Skript zur Vorlesung 29. Juli 2024

Funktionalanalysis 1

Sommersemester 2024

Gandalf Lechner[1]

Gegenstand dieser Vorlesung ist eine Einführung in die Funktionalanalysis. Wir beginnen mit Banach- und Hilberträumen und linearen Abbildungen (Operatoren) zwischen solchen Räumen, und besprechen dann die allgemeinen Prinzipien der Funktionalanalysis. Ein weiteres Thema ist eine Einführung in die Spektraltheorie von beschränkten Operatoren auf Hilberträumen.

[1]Department Mathematik, FAU Erlangen-Nürnberg, Cauerstr. 11 Erlangen, gandalf.lechner@fau.de

Inhaltsverzeichnis

0 Einleitung und Motivation

In der Funktionalanalysis werden die beiden Themenstränge aus der ersten Studienphase, Lineare Algebra und Analysis, zusammengeführt. In diesem einführenden Abschnitt soll dies anhand einiger Beispiele und Fragen erläutert werden.

In der Linearen Algebra stehen Vektorräume und ihre lineare Abbildungen im Vordergrund. Es werden lineare Gleichungssysteme diskutiert und die Struktur linearer Abbildungen geklärt (Eigenwerte, Diagonalisierung, Jordan-Normalform). Weiterhin wird die geometrische Bedeutung dieses Formalismus verdeutlicht, indem über lineare Untermannigfaltigkeiten (Unterräume), Winkel (Skalarprodukte), und einfache geometrische Objekte wie Quadriken gesprochen wird. Bei den betrachteten Vektorräumen ist die Dimension entweder irrelevant oder, in den meisten Fällen, endlich.

In der Analysis hingegen stehen nichtlineare Abbildungen im Vordergrund (z.B. x^2, e^x, ...), und es werden stetige, differenzierbare, messbare, und integrierbare Funktionen betrachtet. Die geometrische Natur nichtlinearer Abbildungen, also die ihrer Graphen bzw. die von lokal durch Graphen beschriebene Untermannigfaltigkeiten, ist wesentlich komplizierter als die in der Linearen Algebra betrachteten Geometrien. Lokal, also näherungsweise in der Nähe eines Punktes, finden im Zusammenhang mit differenzierbaren Funktionen aber oft Methoden der Linearen Algebra Anwendung (zB Jacobi-Matrix, Beschreibung eines Tangentialraums, etc).

In dieser Vorlesung werden wir auf andere Verbindungen zwischen Linearer Algebra und Analysis eingehen. Um dies zu motivieren, bemerken wir zunächst, dass in der Analysis ständig Vektorräume auftreten, nämlich Funktionenräume (denken Sie an $C(\mathbb{R}), C^n(\mathbb{R}), L^2(\mathbb{R})$) oder Vektorräume von Folgen (z.B. beschränkte Folgen, Cauchy-Folgen, oder konvergente Folgen). Das folgende Beispiel beleuchtet die Frage nach (Orthonormal-)basen in einem solchen Raum.

Beispiel 0.1 (Fourierreihen). Wir betrachten den Vektorraum $V := C([0, 2\pi])$ aller stetigen Funktionen $f : [0, 2\pi] \to \mathbb{C}$, den wir mit dem Skalarprodukt

$$\langle f, g\rangle := \int_0^{2\pi} \overline{f(x)}g(x)\, dx$$

ausstatten. In V betrachten wir die Vektoren (Funktionen) $f_n(x) := (2\pi)^{-1/2}e^{inx}$, $n \in \mathbb{Z}$, und machen zwei interessante Beobachtungen (siehe auch die folgende Übung):

- Die f_n sind Einheitsvektoren bzgl. der von $\langle\,\cdot\,,\,\cdot\,\rangle$ induzierten Norm:

$$\|f_n\|^2 = \langle f_n, f_n\rangle = \frac{1}{2\pi}\int_0^{2\pi} |f_n(x)|^2 dx = 1.$$

- Die f_n stehen paarweise orthogonal aufeinander: Für $n, m \in \mathbb{Z}$, $n \neq m$, gilt

$$\langle f_n, f_m\rangle = \frac{1}{2\pi}\int_0^{2\pi} e^{i(m-n)x} dx = \left[\frac{e^{i(m-n)x}}{i(m-n)}\right]_0^{2\pi} = 0$$

aufgrund der 2π-Periodizität von $x \mapsto e^{inx}$.

Diese Eigenschaften erinnern uns sehr stark an eine Orthonormalbasis, wie wir sie aus Linearer Algebra kennen. Allerdings haben wir es hier mit einem unendlichdimensionalen Vektorraum zu tun, spätestens die zweite Beobachtung oben macht das klar (wir haben unendlich viele paarweise aufeinander senkrecht stehende Vektoren). Wenn wir also überlegen, ob ein allgemeiner Vektor $v \in V$ "in diese Basis entwickelt werden kann" (einen Begriff, den wir genauer klären müssen), so sollten wir $v = \sum_{n\in\mathbb{Z}} c_n \cdot f_n$ erwarten. Aber das macht nur Sinn, wenn wir in V einen Konvergenzbegriff haben.

Da V mit einer Norm ausgestattet ist, steht ein solcher Begriff zur Verfügung (dazu später mehr).

Dieses Beispiel macht klar, dass wir für grundlegende Fragen nicht nur einen rein algebraisch definierten Vektorraum V, sondern zusätzlich einen Konvergenzbegriff auf V betrachten sollten. In Funktionalanalysis 1 wird dieser fast immer durch eine Norm gegeben sein.

Falls Sie sich nicht mehr aus Analysis erinnern:

- Zeigen Sie, dass $\langle f, g\rangle := \int_0^1 \overline{f(x)}g(x)\,dx$ tatsächlich ein Skalarprodukt auf $C([0,1])$ ist.
- Schlagen Sie die Definition eines Skalarproduktes und einer Norm nach und prüfen Sie nach, dass $\|v\| = \sqrt{\langle v, v\rangle}$ eine Norm ist, wenn $\langle\,\cdot\,,\,\cdot\,\rangle$ ein Skalarprodukt ist.

Genau wie in Linearer Algebra haben wir es in der Analysis oft mit *linearen* Abbildungen zu tun. Zwei wohlbekannte Beispiele sind:

- Differentiation ist linear, d.h. $D : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$, $D(f) := f'$ ist eine lineare Abbildung,
- Integration ist linear, d.h. $I : L^1(\mathbb{R}) \to \mathbb{R}$, $I(f) := \int_{\mathbb{R}} f(x)\,dx$, ist eine lineare Abbildung,

und es gibt noch viele weitere Beispiele. Zum Beispiel sind die Abbildung

$$M_g : C(\mathbb{R}) \to C(\mathbb{R}), \qquad (M_g f)(x) := g(x)f(x),$$

die mit einer festen Funktion $g \in C(\mathbb{R})$ multipliziert, oder die Abbildung

$$K_h : L^2(\mathbb{R}) \to \mathbb{C}, \qquad K_h(f) := \int_{\mathbb{R}} h(x)f(x)dx,$$

die gegen eine feste Funktion $h \in L^2(\mathbb{R})$ integriert, oder die Abbildung

$$\delta_x : C(\mathbb{R}) \to \mathbb{R}, \qquad \delta_x(f) = f(x),$$

die stetige Funktionen an einem festen Punkt $x \in \mathbb{R}$ auswertet, allesamt linear.

