) $\Rightarrow$ (ii) folgt aus Theorem 3.10, weil dort (S) nachgewiesen wurde.

(ii) $\Rightarrow$ (i): Zunächst ist wegen $P^2 = P$ der Teilraum $\operatorname{im} P$ abgeschlossen, denn falls $Px_n \to y$ in H gilt, dann folgt $Py = P(\lim Px_n) = \lim P^2 x_n = \lim Px_n = y$, also $y = Py \in \operatorname{im} P$.

Weiters erhalten wir aus 4.1(vi) mit der Selbstadjungiertheit, dass $\ker P = (\operatorname{im} P^*)^\perp = (\operatorname{im} P)^\perp$ gilt. Somit folgt bereits $Px = u$ für die eindeutige Zerlegung[1] $x = u + u_\perp$ mit $u \in \operatorname{im} P$, $u_\perp \in (\operatorname{im} P)^\perp$, also dass P die Orthogonalprojektion auf $\operatorname{im} P$ ist. □

Wir werfen im abschließenden Teil dieses Kapitels noch einen kurzen Blick auf die Spektraltheorie für spezifische Klassen von Operatoren auf Hilberträumen.

4.6. Einfache spektrale Eigenschaften: Es sei hier stets H ein Hilbertraum, $A \in L(H)$ und $\lambda, \mu \in \mathbb{K}$. Wir erinnern daran, dass allgemein $\sigma(A) \subseteq \{z \in \mathbb{K} \mid |z| \leq \|A\|\}$ und kompakt ist.

(i) $\sigma(A^*) = \overline{\sigma(A)}$, d.h. $\lambda \in \sigma(A) \Leftrightarrow \overline{\lambda} \in \sigma(A^*)$.

Wir müssen für den Beweis nur anmerken, dass $(\lambda - A)^* = \overline{\lambda} - A^*$ sowie für einen (stetig) invertierbaren Operator $B \in L(H)$ die Relation $(B^{-1})^* = (B^*)^{-1}$ gilt. Somit ist die (stetige) Invertierbarkeit von $\lambda - A$ äquivalent zu jener von $\overline{\lambda} - A^*$.

(ii) Sei A normal. Ist $v \neq 0$ ein Eigenvektor zum Eigenwert λ von A, dann ist v auch Eigenvektor von A^* mit Eigenwert $\overline{\lambda}$. Ist $w \neq 0$ ein Eigenvektor zum Eigenwert μ von A und $\mu \neq \lambda$, dann ist $w \perp v$.

Es ist dann nämlich auch $\lambda - A$ normal und somit Proposition 4.3 mit $(\lambda - A)^* = \overline{\lambda} - A^*$ anwendbar, d.h. $\left\|(\overline{\lambda} - A^*)v\right\| = \|(\lambda - A)v\| = 0$, also $A^*v = \overline{\lambda}v$. Die zweite Aussage folgt nun aus

$$\lambda\langle v, w\rangle = \langle \lambda v, w\rangle = \langle Av, w\rangle = \langle v, A^* w\rangle = \langle v, \overline{\mu} w\rangle = \mu\langle v, w\rangle,$$

weil dies nur für $\langle v, w\rangle = 0$ haltbar ist.

(iii) Ist A selbstadjungiert und kompakt, dann gilt $\sigma(A) \subseteq \mathbb{R}$ und zumindest $-\|A\|$ oder $\|A\|$ ist ein Eigenwert von A, d.h. $\{-\|A\|, \|A\|\} \cap \sigma_p(A) \neq \emptyset$.

[1] Erinnerung an die Lineare Algebra: $P^2 = P$ impliziert $Pu = u$ für alle $u \in \operatorname{im} P$, denn mit $u = Pv$ folgt $Pu = P(Pv) = P^2 v = Pv = u$.

Beweis. Nach Korollar 2.45 besteht $\sigma(A) \setminus \{0\}$ nur aus Eigenwerten; und falls $\lambda \in \sigma_p(A)$, dann gibt es $v \neq 0$ mit $Av = \lambda v$ und es folgt

$$\lambda\langle v, v\rangle = \langle Av, v\rangle = \langle v, Av\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle,$$

was wegen $\langle v, v\rangle > 0$ natürlich $\overline{\lambda} = \lambda$ nach sich zieht, also ist λ reell.

Gemäß Proposition 4.4 gilt $\|A\| = \sup_{\|x\|\leq 1} |\langle Ax, x\rangle|$, weshalb es eine Folge (x_n) in $K_1(0)$ gibt mit $\|A\| = \lim |\langle Ax_n, x_n\rangle|$. Wegen der Kompaktheit von A gibt es eine Teilfolge (x_{n_k}), sodass $Ax_{n_k} \to y$ in H gilt. OBdA gilt auch $\langle Ax_{n_k}, x_{n_k}\rangle \to \mu$ in $\mathbb{K}$ (weil die ursprüngliche Zahlenfolge $(\langle Ax_n, x_n\rangle)_{n\in\mathbb{N}}$ beschränkt in $\mathbb{K}$ ist). Beachte, dass wegen $\langle Ax, x\rangle \in \mathbb{R}$ auch μ reell ist und weiters laut Konstruktion $|\mu| = \|A\|$ gilt. Wir erhalten

$$\begin{aligned}\|Ax_{n_k} - \mu x_{n_k}\|^2 = \langle Ax_{n_k} - \mu x_{n_k}, Ax_{n_k} - \mu x_{n_k}\rangle = \|Ax_{n_k}\|^2 - 2\mu\langle Ax_{n_k}, x_{n_k}\rangle + \mu^2\|x_{n_k}\|^2 \\ \leq \|A\|^2\|x_{n_k}\|^2 - 2\mu\langle Ax_{n_k}, x_{n_k}\rangle + \mu^2 \leq \|A\|^2 - 2\mu\langle Ax_{n_k}, x_{n_k}\rangle + \mu^2 \\ = \mu^2 - 2\mu\langle Ax_{n_k}, x_{n_k}\rangle + \mu^2 \to 0,\end{aligned}$$

also wegen $Ax_{n_k} \to y$ auch $\mu x_{n_k} \to y$. Daher folgt $Ay = \lim \mu Ax_{n_k} = \mu y$, also ist μ ein Eigenwert von A mit $|\mu| = \|A\|$, falls $y \neq 0$ gilt. Aber die Annahme $y = 0$ führt zu $Ax_{n_k} \to 0$ und daher zu $\|A\| = \lim |\langle Ax_{n_k}, x_{n_k}\rangle| = 0$, was $A = 0$ bedeutet, in welchem Fall aber die Aussage ohnehin richtig ist. Für $A \neq 0$ ist also $y \neq 0$ und somit ein Eigenvektor. □