Prüfen Sie, dass all diese Abbildungen wohldefiniert und linear sind.

Die letzte Abbildung ist eine lineare Abbildung von einem Vektorraum in den zugrundeliegenden Körper. Solche Abbildungen werden wir (lineare) *Funktionale* nennen, sie haben der Funktionalanalysis ihren Namen gegeben.

Um einige wichtige Unterschiede zur endlichdimensionalen Situation zu betonen, betrachten wir ein weiteres Beispiel.

Beispiel 0.2. Wir betrachten die lineare Differentiationsabbildung

$$D : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R}), \qquad D(f) := f'.$$

Es ist klar, dass D wohldefiniert ist (die Ableitung einer glatten Funktion ist glatt). Ist D surjektiv? Ja, denn für $f \in C^\infty(\mathbb{R})$ liegt nach dem Hauptsatz der Differential- und Integralrechnung auch F, definiert durch $F(x) := \int_0^x f(t)dt$, in $C^\infty(\mathbb{R})$, und erfüllt $D(F) = f$.

Ist D injektiv? Nein, die konstanten Funktionen $f(x) = c$ bilden den (nichttrivialen) Kern von D.

Während für endlichdimensionale Vektorräume V und Endomorphismen (lineare Abbildungen $A : V \to V$) die Begriffe Injektivität, Surjektivität und Bijektivität aufgrund der Dimensionsformel äquivalent sind, ist dies für unendlichdimensionale Vektorräume nicht richtig. Es lohnt sich, sich diese Beobachtung gut zu merken, wir werden ihr noch öfters begegnen.

Bzgl. der *Eigenwerte* λ von D bemerken wir, dass jede komplexe Zahl λ ein Eigenwert ist, da $f_\lambda(x) := e^{\lambda x}$ in $C^\infty(\mathbb{R})\backslash\{0\}$ liegt und $D(f_\lambda) = \lambda \cdot f_\lambda$ erfüllt. Schränken wir D hingegen auf den Unterraum $C_c^\infty(\mathbb{R})$ aller glatten Funktionen mit kompaktem Träger ein, so sieht man leicht, dass diese Einschränkung $D|_{C_c^\infty(\mathbb{R})} : C_c^\infty(\mathbb{R}) \to C_c^\infty(\mathbb{R})$ gar keine Eigenwerte hat.

Das sind im ersten Fall zu viele und im zweiten Fall zu wenige Eigenwerte, um die Methoden der Linearen Algebra anwenden zu können. Hier werden wir also Verallgemeinerungen brauchen.

- Zeigen Sie, dass $D|_{C_c^\infty(\mathbb{R})} : C_c^\infty(\mathbb{R}) \to C_c^\infty(\mathbb{R})$ keine Eigenwerte hat.
- In obigem Beispiel haben wir einen Endomorphismus gesehen, der surjektiv aber nicht injektiv ist. Geben Sie nun ein Beispiel einer linearen Abbildung $A : V \to V$ an, die injektiv aber nicht surjektiv ist.

In dem nächsten Beispiel sehen wir, dass eine höhere/abstraktere Perspektive auf ein konkretes Problem zu interessanten Einsichten führen kann.

Beispiel 0.3. Wir suchen eine (sagen wir, stetige) Funktion $f : [0,1] \to \mathbb{R}$, die die Integralgleichung

$$f(x) = \int_0^1 \frac{f(y)}{1 + x^2 + y^2}\, dy + g(x)$$

erfüllt, wobei $g \in C([0,1])$ eine vorgegebene Funktion ist. Es erscheint schwierig, direkt zu entscheiden, ob ein solches f existiert oder es sogar zu berechnen.

Wir nehmen deshalb eine höhere Perspektive ein und betrachten den Vektorraum $V := C([0,1])$ der stetigen Funktionen auf $[0,1]$, den Vektor $g \in V$, und die lineare Abbildung $A : V \to V$, $(Af)(x) = \int_0^1 \frac{f(y)}{1+x^2+y^2}\, dy$ (Mit [Satz 2.27, Ana3] sieht man leicht ein, dass A wirklich V in V abbildet.)

So schreibt sich die obige Gleichung als

$$f = A(f) + g \iff (\mathrm{id}_V - A)(f) = g.$$

Wir erkennen:

- Die Integralgleichung ist genau dann lösbar, wenn g im Bild der linearen Abbildung $\mathrm{id}_V - A$ liegt.
- Die Integralgleichung hat genau dann für alle g eine eindeutige Lösung, wenn $(\mathrm{id}_V - A) : V \to V$ bijektiv ist. In diesem Fall ist die Lösung $f = (\mathrm{id}_V - A)^{-1}(g)$.

Für endlichdimensionales V wäre nun die Determinante von $\mathrm{id}_V - A$ ein gutes Werkzeug, um die Bijektivität zu entscheiden. Eine Determinante werden wir für lineare Abbildungen auf unendlichdimensionalen Vektorräumen allerdings nicht zur Verfügung haben (mehr dazu später).

Eine mehr analytische Idee ist, sich von der geometrischen Reihe leiten zu lassen: Ist $A \in \mathbb{C}$ eine Zahl mit $|A| < 1$, so gilt $(1 - A)^{-1} = \sum_{n=0}^{\infty} A^n$ (geometrische Reihe). Ist $A : V \to V$ eine lineare Abbildung auf einem Vektorraum wie $V = C([0, 1])$, so liegen auch die Partialsummen $\sum_{n=0}^{N} A^n$ in $\mathrm{End}(V)$ (wobei $A^0 := \mathrm{id}_V$ per Definition, $A^2 := A \circ A$, etc). Aber wir benötigen einen Konvergenzbegriff auf $\mathrm{End}(V)$, um etwas über eine unendliche Reihe sagen zu können.

Im Unterschied zu Beispiel 0.1 geht es hier um einen Konvergenzbegriff auf einem Vektorraum von linearen Abbildungen von einem Vektorraum V in sich selbst. Hier steht uns typischerweise kein Skalarprodukt zur Verfügung, aber immerhin noch eine Operatornorm (mehr dazu später).

Wir fassen die gemachten Beobachtungen zu einem Arbeitsplan für die Funktionalanalysis zusammen.