Beispiel: Sei $h(t) := t$ für $t \in [0, 1]$, dann ist $h \in C([0, 1])$ und reellwertig, also der zugehörige Multiplikationsoperator M_h gemäß Beispiel 4.4 selbstadjungiert auf $L^2([0, 1)]$ und man sieht leicht, dass $\|M_h\| = 1$ gilt. Für $\lambda \in \mathbb{C} \setminus [0, 1]$ ergibt die Multiplikation mit der Funktion $1/(\lambda - t)$ eine stetige Inverse zu $\lambda - M_h$ und für $\lambda \in [0, 1]$ kann $\lambda - M_h$ nicht invertierbar sein, weil z.B. $(\lambda - t)f(t) = 1$ nicht für fast alle t mit $f \in L^2([0, 1])$ gelten kann[2]. Daher ist $\sigma(M_h) = [0, 1]$. Aber M_h hat keine Eigenwerte, d.h. $\sigma_p(M_h) = \emptyset$, weil für $0 \leq \lambda \leq 1$ die Gleichung $(\lambda - t)f(t) = 0$ f.ü. stets $f(t) = 0$ f.ü. impliziert, also $f = 0$ in $L^2([0, 1])$ sein muss. Nach obiger Eigenschaft (iii) kann daher M_h nicht kompakt sein, weil sonst 1 ein Eigenwert sein müsste. Alternativ könnten wir die Nichtkompaktheit auch aus Korollar 2.46(i) ableiten, wonach ja für kompaktes $A \in L(H)$ die Menge $\sigma(A) \setminus \{0\}$ höchstens abzählbar ist.

Wir kommen nun zu Verallgemeinerungen von Diagonalisierbarkeitssätzen aus der Linearen Algebra. Im Endlichdimensionalen hat ja ein selbstadjungierter Operator A bezüglich einer Orthonormalbasis $e_1, \ldots, e_n$ Diagonalgestalt, falls seine Wirkung auf einen Vektor x mit gewissen reellen Zahlen $\lambda_1, \ldots, \lambda_n$ (nämlich den Eigenwerten mit Vielfachheiten) in der Form $Ax = \sum_{k=1}^n \lambda_k\langle x, e_k\rangle e_k$ geschrieben werden kann.

4.7. Theorem (Spektralsatz A für kompakte selbstadjungierte Operatoren): Sei H ein Hilbertraum und $0 \neq A \in K(H)$ selbstadjungiert. Dann existiert ein höchstens

[2] Und daher also $1 \notin \operatorname{im}(\lambda - M_h)$

abzählbares Orthonormalsystem $S = \{e_k \mid k \in M\}$ (d.h. $M \subseteq \mathbb{N}$) bestehend aus Eigenvektoren zu A zu den Eigenwerten $\lambda_1, \lambda_2, \ldots$ in $\mathbb{R} \setminus \{0\}$, die endlich viele sind oder gegen 0 konvergieren, sodass $H = \ker A \oplus \overline{\operatorname{span} S}$, $\|A\| = \max\{|\lambda_k| \mid k \in M\}$ und

$$\forall x \in H\colon \quad Ax = \sum_{k \in M} \lambda_k \langle x, e_k \rangle e_k.$$

Beweis. Nach 2.45 und 2.46 besteht die Menge $\sigma(A) \setminus \{0\}$ nur aus höchstens abzählbar vielen paarweise verschiedenen Eigenwerten $\mu_1, \mu_2, \ldots$, die sich höchstens bei 0 häufen können und endlichdimensionale Eigenräume $\ker(\mu_j - A)$ besitzen. Gemäß 4.6(iii) sind alle μ_j reell und gemäß 4.6(ii) gilt $\ker(\mu_r - A) \perp \ker(\mu_j - A)$ für $r \neq j$. Es sei $d_j := \dim(\mu_j - A)$ und die Folge $\lambda_1, \lambda_2, \ldots$ so konstruiert, dass zunächst μ_1 genau d_1-mal wiederholt wird, dann μ_2 genau d_2-mal usw. Die für die Aufzählung der λ_k verwendete Indexmenge sei M. Entsprechend wählen wir Orthonormalbasen in jedem Eigenraum $\ker(\mu_j - A)$ und fügen diese zu einem Orthonormalsystem $S := \{e_1, e_2, \ldots\}$ zusammen, indem zunächst die d_1 Vektoren der ersten Basis genommen werden, dann die d_2 Vektoren der zweiten Basis usw. Mit dieser Konvention erhalten wir $Ae_k = \lambda_k e_k$ für jedes $k \in M$.

Im Falle $\ker A = \{0\}$ ist $\ker A \perp e_k$ trivial und andernfalls ist 0 auch ein Eigenwert von A und es gilt $\ker A \perp e_k$ nach 4.6(ii) für jedes $k \in M$. Also gilt $\ker A \perp \overline{\operatorname{span} S}$.

Behauptung 1: $\ker A \oplus \overline{\operatorname{span} S} = H$

Setze $U := (\ker A \oplus \overline{\operatorname{span} S})^\perp$, dann ist die Behauptung äquivalent zu $U = \{0\}$.
Sei $y \in U$, dann ist insbesondere $y \perp \ker A$ und $y \perp e_k$ für jedes $k \in M$. Wegen $y \perp \ker A$ ist für jedes $x \in \ker A$ auch $\langle Ay, x \rangle = \langle y, Ax \rangle = \langle y, 0 \rangle = 0$, also $Ay \perp \ker A$. Wegen $y \perp e_k$ ist $\langle Ay, e_k \rangle = \langle y, Ae_k \rangle = \langle y, \lambda_k e_k \rangle = \lambda_k \langle y, e_k \rangle = 0$, also auch $Ay \perp e_k$. Zusammen folgt $Ay \in (\ker A \oplus \overline{\operatorname{span} S})^\perp = U$. Somit haben wir $A(U) \subseteq U$ nachgewiesen, d.h. die Einschränkung von A auf U kann als kompakter selbstadjungierter Operator $B \in L(U)$ aufgefasst werden. Angenommen $U \neq \{0\}$, dann gilt auch $B \neq 0$ (weil $U \perp \ker A$). Nach 4.6(iii) besitzt B einen Eigenwert $\lambda \neq 0$ mit zugehörigem Eigenvektor $0 \neq u \in U$. Es müsste daher auch $\lambda \in \sigma(A) \setminus \{0\} \subseteq \sigma_p(A)$ gelten und somit $u \in \operatorname{span} S \subseteq U^\perp$, ein Widerspruch. Also ist $B = 0$ und somit $U = \{0\}$.

Behauptung 2: $Ax = \sum_{k \in M} \lambda_k \langle x, e_k \rangle e_k$ für jedes $x \in H$.