Einige Fragen und Ziele der Funktionalanalysis:

- Entwicklung einer allgemeinen Theorie von normierten Vektorräumen $(V, \|\cdot\|)$: Konvergenz- und Vollständigkeitsbegriffe. Verständnis von Basen bzw. Orthonormalbasen im Falle eines Skalarproduktes, Dualräume.
- Geometrische Betrachtungen in normierten Vektorräumen: Unterräume, konvexe Mengen, Orthogonalität.
- Lineare Abbildungen zwischen normierten Vektorräumen: Strukturanalyse, zentrale Eigenschaften, Verallgemeinerung des Eigenwertkalküls zur auf unendlichdimensionale Räume passenden Spektraltheorie.
- Diskussion von vielen konkreten Beispielen und Anwendungen zu all diesen abstrakten Ideen.

Voraussetzungen

Diese Vorlesung baut auf den Grundvorlesungen Analyis 1–3 und Lineare Algebra 1–2 auf, weitere Voraussetzungen bestehen nicht. Gelegentlich werden Bezüge zur Topologie verwendet, Vorkenntnisse in Topologie werden aber nicht erwartet. An einigen Stellen wird auf meine Analysisskripte Analysis 1–3 als [Ana1], [Ana2], [Ana3] verwiesen. Verweise auf die begleitenden Präsenzübungs- und Hausaufgaben in diesem Semester erfolgen in dem Schema P$n.m$ bzw. H$n.m$ (Aufgabe m auf Blatt n).

Zu diesem Skript

Die Gestaltung des Skripts enthält einige selbsterklärende farblich markierte Elemente (grau für Definitionen, grün für Beispiele, etc.) Zwei weitere Elemente sind pink markierte Elemente:

> Diese Abschnitte erhalten kleine Übungsaufgaben bzw. Anregungen zum Nachdenken, die nur manchmal in den Präsenz- oder Hausaufgaben aufgegriffen werden.

und gelb markierte Elemente:

> Diese Abschnitte enthalten nicht prüfungsrelevantes Zusatzmaterial oder Hinweise auf Bezüge zu weiterführenden oder anderen Vorlesungen.

Dieses Skript ist eine Hilfestellung für Sie, indem es den wesentlichen Stoff der Vorlesungen zusammenfasst. Es ist aber kein Ersatz für die Lektüre von Lehrbüchern, und sicher kein Ersatz für den Besuch der Vorlesung. Zwei empfehlenswerte Bücher sind unten aufgeführt.

Dieses Skript wird während des Semesters geschrieben und fortlaufend aktualisiert. Auf der Titelseite oben rechts finden Sie das Datum der Version, die Sie vor sich haben. Wenn Sie in diesem Skript Tippfehler oder andere Ungereimtheiten finden, kontaktieren Sie mich bitte oder posten Sie eine Frage / einen Kommentar in unserem StudOn-Forum, damit ich diese Fehler ausbessern kann.

Gandalf Lechner

Literatur

[Hal13] B. Hall. *Quantum Theory for Mathematicians*. Springer, 2013.

[RS72] M. Reed und B. Simon. *Methods of Modern Mathematical Physics I - Functional Analysis*. Academic Press, 1972.

[Rud91] W. Rudin. *Functional Analysis*. 2. Aufl. McGraw-Hill, 1991.

[SC95] F. C. Sánchez und J. M. F. Castillo. “Isometries of Finite-Dimensional Normed Spaces”. *Extr. Math.* 10.2 (1995), 146–151. URL: `http://matematicas.unex.es/~fcabello/files/printable/01.pdf`.

[Wer95] D. Werner. *Funktionalanalysis*. Springer, 1995.

1 Normierte Räume und Banachräume

In diesem Kapitel geht es um die für die Funktionalanalysis zentralen Banachräume, die uns das ganze Semester über beschäftigen werden.

1.1 Metrische und normierte Räume

Wir erinnern zunächst an einige Grundbegriffe, die Ihnen aus den Analysis-Vorlesungen bekannt sein sollten.

Definition 1.1. Sei X eine Menge. Eine *Metrik* auf X ist eine Abbildung

$$d : X \times X \to [0, \infty),$$

so dass für alle $x, y, z \in X$ die folgenden Eigenschaften gelten:

a) $d(x, y) = d(y, x)$ (Symmetrie),

b) $d(x, y) = 0 \Leftrightarrow x = y$,

c) $d(x, y) \leq d(x, z) + d(z, y)$ (Dreiecksungleichung).

Ein *metrischer Raum* ist ein Paar (X, d) bestehend aus einer Menge X und einer Metrik d auf X.

Die Idee ist, dass $d(x, y)$ den Abstand zwischen x und y angibt.

Metrische Räume müssen nicht Vektorräume sein. Im Falle eines Vektorraums haben wir den Nullvektor 0 und können die Idee des Abstands eines beliebigen Vektors zum Nullvektor formalisieren. Dies führt zum Konzept einer Norm. Wir arbeiten hier und in der gesamten Vorlesung nur mit Vektorräumen über $\mathbb{R}$ oder $\mathbb{C}$. Falls nichts anderes gesagt wird, ist mit "Vektorraum" immer Vektorraum über $\mathbb{C}$ gemeint. Wenn wir $\mathbb{K}$ schreiben, meinen wir einen beliebigen der beiden Körper $\mathbb{R}$ oder $\mathbb{C}$.

Definition 1.2. Sei V ein Vektorraum über $\mathbb{K}$. Eine *Norm* auf V ist eine Abbildung

$$\|\cdot\| : V \to [0, \infty), \qquad v \mapsto \|v\|,$$

mit den Eigenschaften ($\lambda \in \mathbb{K}$, $v \in V$)

a) $\|v\| = 0 \Leftrightarrow v = 0$ (Definitheit),

b) $\|\lambda v\| = |\lambda| \|v\|$ (positive Homogenität),

c) $\|v + w\| \leq \|v\| + \|w\|$ (Dreiecksungleichung).

Ein *normierter Raum* ist ein Paar $(V, \|\cdot\|)$ bestehend aus einem Vektorraum V und einer Norm $\|\cdot\|$ auf V.

Wenn nur die Eigenschaften b) und c) gelten, spricht man allgemeiner von einer *Halbnorm*. Das Symbol $\|\cdot\|$ wir üblicherweise für Normen reserviert.

Wir illustrieren das Konzept eines normierten Raumes mit vielen Beispielen.

Beispiel 1.3 (Endlichdimensionale normierte Räume).

a) Der Betrag $|\cdot|$ auf $\mathbb{K}$ macht $(\mathbb{K}, |\cdot|)$ zu einem normierten (eindimensionalen) Vektorraum. Der Beweis der Normeigenschaften ist hier trivial.

b) Auf Vektoren $x = (x_1, \ldots, x_d) \in \mathbb{K}^d$ betrachten wir

$$\|x\|_1 := \sum_{k=1}^{d} |x_k|.$$

Es lässt sich leicht prüfen, dass dies tatsächlich eine Norm ist. Allgemeiner betrachten wir die p-Normen (mit $p \geq 1$)

$$\|x\|_p := \left(\sum_{k=1}^{d} |x_k|^p \right)^{1/p}, \qquad p \in [1, \infty),$$

und für $p = \infty$

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

Der Beweis der Dreiecksungleichung ist für allgemeines $p \in [1, \infty]$ nicht offensichtlich; er benutzt die für alle $x = (x_1, \ldots, x_d), y = (y_1, \ldots, y_d) \in \mathbb{K}^d$ gültige *Hölder-Ungleichung* [Ana 2, Satz 8.8]

$$\sum_{k=1}^{d} |x_k y_k| \leq \|x\|_p \cdot \|y\|_q, \qquad \frac{1}{p} + \frac{1}{q} = 1,\ p, q > 1.$$

Dieses Beispiel zeigt insbesondere, dass es auf einem Vektorraum viele unterschiedliche interessante Normen geben kann.

c) Ist V ein d-dimensionaler $\mathbb{K}$-Vektorraum, so erhalten wir durch Wahl einer Basis einen Isomorphismus $\Phi : V \to \mathbb{K}^d$ und damit die Normen $\|\cdot\|_p \circ \Phi$ auf V.

Die einfachsten unendlichdimensionalen normierten Räume sind Räume von Folgen.

Beispiel 1.4 (Die Folgenräume $\ell^p(\mathbb{N})$). Wie im vorigen Beispiel betrachten wir $p \in [1, \infty]$ und für eine $\mathbb{K}$-wertige Folge $(x_n)_{n \in \mathbb{N}}$

$$\|x\|_p := \begin{cases} \left(\sum_{n=1}^{\infty} |x_n|^p\right)^{1/p} & 1 \leq p < \infty \\ \sup\limits_{n \in \mathbb{N}} |x_n| & p = \infty \end{cases}.$$

Diese Reihen konvergieren absolut oder gehen gegen ∞, dh $\|x\|_p \in [0, \infty]$ ist wohldefiniert. Wir betrachten nun die Menge[a]

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

aller "p-summierbaren" $\mathbb{K}$-wertigen Folgen $(x_n)_{n \in \mathbb{N}}$.

Dann ist $(\ell^p(\mathbb{N}), \|\cdot\|_p)$ ein normierter Raum. Der Beweis dieser Aussage wird in den Hausaufgaben geführt.

[a]Auch die Schreibweisen $\ell^p_{\mathbb{N}}$ und ℓ^p anstelle von $\ell^p(\mathbb{N})$ sind weit verbreitet.

Wie schon im einführenden Kapitel angedeutet, bilden Funktionenräume eine reichhaltige Klasse von normierten Räumen. Hier betrachten wir einige Beispiele.

Beispiel 1.5 (Einige Funktionenräume).

a) Sei S eine Menge und $\ell^\infty(S)$ die Menge aller beschränkten Funktionen $f : S \to \mathbb{C}$. Dann ist $\ell^\infty(S)$ ein Vektorraum, und

$$\|f\|_\infty := \sup_{s\in S} |f(s)|$$

eine Norm, die sogenannte *Supremumsnorm*. Dieser Raum ist unendlichdimensional genau dann, wenn S unendlich ist.
Bei den Normeigenschaften ist nur die Dreiecksungleichung vielleicht nicht ganz offensichtlich. Dazu betrachten wir $s \in S$ und $f, g \in \ell^\infty(S)$. Dann gilt

$$|f(s) + g(s)| \leq |f(s)| + |g(s)| \leq \|f\|_\infty + \|g\|_\infty,$$

also $\|f + g\|_\infty = \sup_{s\in S} |f(s) + g(s)| \leq \|f\|_\infty + \|g\|_\infty$.

Dieses Beispiel verallgemeinert den Folgenraum $\ell^\infty(\mathbb{N})$, insbesondere können wir $\ell^\infty(\mathbb{R})$ etc betrachten. Auch eine Kombination mit weiteren Eigenschaften wie zB Stetigkeit, also etwa der Raum

$$C_b(\mathbb{R}) := C(\mathbb{R}) \cap \ell^\infty(\mathbb{R})$$

aller beschränkten stetigen Funktionen, kommen häufig vor.

b) Auf dem Raum $C(\mathbb{R})$ aller stetigen Funktionen $\mathbb{R} \to \mathbb{C}$ ist

$$p(f) := \sup_{x\in[0,1]} |f(x)|$$

eine Halbnorm und keine Norm. Der Unterschied zu a) ist, dass es Elemente $f \neq 0$ gibt mit $p(f) = 0$ (betrachte f mit Träger $\operatorname{supp} f$ disjunkt zu $[0, 1]$).

c) Auf $C^1([0, 1])$ ist $p(f) := \|f'\|_\infty$ eine Halbnorm und keine Norm, da konstante Funktionen f verschwindende Ableitung haben ($\|f'\|_\infty = 0$).

d) Auf dem Raum $C^k([a, b])$ der k-mal stetig differenzierbaren Funktionen $f : [a, b] \to \mathbb{C}$ ist

$$\|f\| := \sum_{j=0}^{k} \|f^{(j)}\|_\infty$$

eine Norm, wie man leicht nachprüft.

e) Auf $C([a,b])$ ist

$$\|f\|_{L^1} := \int_a^b |f(x)|dx$$

eine Norm, die sogenannte L^1-Norm. Der Nachweis der Normeigenschaften ist wieder einfach.

Das letzte Beispiel lässt sich in verschiedene Richtungen verallgemeinern. Wir geben im nächsten Beispiel eine maßtheoretische Variante an (siehe Analysis 3). Ein Beispiel davon ist das Lebesguemaß.

Beispiel 1.6 (Die L^p-Räume). Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum. Dann ist für jede messbare Funktion $f : \Omega \to \mathbb{C}$ das Integral (mit $1 \leq p < \infty$)

$$\|f\|_{L^p} := \left(\int_\Omega |f(x)|^p d\mu(x)\right)^{1/p} \in [0, \infty]$$

definiert. Auf dem Vektorraum

$$\mathcal{L}^p(\Omega, \mu) := \{f : \Omega \to \mathbb{C} \text{ messbar} : \ \|f\|_{L^p} < \infty\}$$

ist $\|\cdot\|_{L^p}$ eine Halbnorm (Der Beweis der Dreiecksungleichung beruht dabei auf einer Integralversion der Hölderschen Ungleichung, siehe Analyis 3). Im Allgemeinen ist $\|\cdot\|_{L^p}$ allerdings keine Norm, da $\|f\|_{L^p} = 0$ nur impliziert, dass f außerhalb einer μ-Nullmenge verschwindet. (Zb ist für die Funktion $f = \chi_{\{0\}}$ das Integral $\|f\|_{L^1} = \int_{\mathbb{R}} |f(x)|d\lambda(x) = 0$ (mit Lebesguemaß λ), aber $f \neq 0$.)