Nach Behauptung 1 kann jedes $x \in H$ eindeutig zerlegt werden in der Form $x = y + \sum_{k \in M} \langle x, e_k \rangle e_k$ mit $y \in \ker A$. Anwendung von A darauf ergibt nun

$$Ax = Ay + \sum_{k \in M} \langle x, e_k \rangle Ae_k = 0 + \sum_{k \in M} \langle x, e_k \rangle \lambda_k e_k = \sum_{k \in M} \lambda_k \langle x, e_k \rangle e_k.$$

Nachdem $|\lambda| \leq \|A\|$ für $\lambda \in \sigma(A)$ allgemein gilt, erhalten wir zusammen mit 4.6(iii) schließlich auch die Gleichung $\|A\| = \max\{|\mu_1|, |\mu_2|, \ldots\} = \max\{|\lambda_k| \mid k \in M\}$. □

4.8. Korollar (Spektralsatz B für kompakte selbstadjungierte Operatoren): Sei H ein Hilbertraum, $A \in K(H)$ selbstadjungiert und $\{\mu_k \mid k \in M_0\}$ die endliche oder abzählbare Menge (d.h. $M_0 \subseteq \mathbb{N}$) der paarweise verschiedenen reellen Eigenwerte von A ungleich 0. Für jedes $k \in M_0$ ist der Eigenraum $\ker(\mu_k - A)$ endlichdimensional und es bezeichne P_k die Orthogonalprojektion darauf. Dann gilt im Sinne der Konvergenz in der Operatornorm

$$A = \sum_{k \in M_0} \mu_k P_k.$$

Beweis. Wegen der Darstellung im Theorem 4.7 wissen wir bereits, dass die Summe punktweise auf H konvergiert. Für endliches M_0 ist die Aussage des Korollars trivial, also dürfen wir $M_0 = \mathbb{N}$ annehmen und somit auch $\mu_k \to 0$. Wir betrachten die Partialsummen $S_m := \sum_{k=1}^{m} \mu_k P_k$ und bemerken, dass für $n > m$ der Operator $S_n - S_m = \sum_{k=m+1}^{n} \mu_k P_k$ eine endliche reelle Linearkombination der endlichdimensionalen selbstadjungierten Operatoren $P_{m+1}, \ldots, P_n$ ist und somit selbst kompakt und selbstadjungiert. Daher ist $\|S_n - S_m\|$ gemäß 4.6(iii) gleich dem betragsmäßig größten Eigenwert. Letzterer ist $\max_{m+1 \le k \le n} |\mu_k|$. Daher folgt für $n > m$ stets

$$\|S_n - S_m\| \le \max_{m+1 \le k \le n} |\mu_k| \to 0 \quad (m, n \to \infty),$$

also ist (S_n) eine Cauchyfolge im Banachraum $L(H)$. Weil der gleichmäßige Limes gleich dem punktweisen sein muss, konvergiert (S_n) tatsächlich gegen A. □

Wir hatten in 2.35 gesehen, dass allgemein in normierten Räumen der Operatornorm-Limes einer Folge von endlichdimensionalen Operatoren stets kompakt ist. Das obige Korollar zeigt also nun teilweise umgekehrt, dass selbstadjungierte kompakte Operatoren auf Hilberträumen immer als Limes von endlichdimensionalen Operatoren geschrieben werden können. Diese Aussage lässt sich recht mühelos auf alle kompakten Operatoren auf einem Hilbertraum ausdehnen (vgl. z.B. [Wer18, Korollar VI.3.7]) und die entsprechende Eigenschaft wird *kompakte Approximationseigenschaft* genannt. Für Banachräume ist dies im Allgemeinen nicht mehr richtig, wobei Gegenbeispiele schwer zu finden waren (vgl. dazu auch die historischen Anmerkungen in [Wer18, Abschnitt II.6]).

4.9. Beispiel: Es sei $H = L^2([0,1])$ und A der Integraloperator

$$\forall f \in L^2([0,1]): \quad (Af)(s) = \int_0^1 k(s,t) f(t)\, dt \quad (s \in [0,1])$$

mit Kernfunktion $k \in C([0,1] \times [0,1])$, die reellwertig ist und symmetrisch, d.h. $k(t,s) = k(s,t)$ für alle $s, t \in [0,1]$. In den UE wird/wurde gezeigt, dass $\|A\| \le \|k\|_\infty$ gilt und A *selbstadjungiert* ist. Wir skizzieren nun kurz, warum A auch *kompakt* ist: Auf $[0,1]^2$ kann nämlich k gleichmäßig durch eine Folge von reellen Polynomfunktionen (p_n) in zwei Variablen approximiert werden, wobei der Grad von p_n höchstens n sein soll. (Dies folgt aus dem Satz von Stone-Weierstraß, siehe z.B. [Wer18, Satz VIII.4.7] oder [Hoe20, 8.10]). Es bezeichne A_n den Integraloperator mit der Kernfunktion p_n. Es ist $p_n(s,t)$ eine endliche reelle Linearkombination von Monomen $s^k t^l$ mit $k + l \leq n$. Daher ist leicht zu sehen, dass $\mathrm{im}(A_n)$ enthalten ist in der $\mathbb{C}$-linearen Hülle der Polynomfunktionen auf $[0,1]$ vom Grad höchstens n, also ein endlichdimensionaler Teilraum von $L^2([0,1])$ ist. Wir erhalten wegen $\|A - A_n\| \leq \|k - p_n\|_\infty$ somit eine Approximation von A durch Operatoren mit endlichdimensionalem Bild bzgl. der Operatornorm, woraus nach Lemma 2.35 nun die Kompaktheit von A folgt.

Diese Klasse von Beispielen bzw. deren Erweiterung auf Integraloperatoren mit Kernfunktionen $k \in L^2([0,1] \times [0,1])$ öffnet die Türen für eine Fülle von klassischen Anwendungen auf Integralgleichungen und Randwertprobleme für Differentialgleichungen, mehr dazu z.B. in [Schr00, Appendix 1.A], [Heu06, Kapitel VI und XII], [Wer18, Abschnitte VI.4].

4.10. Das Spektrum jedes selbstadjungierten Operators ist reell: Sei H ein *komplexer* Hilbertraum, d.h. $\mathbb{K} = \mathbb{C}$, und $A \in L(H)$. Wir werden hier die erste Aussage in 4.6(iii), nämlich $\sigma(A) \subseteq \mathbb{R}$, vom Fall eines kompakten selbstadjungierten Operators auf die Klasse aller selbstadjungierten Operatoren ausdehnen.

Lemma: Es sei der *numerische Wertebereich* von A durch

$$W(A) := \{\langle Ax, x\rangle \mid x \in H, \|x\| = 1\}$$

definiert. Dann ist $W(A) \subseteq \mathbb{C}$ beschränkt und es gilt $\sigma(A) \subseteq \overline{W(A)}$.