Dieses Problem lässt sich dadurch beheben, dass man zu dem Quotientenraum

$$L^p(\Omega, \mu) := \mathcal{L}^p(\Omega, \mu)/\sim$$

bzgl der Äquivalenzrelation $f \sim g :\Leftrightarrow (f(x) = g(x)$ für μ-fast alle $x \in \Omega)$ übergeht.

Dann ist $(L^p(\Omega, \mu), \|\cdot\|_{L^p})$ ein normierter Raum.

Für den Fall $p = \infty$ definieren wir das *wesentliche Supremum* einer messbaren Funktion $f : \Omega \to \mathbb{C}$ als (siehe Analysis 3)

$$\|f\|_{L^\infty} := \mu\text{-esssup}|f| = \inf_{N \subset \Omega, \mu(N)=0} \sup_{x \in \Omega \setminus N} |f(x)|,$$

$$\mathcal{L}^\infty(\Omega, \mu) := \{f : \Omega \to \mathbb{C} \text{ messbar} : \ \|f\|_{L^\infty} < \infty\}.$$

Es gilt also $\|f\|_{L^\infty} < \infty$ genau dann, wenn f *wesentlich beschränkt* ist, dh wenn f auf dem Komplement einer μ-Nullmenge beschränkt ist. Auf dem Quotientenraum $L^\infty(\Omega, \mu) := \mathcal{L}^\infty(\Omega, \mu)/\sim$ bzgl. der gleichen Äquivalenzrelation wie oben ist $\|\cdot\|_{L^\infty}$ eine Norm.

Wenn die Metrik bzw. Norm aus dem Kontext klar ist, sprechen wir kürzer von einem metrischen Raum X bzw. normierten Raum V und unterdrücken die Metrik bzw. Norm in unserer Notation. Die Eigenschaften des metrischen bzw. normierten Raumes werden aber oft wesentlich von der Wahl der Metrik/Norm abhängen.

Angesichts der vielen Beispiele von normierten Vektorräumen kann man sich fragen, ob es überhaupt Vektorräume ohne Norm gibt. Dies ist tatsächlich nicht der Fall – auf *jedem* Vektorraum gibt es eine (sogar unendlich viele) Normen. Um diesen Punkt kurz zu erläutern, erinnern wir zunächst an das Konzept einer Basis aus der linearen Algebra. Um es von anderen Basiskonzepten zu unterscheiden, die wir später einführen werden, verwenden wir den Begriff "Hamelbasis" statt "Basis".

Definition 1.7. Eine *Hamelbasis* eines Vektorraums V ist eine Teilmenge $B \subset V$, so dass sich jeder Vektor $v \in V$ in eindeutiger Weise als eine (endliche!) Linearkombination von Vektoren aus B schreiben lässt.

Mit Hilfe des Lemmas von Zorn lässt sich zeigen, dass jeder Vektorraum (auch jeder unendlichdimensionale Vektorraum) eine Hamelbasis besitzt (siehe Lineare Algebra). Da der Beweis auf dem Auswahlaxiom beruht, lässt sich eine solche Basis aber oft nicht angeben, und ist häufig von nur geringem praktischen Nutzen.

Immerhin können wir folgern, dass jeder Vektorraum V eine Norm besitzt. Sei dazu B eine Hamelbasis von V, und definiere für $v = \sum_n c_n \cdot b_n$ (mit einer endlichen Summe über Koeffizienten $c_n \in \mathbb{K}$ und Basisvektoren $b_n \in B$)

$$\|v\| := \max_n |c_n|.$$

Dann prüft man leicht nach, dass $\|\cdot\|$ tatsächlich eine Norm auf V ist.

Zum Vergleich von zwei Normen erinnern wir an das aus Analysis 2 bekannte Konzept der Äquivalenz von Normen:

Definition 1.8. Zwei Normen $\|\cdot\|, \|\cdot\|'$ auf einem Vektorraum X heißen äquivalent, wenn es Konstanten $0 < c, C < \infty$ gibt, so dass

$$c\|x\| \leq \|x\|' \leq C\|x\|, \qquad x \in X.$$

Beispiel 1.9.

a) Auf einem endlichdimensionalen Vektorraum sind alle Normen äquivalent [Ana2, Satz 10.9].

b) Die Supremumsnorm $\|\cdot\|_\infty$ und die L^1-Norm $\|\cdot\|_{L^1}$ auf $C([0,1])$ sind nicht äquivalent.
Es gilt zwar $\|f\|_{L^1} \leq \|f\|_\infty$ für alle $f \in C([0,1])$, aber mit Hilfe der Folge $f_n(x) := x^n$ zeigt man, dass es keine endliche Konstante C gibt, so dass $\|f\|_\infty \leq C\|f\|_{L^1}$ für alle $f \in C([0,1])$ gilt.

c) Auf $C^k([a,b])$ können wir statt der in Beispiel 1.5 d) definierten Norm $\|f\| := \sum_{j=0}^k \|f^{(j)}\|_\infty$ auch die Norm

$$\|f\|' := \max_{j=0,\ldots,k} \|f^{(j)}\|_\infty$$

betrachten. Offenbar gilt $\|f\| \leq (k+1)\|f\|'$ und $\|f\|' \leq \|f\|$. Also sind $\|\cdot\|$ und $\|\cdot\|'$ äquivalent.

1.2 Konvergenz in normierten Räumen und beschränkte lineare Abbildungen

In einem normierten Raum stehen uns nicht nur die algebraischen Operationen (Linearkombinationen), sondern zusätzlich ein Abstands- und Konvergenzbegriff zur Verfügung. Wir nun erinnern an die grundlegenden Konzepte. Dies sind Erinnerungen an Material aus Analysis 1 und 2, dass Sie ggf. als Übungsaufgabe auffrischen können.

- Ein normierter Raum ist auf kanonische Art und Weise auch ein metrischer Raum mit der Metrik
$$d(v,w) := \|v-w\|.$$
- In einem metrischen Raum (X,d) heißt eine Folge $(x_n)_{n\in\mathbb{N}}$ *konvergent*, wenn es $x \in X$ gibt, so dass
$$\lim_{n\to\infty} d(x_n,x) = 0.$$
(d.h. für jedes $\varepsilon > 0$ gibt es $N_\varepsilon \in \mathbb{N}$, so dass $d(x_n,x) < \varepsilon$ für alle $n \geq N_\varepsilon$.)