Beweis. Die Beschränktheit von $W(A)$ folgt aus $|\langle Ax, x\rangle| \leq \|Ax\| \|x\| \leq \|A\| \|x\|^2$. Sei nun $\lambda \in \mathbb{C} \setminus \overline{W(A)}$ und $d := \inf\{|\lambda - \mu| \mid \mu \in W(A)\}$, dann ist wegen der Kompaktheit von $\overline{W(A)}$ sicherlich $d > 0$. Für $x \in H$ mit $\|x\| = 1$ gilt also

$$0 < d\|x\| = d \leq |\lambda - \langle Ax, x\rangle| = |\langle(\lambda - A)x, x\rangle| \leq \|(\lambda - A)x\| \|x\| = \|(\lambda - A)x\|.$$

Dies zeigt, dass $\lambda - A$ injektiv ist und eine Inverse $(\lambda - A)^{-1}$ als Abbildung $\mathrm{im}(\lambda - A) \to H$ mit Operatornorm höchstens $1/d$ besitzt, also ein Isomorphismus normierter Räume ist. Aus der Vollständigkeit von H folgt daher jene von $\mathrm{im}(\lambda - A)$ und somit auch dessen Abgeschlossenheit. Wir überlegen unten abschließend noch, dass $\mathrm{im}(\lambda - A)$ auch dicht in H ist, dann folgt $\mathrm{im}(\lambda - A) = H$ und daher $\lambda \in \mathbb{C} \setminus \sigma(A)$, was den Beweis beendet.

Ist $\mathrm{im}(\lambda - A)$ nicht dicht, dann gibt es ein $z \in (\mathrm{im}(\lambda - A))^\perp$ mit $\|z\| = 1$ und es folgt insbesondere

$$0 = \langle(\lambda - A)z, z\rangle = \lambda\langle z, z\rangle - \langle Az, z\rangle = \lambda - \langle Az, z\rangle,$$

d.h. $\lambda \in W(A)$ im Widerspruch zur ursprünglichen Wahl $\lambda \in \mathbb{C} \setminus \overline{W(A)}$. □

Aus obigem Lemma zusammen mit der Bemerkung in 4.4 erhalten wir unmittelbar das folgende Resultat.

Theorem: Wenn $A \in L(H)$ selbstadjungiert ist, dann gilt $\sigma(A) \subseteq \mathbb{R}$. Ist A ein *positiver* Operator, d.h. $\langle Ax, x\rangle \geq 0$ für alle $x \in H$, dann folgt $\sigma(A) \subseteq [0, \infty[$.

Bemerkung: Noch genauer erhalten wir für selbstadjungiertes A sogar $\sigma(A) \subseteq [m, M]$ mit $m := \inf W(A)$ und $M := \sup W(A)$.

4.11. Spektralradius: Sei H ein *komplexer* Hilbertraum, d.h. $\mathbb{K} = \mathbb{C}$, und $A \in L(H)$. Wir definieren den *Spektralradius* von A durch $r(A) := \max\{|\lambda| \mid \lambda \in \sigma(A)\}$. Das ist also der Radius der kleinsten abgeschlossenen Kreisscheibe in $\mathbb{C}$ um 0, die das Spektrum von A enthält, und klarerweise gilt immer $r(A) \leq \|A\|$. Wir zeigen hier zwei wichtige Eigenschaften des Spektralradius.

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

Beweis. Wir erinnern daran, dass die Resolvente auf der offenen Menge $\varrho(A) \subseteq \mathbb{C}$ lokal immer in eine Potenzreihe (mit Koeffizienten in $L(H)$) entwickelt werden kann. Speziell haben wir natürlich für $|\mu| > \|A\| \geq r(A)$ die Neumannsche Reihe

$$(\mu - A)^{-1} = \frac{1}{\mu} \sum_{n=0}^{\infty} A^n \left(\frac{1}{\mu}\right)^n$$

als Potenzreihe in $z := 1/\mu$ und analog zu den Überlegungen in der Komplexen Analysis ergibt sich die Konvergenz obiger Reihenentwicklung dann sogar für $|\mu| > r(A)$, d.h. $|z| < 1/r(A)$, während für ein gewisses $\mu_0 \in \sigma(A)$ mit $|\mu_0| = r(A)$ sicherlich Divergenz vorliegt. Deshalb ist der Konvergenzradius R der Potenzreihe in z gerade $1/r(A)$ (was wir im Falle $r(A) = 0$ als ∞ lesen) und die klassische Formel für R ergibt zunächst

$$r(A) = \limsup \|A^n\|^{1/n}.$$

Behauptung: Für $\lambda \in \sigma(A)$ und $n \in \mathbb{N}$ gilt $|\lambda| \leq \|A^n\|^{1/n}$.

Wegen $\lambda^n - A^n = (\lambda - A)(A^{n-1} + \lambda A^{n-2} + \ldots + \lambda^{n-1})$ ist nämlich $\lambda^n \in \sigma(A^n)$, daher $|\lambda|^n = |\lambda^n| \leq \|A^n\|$ und weiter $|\lambda| \leq \|A^n\|^{1/n}$.

Aus der Behauptung folgt nun die Ungleichung $r(A) \leq \liminf \|A^n\|^{1/n}$ und daher insgesamt

$$r(A) \leq \liminf \|A^n\|^{1/n} \leq \limsup \|A^n\|^{1/n} = r(A),$$

weshalb die Konvergenz von $(\|A^n\|^{1/n})_{n\in\mathbb{N}}$ gegen $r(A)$ folgt. □

(ii) Ist A normal, dann gilt $\|A\| = r(A)$ und somit auch $\|A\| = \max\{|\lambda| \mid \lambda \in \sigma(A)\}$.

Beweis. Wegen der Normalität von A ist $(A^*A)^m = (A^m)^*A^m$ für alle $m \in \mathbb{N}$ sowie $\|(A^*A)^2\| = \|(A^*A)^*(A^*A)\| = \|A^*A\|^2$ und induktiv auch $\|A^*A\|^{2^m} = \left\|(A^*A)^{2^m}\right\|$. Daher gilt

$$\|A\|^{2^n} = \left(\|A\|^2\right)^{2^{n-1}} = \|A^*A\|^{2^{n-1}} = \left\|(A^*A)^{2^{n-1}}\right\| = \left\|\left(A^{2^{n-1}}\right)^* A^{2^{n-1}}\right\| = \left\|A^{2^{n-1}}\right\|^2$$

und somit $\|A\| = \left\|A^{2^{n-1}}\right\|^{1/2^{n-1}} \to r(A)$ für $n \to \infty$. □

4.12. Stetiger Funktionalkalkül für selbstadjungierte Operatoren: Es sei H ein komplexer Hilbertraum und $T \in L(H)$ selbstadjungiert. Wir wissen aus 4.10, dass $\sigma(T) \subseteq I$ für ein reelles kompaktes Intervall $I \subseteq \mathbb{R}$ gilt. Wir wollen hier für jedes $f \in C(\sigma(T))$ einen Operator $f(T) \in L(H)$ erklären.