  Die Folge $(x_n)_{n\in\mathbb{N}}$ heißt *Cauchyfolge*, wenn $d(x_n,x_m) \to 0$ für $n,m \to \infty$ (d.h. für jedes $\varepsilon > 0$ gibt es $N_\varepsilon \in \mathbb{N}$, so dass $d(x_n,x_m) < \varepsilon$ für alle $n,m \geq N_\varepsilon$.)
- Jede konvergente Folge ist eine Cauchyfolge. Die Umkehrung dieser Aussage ist im Allgemeinen falsch, dazu später mehr.

Es sei betont, dass der Konvergenzbegriff stark von der gewählten Norm abhängen kann. Zum Beispiel konvergiert die Folge der stetigen Funktionen $f_n \in C([0,1])$, $f_n(x) := x^n$, bzgl der L^1-Norm gegen Null (denn $\|f_n\|_{L^1} = \int_0^1 x^n dx = \frac{1}{n+1} \to 0$), aber nicht bzgl. der Supremumsnorm (denn $\|f_n\|_\infty = \sup_{x\in[0,1]} x^n = 1$ für alle n).

Andererseits ist es klar, dass äquivalente Normen den gleichen Konvergenzbegriff induzieren (siehe auch [Ana2, Kapitel 10.2]): Sind $\|\cdot\|$ und $\|\cdot\|'$ zwei äquivalente Normen auf einem Vektorraum V, so ist eine Folge $(v_n) \subset V$ konvergent gegen $v \in V$ (bzw Cauchy, bzw beschränkt) bzgl $\|\cdot\|$ genau dann, wenn sie konvergent gegen $v \in V$ (bzw Cauchy, bzw beschränkt) bzgl $\|\cdot\|'$ ist.

Wenn die Metrik/Norm aus dem Zusammenhang klar ist, schreiben wir $x_n \to x$ und $\lim\limits_{n\to\infty} x_n = x$ für Konvergenz bzgl dieser Metrik/Norm.

Weiterhin erinnern wir an den Begriff einer beschränkten Menge:

- Eine Teilmenge $A \subset X$ eines metrischen Raumes (X,d) heißt *beschränkt*, wenn
$$\sup_{a,b\in A} d(a,b) < \infty.$$
Im Falle eines normierten Raumes $(X,\|\cdot\|)$ ist Beschränktheit von $A \subset X$ äquivalent zu
$$\sup_{a\in A} \|a\| < \infty.$$

Beweis: leichte Übung.

- In einem metrischen Raum (X, d) ist die *offene Kugel mit Radius* $r \geq 0$ *und Mittelpunkt* $x \in X$ die Menge

  $$U_r(x) := \{y \in X : \ d(x, y) < r\}.$$

  Eine Teilmenge A eines normierten Raumes ist also genau dann beschränkt, wenn sie in einer Kugel mit hinreichend großem Radius enthalten ist.

Um etwas mit diesen Begriffen zu arbeiten, zeigen wir ein kleines Lemma.

Lemma 1.10. *Sei* $(X, \|\cdot\|)$ *ein normierter Raum.*

a) Falls $x_n \to x$ *und* $y_n \to y$*, dann auch* $x_n + y_n \to x + y$.

b) Falls $x_n \to x$ *in* X *und* $\lambda_n \to \lambda$ *in* $\mathbb{K}$*, dann auch* $\lambda_n x_n \to \lambda x$.

c) Falls $x_n \to x$*, so ist* (x_n) *beschränkt, und es gilt* $\|x_n\| \to \|x\|$.

Beweis. a) $\|(x_n + y_n) - (x + y)\| \leq \|x_n - x\| + \|y_n - y\| \to 0$.

b) $\|\lambda_n x_n - \lambda x\| \leq \|\lambda_n x_n - \lambda_n x\| + \|\lambda_n x - \lambda x\| = |\lambda_n|\|x_n - x\| + |\lambda_n - \lambda|\|x\| \to 0$.

c) Jede Norm erfüllt die umgekehrte Dreiecksungleichung $\big|\|x\| - \|y\|\big| \leq \|x - y\|$. Damit ergibt sich $\big|\|x_n\| - \|x\|\big| \leq \|x_n - x\| \to 0$. Da konvergente Folgen in $\mathbb{R}$ beschränkt sind, folgt die Behauptung. $\square$

Sie können dieses Lemma so lesen: In einem normierten Raum sind die Addition, Skalarmultiplikation und Norm stetig, dh

$$\begin{aligned} X \times X \to X, &\qquad (x, y) \mapsto x + y, \\ \mathbb{K} \times X \to X, &\qquad (\lambda, y) \mapsto \lambda \cdot x, \\ X \to \mathbb{R}, &\qquad x \mapsto \|x\|, \end{aligned}$$

sind stetige Abbildungen.

Wir fahren mit einigen Erinnerungen zu linearen Abbildungen zwischen normierten Räumen fort (siehe Analysis 2). Wir schreiben

$$\mathrm{Hom}(X, Y) := \{A : X \to Y : \ A \,\text{linear}\}$$

für den Vektorraum aller linearen Abbildungen zwischen zwei Vektorräumen X und Y. Für $A \in \mathrm{Hom}(X, Y)$ setzen wir meistens keine Klammern um das Argument, dh schreiben $Ax := A(x)$.

Definition 1.11. Seien $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ normierte Räume. Dann definieren wir für $A \in \mathrm{Hom}(X, Y)$ die *Operatornorm*

$$\|A\| := \sup_{x \in X \setminus \{0\}} \frac{\|Ax\|_Y}{\|x\|_X} = \sup_{\|x\|_X = 1} \|Ax\|_Y \in [0, \infty].$$

Eine *beschränkte* lineare Abbildung (oder: beschränkter linearer *Operator*) ist ein $A \in \mathrm{Hom}(X, Y)$ mit $\|A\| < \infty$. Die Menge aller beschränkten linearen Abbildungen $X \to$

Y wird mit $\mathcal{B}(X,Y)$ bezeichnet, und $\mathcal{B}(X) := \mathcal{B}(X,X)$ gesetzt.

Eine andere verbreitete Notation ist $\mathcal{L}(X,Y)$ anstelle von $\mathcal{B}(X,Y)$. Wir verwenden jedoch $\mathcal{B}$, um auf die definierende **B**eschränktheit der linearen Abbildungen in $\mathcal{B}(X,Y)$ hinzuweisen.

Bemerken Sie, dass Beschränktheit von A nicht Beschränktheit des Bildes bedeutet. Als lineare Abbildung hat A nur dann ein beschränktes Bild, wenn $A = 0$. Beschränktheit von A bedeutet, dass die Norm von $x \in X$ durch Anwendung von A höchstens um den Faktor $\|A\|$ gestreckt wird.

a) Zeigen Sie, dass $\mathcal{B}(X,Y)$ mit der Operatornorm ein normierter Raum ist.

b) Zeigen Sie, dass $A \in \mathrm{Hom}(X,y)$ genau dann beschränkt ist, wenn für jede beschränkte Menge $M \subset X$ das Bild $A(M) := \{Am : m \in M\} \subset Y$ beschränkt ist.