Den Anfang könnten wir mit Polynomen $p \in \mathbb{C}[t]$ machen (wobei wir aber die Variable t eigentlich „nur reell" betrachten werden), indem wir für $p(t) = a_0 + a_1 t + \ldots + a_n t^n$ mit $a_j \in \mathbb{C}$ $(j = 0, \ldots, n)$ natürlich $p(T) := a_0 + a_1 T + \ldots + a_n T^n$ setzen. Weil die Monome T^l und T^k vertauschen, ergibt sich daraus schon mal ein Algebrenhomomorphismus

$$\Phi_0 \colon \mathbb{C}[t] \to L(H), p \mapsto p(T),$$

d.h. Φ_0 ist $\mathbb{C}$-linear und multiplikativ, also $(pq)(T) = p(T)q(T)$ für $p, q \in \mathbb{C}[t]$. Die Grundidee für das Weitere ist dann gleichmäßige Approximation (nach dem Satz von Weierstraß) einer stetigen Funktion f durch eine Folge von Polynomen p_n und „Übersetzung" in $f(T) := \lim p_n(T)$ im Sinne der Operatornorm.

Wegen $\sigma(T) \subseteq I$ ergibt jedes Element in $C(I)$ durch Einschränkung eine Funktion in $C(\sigma(T))$. Andererseits lässt sich gemäß Ausdehnungssatz von Tietze ([Hoe20, Theorem 6.10]) jede auf $\sigma(T)$ definierte stetige Funktion zu einer stetigen Funktion auf I ausdehnen. Es gibt aber ein technisches Ärgernis, weil selbst für zwei Polynome $p \neq q$ die Einschränkungen der Polynomfunktionen auf $\sigma(T)$ identisch sein können (z.B. ist nicht ausgeschlossen, dass $\sigma(T)$ für gewisse T eine endliche diskrete Menge ist). Daher ist a priori nicht einmal klar, ob die Zuordnung $p|_{\sigma(T)} \mapsto p(T)$ wohldefiniert ist. Klärung bringt aber das folgende

Lemma: Für jedes $p \in \mathbb{C}[t]$ gilt

(i) $\sigma(p(T)) = \{p(\lambda) \mid \lambda \in \sigma(T)\} = p(\sigma(T))$,

(ii) $\|p(T)\| = \sup_{\lambda \in \sigma(T)} |p(\lambda)| = \left\| p|_{\sigma(T)} \right\|_\infty$.

Beweis. (i): Dies ist klar für konstante Polynome, also nehmen wir gleich $\deg(p) \geq 1$ an.

<u>Behauptung $\sigma(p(T)) \subseteq p(\sigma(T))$</u>: Sei $\mu \in \sigma(p(T))$. Es muss $\deg(p - \mu) \geq 1$ sein, also gibt es eine Faktorisierung $p(t) - \mu = c(t - \lambda_1) \cdots (t - \lambda_m)$ mit $\lambda_1, \ldots, \lambda_m \in \mathbb{C}$ und $c \in \mathbb{C} \setminus \{0\}$. Daraus folgt die Gleichung $p(T) - \mu = c(T - \lambda_1) \cdots (T - \lambda_m)$ für Operatoren. Für mindestens ein j muss $\lambda_j \in \sigma(T)$ sein, denn andernfalls wäre $p(T) - \mu$ nach dieser Operatorleichung invertierbar. Es folgt somit $p(\lambda_j) - \mu = 0$, d.h. $\mu = p(\lambda_j) \in p(\sigma(T))$.

<u>Wir zeigen noch $p(\sigma(T)) \subseteq \sigma(p(T))$</u>: Sei $\lambda \in \sigma(T)$. Dann ist λ eine Nullstelle von $p(t) - p(\lambda)$, also gibt es ein $q \in \mathbb{C}[t]$, sodass $p(t) - p(\lambda) = (t - \lambda)q(t)$ gilt. Daraus ergibt sich $p(T) - p(\lambda) = (T - \lambda)q(T)$ als Gleichung in $L(H)$. Nun kann $p(T) - p(\lambda)$ nicht invertierbar sein, weil andernfalls $(T - \lambda)q(T)$ invertierbar wäre und somit insbesondere auch $\lambda - T$. Also ist $p(\lambda) \in \sigma(p(T))$.

(ii): Offensichtllich ist Φ_0 auch *involutiv*, d.h. $\overline{p}(T) = p(T)^*$ und wir erhalten

$$\|p(T)\|^2 = \|p(T)^* p(T)\| = \|\overline{p}(T)p(T)\| = \|(\overline{p}p)(T)\|.$$

Nachdem $(\overline{p}p)(T)$ selbstadjungiert ist, dürfen wir 4.11(ii) anwenden, sodass weiter

$$\|p(T)\|^2 = \sup\{|\mu| \mid \mu \in \sigma((\overline{p}p)(T))\}$$

folgt. Wegen (i) gilt $\sigma((\overline{p}p)(T)) = (\overline{p}p)(\sigma(T))$ und daher

$$\|p(T)\|^2 = \sup\{|p(\lambda)|^2 \mid \lambda \in \sigma(T)\} = \left(\sup_{\lambda\in\sigma(T)} |p(\lambda)|\right)^2.$$

$\square$

Dank dieses Lemmas schließen wir, dass Φ_0 einen wohldefinierten involutiven Algebrenhomomorphismus vom Teilraum $U \subseteq C(\sigma(T))$ der auf $\sigma(T)$ eingeschränkten Polynomfunktionen nach $L(H)$ induziert, den wir weiterhin mit Φ_0 bezeichnen. Außerdem ist Φ_0 nach (ii) sogar isometrisch bzgl. $\|.\|_\infty$ und Operatornorm, daher insbesondere (gleichmäßig) stetig $U \to L(H)$. Wegen der $\|.\|_\infty$-Dichtheit[3] von U in $C(\sigma(T))$ gibt es eine eindeutige (gleichmäßig) stetige Fortsetzung $\Phi\colon C(\sigma(T)) \to L(H)$ von Φ_0. Durch Routineüberlegungen mit Limiten ergibt sich dann, dass auch Φ ein isometrischer involutiver Algebrenhomomorphismus ist. Wir bemerken, dass aus dem Polynom t via Polynomfunktion und Einschränkung auf $\sigma(T)$ die Funktion $\mathrm{id}_{\sigma(T)} \in C(\sigma(T))$ entsteht. Mit der Eigenschaft $\Phi(\mathrm{id}_{\sigma(T)}) = T$ und $\Phi(1) = 1$ (links die konstante Funtion 1 und rechts eigentlich id_H) ist Φ übrigens auf den Polynomfunktionen festgelegt und wegen deren Dichtheit und der Stetigkeit von Φ ist daher Φ auch auf $C(\sigma(T))$ eindeutig. Wir fassen zusammen.