Wir erinnern auch, dass $A \in \mathrm{Hom}(X,Y)$ genau dann beschränkt ist, wenn A stetig ist. Erinnern Sie sich dazu zunächst daran, dass eine Abbildung $f : (X,d_X) \to (Y,d_Y)$ zwischen zwei metrischen Räumen genau dann stetig in einem Punkt $x \in X$ ist, wenn sie eine der folgenden äquivalenten Bedingungen erfüllt:

- Für alle $\varepsilon > 0$ existiert $\delta > 0$, so dass $d_Y(f(x), f(x')) < \varepsilon$ für alle $x' \in X$ mit $d_X(x,x') < \delta$ gilt.
  Anders formuliert: Für alle $\varepsilon > 0$ gibt es $\delta > 0$, so dass $f(U_\delta(x)) \subset U_\varepsilon(f(x))$.
- Aus $x_n \to x$ (Konvergenz einer Folge in X) folgt $f(x_n) \to f(x)$ (Konvergenz in Y).

Satz 1.12. *Seien $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ zwei normierte Vektorräume und $A \in \mathrm{Hom}(X,Y)$. Dann sind äquivalent:*

a) A ist stetig (auf ganz X).

b) A ist stetig in $x = 0$.

c) $A \in \mathcal{B}(X,Y)$.

d) Es existiert eine Konstante $C > 0$ mit $\|Ax\|_Y \leq C\|x\|_X$ für alle $x \in X$.

Beweis. c)⇒d). Sei $X \ni x \neq 0$. Dann gilt $\|Ax\|_Y = \frac{\|Ax\|_Y}{\|x\|_X} \cdot \|x\|_X \leq \|A\| \cdot \|x\|_X$, wir können die Konstante also als $C = \|A\|$ wählen.

d)⇒a). Für $x, y \in X$ haben wir $\|Ax - Ay\|_Y = \|A(x-y)\|_Y \leq C\|x-y\|_X$, was sofort die Stetigkeit von A zeigt. a)⇒b) ist trivial.

b)⇒c) Sei $\varepsilon > 0$. Aufgrund der Stetigkeit von A bei $x = 0$ gibt es $\delta > 0$ so, dass $\|x\|_X \leq \delta \Rightarrow \|Ax\|_Y < \varepsilon$. Für beliebiges $x \in X \backslash \{0\}$ haben wir dann

$$\|Ax\|_Y = \frac{2}{\delta}\|x\|_X \|A(\frac{\delta}{2\|x\|_X}x)\|_Y < \frac{2\varepsilon}{\delta} \cdot \|x\|_X,$$

also $\|A\| \leq \frac{2\varepsilon}{\delta} < \infty$. □

Wie Sie aus der Analyis 2 wissen, ist im Falle von endlichdimensionalem X jede lineare Abbildung $A \in \mathrm{Hom}(X,Y)$ (mit beliebigem Y) stetig. Für unendlichdimensionales X muss $A \in \mathrm{Hom}(X,Y)$ aber nicht stetig sein. Wir geben ein Beispiel, weitere folgen in den Übungen.

Beispiel 1.13. Wir betrachten die normierten Räume $(C([-1,1]), \|\cdot\|_{L^1})$ und $(\mathbb{C}, |\cdot|)$ sowie die lineare Abbildung

$$A : C([-1,1]) \to \mathbb{C}, \qquad Af := f(0).$$

Dann ist A nicht stetig.
Beweis: Wir betrachten die in $C([-1,1])$ liegende Funktionenfolge

$$f_n(x) := \begin{cases} n(x+\frac{1}{n}) & -\frac{1}{n} \le x \le 0 \\ -n(x-\frac{1}{n}) & 0 < x \le \frac{1}{n} \\ 0 & \text{sonst} \end{cases}.$$

Offenbar gilt $f_n(0) = 1$ für alle $n \in \mathbb{N}$, und

$$\|f_n\|_{L^1} = \int_{-1}^{1} |f_n(x)| dx = 2n \int_{-\frac{1}{n}}^{0} (x + \tfrac{1}{n}) dx = \frac{1}{n}.$$

Also gilt $\|A\| \ge \frac{|f_n(0)|}{\|f_n\|_{L^1}} = n$ und damit $\|A\| = \infty$.

Die lineare Abbildung in diesem Beispiel bildet von einem normierten Vektorraum X in den zugrundeliegenden Körper (in diesem Fall $\mathbb{C}$) ab. Solche Abbildungen nennt man auch *lineare Funktionale*. Die Existenz nicht stetiger linearer Funktionale belegt, dass wir für unendlichdimensionale Räume zwischen dem aus der linearen Algebra bekannten *algebraischen Dualraum* $\{\varphi : X \to \mathbb{K} \text{ linear}\}$ und dem kleineren *topologischen Dualraum*

$$X' := \mathcal{B}(X, \mathbb{K}) = \{\varphi : X \to \mathbb{K} \text{ linear und stetig}\}$$

unterscheiden müssen. Wir werden uns später eingehend mit diesem Raum befassen.

1.3 Beispiele und Isomorphie von Banachräumen

Die meisten nicht trivialen Konzepte der Analysis beruhen auf einem Grenzwertprozess. Dabei ist es meistens wesentlich, mit *vollständigen* Räumen zu arbeiten, also mit solchen, in denen jede Cauchyfolge konvergiert. Das elementarste Beispiel dazu liefert ein Vergleich des unvollständigen Zahlenkörpers $\mathbb{Q}$ mit dem vollständigen Zahlenkörper $\mathbb{R}$.

Dies motiviert die folgende Definition.

Definition 1.14. Ein metrischer Raum heißt *vollständig*, wenn in ihm jede Cauchyfolge konvergiert. Ein *Banachraum* ist ein normierter Vektorraum $(V, \|\cdot\|)$, der bzgl. der von seiner Norm definierten Metrik vollständig ist.

Aus Analyis 1 und 2 wissen Sie, dass $(\mathbb{K}^d, \|\cdot\|_p)$ für alle $d \in \mathbb{N}$ und alle $p \in [1,\infty]$ ein Banachraum ist. Ein erstes unendlichdimensionales Beispiel ist das folgende.