Theorem: Die Abbildung Φ ist der einzige stetige involutive Algebrenhomomorphismus $\Phi\colon C(\sigma(T)) \to L(H)$ mit $\Phi(\mathrm{id}_{\sigma(T)}) = T$ und $\Phi(1) = 1$. Darüberhinaus ist Φ isometrisch.

Wir schreiben meistens etwas suggestiver $f(T)$ statt $\Phi(f)$ für $f \in C(\sigma(T))$.

Weitere Eigenschaften: (a) Für jedes $f \in C(\sigma(T))$ ist $f(T)$ normal, weil ja $\overline{f}f = f\overline{f}$ und $\overline{f}(T) = \Phi(\overline{f}) = \Phi(f)^* = f(T)^*$ gilt. Außerdem ist $f(T)$ selbstadjungiert genau dann, wenn f reellwertig ist.

(b) Sei λ ein Eigenwert von T und $f \in C(\sigma(T))$. Dann gilt $f(T)x = f(\lambda)x$ für jedes x aus dem Eigenraum zu λ, d.h. $f(\lambda)$ ist ein Eigenwert von $f(T)$. Das ist für Polynomfunktionen klar und folgt dann durch Bildung eines geeigneten Limes.

(c) Eine Verallgemeinerung von (i) aus dem obigen Lemma: Ist $f \in C(\sigma(T))$ und $\mu \in \mathbb{C} \setminus f(\sigma(T))$, dann ist die Abbildung $t \mapsto 1/(\mu - f(t))$ stetig und eine multiplikativ Inverse zur Abbildung $t \mapsto \mu - f(t)$. Das übersetzt sich mittels Φ dann in die Invertierbarkeit von $\mu -$

[3] Das kann zusammen mit (ii) entweder über stetige Ausdehnungen auf I wie oben und dem Satz von Weierstraß (auf Intervallen) argumentiert werden oder mittels Satz von Stone-Weierstraß (auf beliebigen kompakten Teilmengen von $\mathbb{R}$).

$f(T)$, d.h. $\mu \in \varrho(f(T))$. Somit gilt $\sigma(f(T)) \subseteq f(\sigma(T))$. Durch polynomiale Approximation lässt sich auch die umgekehrte Inklusion problemlos zeigen ([Wer18, Satz VII.1.4(d)]) und wir erhalten den sogenannten *Spektralabbildungssatz*

$$\sigma(f(T)) = f(\sigma(T)).$$

Bemerkung: Viele Aspekte des oben gegebenen Kalküls für „stetige Funktionen von T" können sogar ausgedehnt werden, um $f(T) \in L(H)$ allgemeiner für eine (Borel-)messbare beschränkte Funktion $f\colon \sigma(T) \to \mathbb{C}$ zu erklären (siehe z.B. [Wer18, Abschnitt VII.1]).

4.13. Ein kurzer Ausblick auf C^*-Algebren: Die Menge aller stetigen Operatoren $L(H)$ auf einem Hilbertraum H (ausgestattet mit der Operatornorm) ist nach Korollar 2.6 eine (i.A. nichtkommutative) Banachalgebra mit Einselement. Zusätzlich gibt es aber noch die Abbildung $L(H) \to L(H)$, $T \mapsto T^*$, die laut 4.1(i) konjugiert linear ist und sich nach 4.1(ii) mit der Multiplikation wie folgt verträgt: $(ST)^* = T^*S^*$. Eine Abbildung einer komplexen Algebra in sich mit diesen beiden Eigenschaften heiß *Involution*.

Im Falle von $L(H)$ ist die Involution aber sogar gemäß 4.1(v) noch besonders gut mit der Norm verträglich, weil $\|T^*T\| = \|T\|^2$ gilt. Diese Eigenschaft hat übrigens gerade in der oben skizzierten Spektraltheorie eine große Rolle gespielt. Wir haben damit bereits eine wichtige Klasse an Beispielen für den folgenden Begriff.

Definition: Eine komplexe Banachalgebra A heißt *C^*-Algebra*, wenn sie eine Involution $x \mapsto x^*$ mit der zusätzlichen Eigenschaft

$$\|x^*x\| = \|x\|^2 \quad (x \in A)$$

besitzt.

Die entsprechenden strukturerhaltenden Algebrenhomomorphismen zwischen C^*-Algebren, d.h. $\mathbb{C}$-lineare multiplikative Abbildungen $\Phi\colon A \to B$, die auch *involutiv* sind, also $\Phi(x^*) = \Phi(x)^*$ erfüllen, heißen **-Homomorphismen*. (Stetigkeit folgt dann übrigens automatisch, siehe z.B. [Wer18, Lemma IX.2.2].)

$C(K)$ ist für jeden kompakten topologischen Raum K mit der Norm $\|.\|_\infty$ und der Involution $f \mapsto \overline{f}$ eine kommutative C^*-Algebra mit Einselement. Der stetige Funktionalkalkül ist für $K = \sigma(T)$ für selbstadjungiertes $T \in L(H)$ gerade ein Beispiel für einen isometrischen *-Homomorphismus $C(\sigma(T)) \to L(H)$, der die Einselemente aufeinander abbildet.

Für kommutative C^*-Algebren mit Einselement sind die Beispiele der stetigen Funktionen auf einem Kompaktum nach einem Satz von Gelfand-Naimark (vgl. [Wer18, Theorem IX.3.4] oder [Kab14, Theorem 15.3]) sogar typisch, weil es für jede solche C^*-Algebra A einen kompakten topologischen Raum Γ_A und einen isometrischen *-Isomorphismus $A \to C(\Gamma_A)$ gibt. Es ist hier Γ_A definiert als die Menge aller multiplikativen linearen Funktionale $A \to \mathbb{C}$, ausgenommen das Nullfunktional, ausgestattet mit der Topologie der punktweisen Konvergenz. Jedem $x \in A$ wird dann die *Gelfandtransformierte* $\hat{x} \in C(\Gamma_A)$ zugeordnet, die einfach durch $\hat{x}(\varphi) := \varphi(x)$ für alle $\varphi \in \Gamma_A$ definiert ist.

Betrachtet man speziell die kommutative C^*-Algebra A, die in $L(H)$ durch einen selbstadjungierten Operator T und 1 erzeugt[4] wird, dann erhält man durch den obigen *-Isomorphismus sogar auf abstraktere Art einen weiteren Beweis für den stetigen Funktionalkalkül, weil sich in dem Fall Γ_A als homöomorph zu $\sigma(T)$ erweist ([Wer18, Satz IX.3.6 und Korollar IX.3.8]; siehe auch [Kab14, Abschnitt 15.2]). Eigentlich steckt dahinter zunächst noch die Definition des Spektrums für Elemente x einer beliebigen Banach- oder C^*-Algebra A mit Einselement e, wobei die Resolventenmenge wie erwartet so definiert wird $\varrho(x) := \{\lambda \in \mathbb{C} \mid \lambda e - x \text{ ist invertierbar in } A\}$ und dann $\sigma(x) := \mathbb{C} \setminus \varrho(x)$.

Für allgemeine C^*-Algebren A mit Einselement e gibt es die sogenannte *GNS-Konstruktion* (nach Gelfand, Naimark und Segal) zur Erzeugung von *Darstellungen* $\pi \colon A \to L(H)$ (das sind *-Homomorphismen, die e auf $1 = \mathrm{id}_H$ abbilden) als Teilalgebren von $L(H)$ auf Hilberträumen H. Insbesondere gibt es immer eine *treue*, d.h. injektive, Darstellung, weshalb A für viele Zwecke OBdA als Teilalgebra einer Operatoralgebra auf einem Hilbertraum aufgefasst werden kann (siehe dazu [Wer18, IX.3.12-15], [Conw10, §5], [Ara99, Section 2.3]).

4.14. Hilberträume und C^*-Algebren in der Quantenphysik: In der Quantenmechanik werden *Zustände* durch Einheitsvektoren in einem Hilbertraum H beschrieben und die *Observablen* des Systems durch i.A. unbeschränkte (daher unstetige) selbstadjungierte Operatoren, die auf dichten Teilräumen von H definiert sind. Wir unterdrücken hier für unsere Mini-Skizze den technisch etwas aufwändigen Zusatzaspekt, dass wir eigentlich vorher noch viele Begriffe und Resultate auf den Fall unbeschränkter Operatoren ausweiten[5] müssten, bevor wir die folgenden Formeln überhaupt hinschreiben dürfen. Für die präzise Theorie mit Bezügen zur Quantenmechanik verweisen wir auf die kurze Einführung in [Kab18, Kapitel 16] bzw. auf die ausführliche Behandlung in [Wei00].

Ist $x \in H$ ein Zustand und A eine Observable, dann ist der *Erwartungswert* von A im Zustand x gegeben durch die Zahl $\langle Ax, x\rangle$. Wir sehen, dass die Abbildung $A \mapsto \langle Ax, x\rangle$ linear ist und *reelle* Werte hat (wegen der Selbstadjungiertheit). Wenn andererseits für eine fixe Observable A nun x alle möglichen Zustände in H durchläuft, dann ergeben die entsprechenden Erwartungswerte (zumindest im Fall $A \in L(H)$) gerade den numerischen Wertebereich $W(A)$ aus 4.10, dessen Abschluss wiederum das Spektrum $\sigma(A)$ umfasst. Letzteres wird mit der Menge aller möglichen *Messergebnisse* für A identifiziert. Durch eine vertiefende Spektraltheorie mit sogenannten *Spektralmaßen*[6] zu jeder Observablen, kann jedem Zustand x ein Wahrscheinlichkeitsmaß auf $\sigma(A)$ zugeordnet werden. Im Falle, dass x ein Eigenvektor zum Eigenwert λ von A ist, liefert dies gerade das Dirac-Maß δ_λ konzentriert in $\{\lambda\}$, bedeutet also mit Sicherheit den Messwert λ für A im Zustand x.

[4] Das ist die kleinste abgeschlossene C^*-Teilalgebra von $L(H)$, die T und 1 enthält.

[5] Das kann eine eigene VO füllen.

[6] Das sind sozusagen Borelmaße auf $\mathbb{R}$ mit Werten in der Menge aller Orthogonalprojektionen in $L(H)$.

Die Dynamik wird beschrieben durch eine *stark stetige unitäre Gruppe* $(U(t))_{t\in\mathbb{R}}$ auf H, d.h. $U(t) \in L(H)$ ist für jedes t unitär, $t \mapsto U(t)$ ist ein Gruppenhomomorphismus von $\mathbb{R}$ in die Gruppe der invertierbaren (bzw. unitären) Operatoren auf H, speziell $U(-t) = U(t)^{-1} = U(t)^*$ sowie $U(0) = 1$, und

$$\forall x \in H\colon \quad U(t)x \to x \quad (t \to 0).$$

Für selbstadjungiertes $A \in L(H)$ könnten wir einfach aus der Familie von Funktionen $(e_t)_{t\in\mathbb{R}}$ mit $e_t(s) := \exp(its)$ ($s \in \mathbb{R}$) mittels stetigem Funktionalkalkül $U(t) := e_t(A) = \exp(itA)$ setzen und erhalten eine unitäre Gruppe (die in diesem Fall sogar normstetig ist). Allgemein ist nach einem Satz von Stone jede stark stetige unitäre Gruppe auf H in der Form $\exp(itA)$ mit einem eindeutigen dicht definierten (i.A. unbeschränkten) selbstadjungierten Operator A gegeben, der für alle x in seinem (unter $U(t)$ invarianten) Definitionsbereich die Gleichung

$$\frac{d}{dt}\,(U(t)x) = iAU(t)x$$

erfüllt (mit der Ableitung der entsprechenden Abbildung $\mathbb{R} \to H$ im Sinne eines Limes von Differenzenquotienten bzgl. der Norm in H). Im Falle der Dynamik eines quantenmechanischen Systems heißt A der *Hamilton-Operator* und das der obigen Gleichung für $\psi(t) := U(t)x$ entsprechende Anfangswertproblem für die *Schrödinger-Gleichung* lautet

$$\frac{d}{dt}\psi(t) = iA\psi(t), \quad \psi(0) = x.$$

Inspiriert vom Beispiel des stetigen Operators $\exp(iA)$, der von einem unbeschränkten selbstadjungierten Operator A erzeugt wird, liegt die Idee nicht fern, sozusagen durch Anwendung aller beschränkten stetigen Funktionen auf die Observablen eine C^*-Teilalgebra $\mathcal{A}$ von $L(H)$ zu erhalten. In einem weiteren Schritt kann man auch von der umgebenden Algebra $L(H)$ abstrahieren und überhaupt C^*-Algebren als Observablenalgebren studieren – bei Bedarf kann man ohnehin zu Darstellungen durch Operatoralgebren auf einem geeigneten Hilbertraum übergehen (vgl. die weiter oben erwähnte GNS-Konstruktion). Die Dynamik wird dann durch eine Gruppe von *-Isomorphismen auf der C^*-Algebra $\mathcal{A}$ gegeben und Zustände sind normierte positive lineare Funktionale $\rho \in \mathcal{A}'$, d.h. $\rho\colon \mathcal{A} \to \mathbb{C}$ ist linear mit $\|\rho\| = 1$ und $\rho(a^*a) \geq 0$ für alle $a \in \mathcal{A}$. Mehr über diesen C^*-algebraischen Zugang zur Quantenmechanik und Quantenfeldtheorie findet sich z.B. in [Ara99].

Appendix: Auswahlaxiom, Wohlordnung und Lemma von Zorn

1) Auswahlaxiom: Ist I eine nichtleere Menge und $(A_i)_{i\in I}$ eine Familie von paarweise disjunkten, nichtleeren Mengen, dann gibt es eine Abbildung $s\colon I \to \bigcup_{i\in I} A_i$ mit $s(i) \in A_i$ für alle $i \in I$. Eine Abbildung s mit dieser Eigenschaft heißt *Auswahlfunktion.*

2) Das kartesische Produkt von zwei (oder endlich vielen Mengen) wird mittels geordneter Paare (oder Tupel) definiert. Für eine beliebige nichtleere Indexmenge I ist das *kartesische Produkt* von Mengen X_i $(i \in I)$ definiert durch

$$\prod_{i\in I} X_i := \{x\colon I \to \bigcup_{i\in I} X_i \mid \forall i \in I : x(i) \in X_i\}.$$

Für endliche Indexmengen lässt sich diese Definition mit der Konstruktion über Tupel identifizieren. Insbesondere ist damit bei endlichen Produkten klar, dass nichtleere Faktoren zu einer nichtleeren Produktmenge führen. Für allgemeine Indexmengen folgt diese Aussage zunächst aus dem Auswahlaxiom und ist darüberhinaus aber sogar gleichwertig damit (vgl. z.B. auch [Cie97, Theorem 2.3.2]):

Lemma: Folgende Aussagen sind äquivalent:
(i) Auswahlaxiom
(ii) Ist $I \neq \emptyset$ und $(X_i)_{i\in I}$ eine Familie nichtleerer Mengen, dann ist $\prod_{i\in I} X_i \neq \emptyset$.

Beweis: (i)$\Rightarrow$(ii): Bilde für jedes $i \in I$ die Menge $A_i := \{i\} \times X_i$. Dann ist $(A_i)_{i\in I}$ eine Familie von paarweise disjunkten, nichtleeren[7] Mengen und das Auswahlaxiom garantiert die Existenz einer Auswahlfunktion s mit $s(i) \in \{i\} \times X_i$ für alle $i \in I$. Somit ist immer $s(i) = (i, x(i))$ für geeignetes $x(i)$ und wir erhalten durch $i \mapsto x(i)$ ein Element im Produkt, welches daher nicht die leere Menge sein kann.

(ii)$\Rightarrow$(i): Laut Voraussetzungen ist $A := \prod_{i\in I} A_i \neq \emptyset$. Somit existiert ein $s \in A$, d.h. eine Auswahlfunktion für $(A_i)_{i\in I}$. $\square$

[7] Das kartesische Produkt zweier nichtleerer Mengen ist eben vorher schon als nichtleer erkannt, weil explizit die geordneten Paare gebildet werden können (nämlich z.B. (a, b) als Menge der Form $\{\{a\}, \{a, b\}\}$).

3) Eine *Wohlordnung* auf einer Menge M ist eine Totalordnung (oder lineare Ordnung) $\leq$ auf M mit der Zusatzeigenschaft, dass jede nichtleere Teilmenge $E \subseteq M$ ein kleinstes Element besitzt, d.h. ein $k \in E$, sodass $\forall a \in E$ gilt $k \leq a$.

Wohlordnungssatz (von Zermelo): Jede Menge kann wohlgeordnet werden.

(Der Beweis stützt sich auf das Auswahlaxiom, siehe [Cie97, Theorem 4.3.3].)

4) Es sei $(M, \leq)$ eine partiell geordnete Menge. Eine Teilmenge $K \subseteq M$ heißt *Kette*, falls (die Einschränkung von) $\leq$ auf K eine Totalordnung ist.

Lemma von Zorn: Ist $(M, \leq)$ eine partiell geordnete Menge mit der Eigenschaft, dass jede Kette in M eine obere Schranke besitzt, dann gibt es in M ein maximales Element bzgl. $\leq$, d.h. es gibt ein $m \in M$, sodass gilt: $a \in M$ und $m \leq a \implies a = m$.

(Der Beweis stützt sich ebenfalls auf das Auswahlaxiom, siehe [Cie97, Theorem 4.3.4].)

(Wir verwenden diesen Satz für den Beweis des Fortsetzungssatzes von Hahn-Banach.)

5) Bemerkung: Der Wohlordnungssatz und das Lemma von Zorn sind sogar äquivalent zum Auswahlaxiom, wie z.B. in [Cie97, Corollary 4.3.5 and Remark] gezeigt wird.

Literaturverzeichnis

[Ara99] H. Araki: *Mathematical Theory of Quantum Fields*
Oxford University Press 1999.

[Cie97] K. Ciesielski: *Set Theory for the Working Mathematician.*
Cambridge University Press 1997.

[Coh13] D. L. Cohn: *Measure Theory.*
Birkäuser, 2nd edition 2013.

[Cons16] A. Constantin: *Fourier Analysis. Part I — Theory*
Cambridge University Press 2016.

[Conw10] J. B. Conway: *A Course in Functional Analysis.*
Springer, 2nd edition 1990, reprint 2010.

[El18] J. Elstrodt: *Maß- und Integrationstheorie.*
Springer, 8. Auflage 2018.

[Foll99] G. B. Folland: *Real Analysis.*
Wiley, 2nd edition 1999.

[Fried82] A. Friedman: *Foundations of Modern Analysis.*
Dover 1982.

[Heu06] H. Heuser: *Funktionalanalysis.*
Teubner, 4. Auflage 2006.

[Hoe20] G. Hörmann: *Grundbegriffe der Topologie.* Vorlesungsskriptum.
Fakultät für Mathematik, Universität Wien, WS 2020/21.

[Kab14] W. Kaballo: *Aufbaukurs Funktionalanalysis und Operatortheorie.*
Springer, 2. Auflage 2014.

[Kab18] W. Kaballo: *Grundkurs Funktionalanalysis.*
Springer, 2. Auflage 2018.

[Meg98] R. E.. Megginson: *An Introduction to Banach Space Theory.*
Springer 1998.

[Mus14] J. Muscat: *Functional Analysis.*
Springer 2014.

[Schr00] H. Schröder: *Funktionalanalysis.*
Harri Deutsch, 2. Auflage 2000.

[Wei00] J. Weidmann: *Lineare Operatoren in Hilberträumen.* Teile I und II. Teubner-Verlag 2000 und 2003.

[Wer18] D. Werner: *Funktionalanalysis.* Springer, 8. Auflage 2018.