Satz 1.15. *Sei S eine Menge. Dann ist $(\ell^\infty(S), \|\cdot\|_\infty)$ ein Banachraum.*

Beweis. Um die Vollständigkeit von $\ell^\infty(S)$ zu zeigen, sei (f_n) eine Cauchyfolge. Wir müssen die Existenz eines Vektors $f \in \ell^\infty(S)$ mit $\|f_n - f\|_\infty \to 0$ für $n \to \infty$ zeigen. Dazu bemerken wir zunächst, dass die Ungleichung $|f_n(s)| \leq \|f_n\|_\infty$ impliziert, dass für jedes $s \in S$ die Folge $(f_n(s))_{n\in\mathbb{N}} \in \mathbb{K}$ eine Cauchyfolge ist (denn $|f_n(s) - f_m(s)| \leq \|f_n - f_m\|_\infty \to 0$ für $n, m \to \infty$). Da der zugrundeliegende Körper $\mathbb{K}$ vollständig ist, existieren die Grenzwerte $f(s) := \lim_n f_n(s)$.

Wir behaupten, dass die so definierte Abbildung f in $\ell^\infty(S)$ liegt und $\|f_n - f\|_\infty \to 0$ gilt. Sei $\varepsilon > 0$. Da $(f_n)_n$ eine Cauchyfolge ist, gibt es $N_\varepsilon \in \mathbb{N}$, so dass

$$\|f_n - f_m\|_\infty < \varepsilon, \qquad n, m \geq N_\varepsilon.$$

Sei $s \in S$. Da $f_n(s) \to f(s)$ konvergiert, gibt es $M_{\varepsilon,s} \in \mathbb{N}$, so dass

$$|f_n(s) - f(s)| \leq \varepsilon, \qquad n \geq M_{\varepsilon,s}.$$

Wir dürfen ohne Beschränkung der Allgemeinheit $M_{\varepsilon,s} \geq N_\varepsilon$ annehmen (ansonsten ersetzen wir $M_{\varepsilon,s}$ durch $\max\{M_{\varepsilon,s}, N_\varepsilon\}$). Somit haben wir für $n \geq N_\varepsilon$

$$\begin{aligned} |f_n(s) - f(s)| &\leq |f_n(s) - f_{M_{\varepsilon,s}}(s)| + |f_{M_{\varepsilon,s}}(s) - f(s)| \\ &\leq \|f_n - f_{M_{\varepsilon,s}}\|_\infty + \varepsilon \\ &\leq 2\varepsilon. \end{aligned}$$

Diese Abschätzung gilt nun gleichmäßig in $s \in S$, dh das N_ε ist unabhängig von s. Wir erhalten so für alle $s \in S$

$$|f(s)| \leq |f_{N_\varepsilon}(s)| + |f(s) - f_{N_\varepsilon}(s)| \leq \|f_{N_\varepsilon}\|_\infty + 2\varepsilon,$$

was die Beschränktheit $\|f\|_\infty < \infty$, also unsere erste Behauptung $f \in \ell^\infty(S)$, impliziert. Weiterhin gilt nach obiger Abschätzung $\|f_n - f\|_\infty \leq 2\varepsilon$, also $f_n \to f$ in der Supremumsnorm $\|\cdot\|_\infty$. $\square$

Eine häufig auftretende Situation ist, dass wir einen Banachraum V und einen Unterraum (Untervektorraum, also eine unter Linearkombinationen abgeschlossene Teilmenge) $U \subset V$ vorliegen haben. Dann können wir die Norm $\|\cdot\|$ von V auf U einschränken und den normierten Raum $(U, \|\cdot\|)$ betrachten. Das folgende Lemma klärt, wann U ein Banachraum ist.

Erinnern Sie sich dazu, dass eine Teilmenge $A \subset V$ eines normierten (oder metrischen) Raumes *offen* heißt, wenn es für jeden Punkt $a \in A$ eine ε-Umgebung $U_\varepsilon(a) = \{v \in V : \|v - a\| < \varepsilon\}$ (mit $\varepsilon > 0$) gibt, die ganz in A liegt, dh $U_\varepsilon(a) \subset A$. Komplementär dazu heißt A *abgeschlossen*, wenn das Komplement $V \setminus A$ offen ist. Der *Abschluss* $\overline{A}$ von A ist die kleinste A enthaltene abgeschlossene Teilmenge von V. Es gilt [Satz 8.23, Ana2]

$$\overline{A} = \{v \in V : \text{ es gibt eine Folge } (a_n) \subset A \text{ mit } a_n \to v\}.$$

Es gilt immer $A \subset \overline{A}$. Die Gleichheit $A = \overline{A}$ gilt genau dann, wenn A abgeschlossen ist.

Lemma 1.16. *Eine Teilmenge $A \subset V$ eines Banachraums V ist genau dann vollständig, wenn sie abgeschlossen ist.*

Beweis. Sei $A = \overline{A}$ abgeschlossen und $(a_n) \subset A$ eine Cauchyfolge. Aufgrund der Vollständigkeit von V gibt es einen Grenzwert $v \in V$ mit $a_n \to v$. Dann liegt v in $\overline{A} = A$. Also ist A vollständig.

Sei A vollständig und $x \in \overline{A}$. Dann gibt es eine Folge $(a_n) \subset A$ mit $a_n \to x$. Aufgrund der Vollständigkeit von A folgt $x \in A$. Also $A = \overline{A}$, dh A ist abgeschlossen. □

Beispiel 1.17.

a) Der Banachraum $\ell^\infty(\mathbb{N})$ aller beschränkter Folgen enthält den Raum $c_0(\mathbb{N})$ aller Nullfolgen als Unterraum. Wir behaupten, dass $c_0(\mathbb{N})$ ein Banachraum ist.
Da wir $\ell^\infty(\mathbb{N})$ bereits als Banachraum erkannt haben (Satz 1.15), müssen wir nach Lemma 1.16 nur zeigen, dass $c_0(\mathbb{N})$ abgeschlossen ist. Sei also $(x_n)_n \subset c_0(\mathbb{N})$ eine Cauchyfolge von Nullfolgen, die wir der Übersichtlichkeit halber als Funktionen $x_n(k) := x_{n,k}$ notieren. Sei $x \in \ell^\infty(\mathbb{N})$ der aufgrund der Vollständigkeit von $\ell^\infty(\mathbb{N})$ existierende Grenzwert. Wir müssen $x \in c_0(\mathbb{N})$ zeigen.

   Sei $\varepsilon > 0$. Dann gibt es $N \in \mathbb{N}$ mit $\|x_n - x\|_\infty < \varepsilon$ für $n \geq N$. Wegen $\lim_{k\to\infty} x_N(k) = 0$ (Nullfolge) existiert $M \in \mathbb{N}$ mit $|x_N(k)| \leq \varepsilon$ für $k \geq M$. Damit haben wir

$$|x(m)| \leq |x(m) - x_N(m)| + |x_N(m)| \leq 2\varepsilon, \qquad m \geq M,$$

   was $x \in c_0(\mathbb{N})$ zeigt.

b) In dem normierten Raum $(C^1([-1,1]), \|\cdot\|_\infty)$ der stetig differenzierbaren Funktionen mit Supremumsnorm betrachten wir die Folge $f_n(x) := (x^2 + \frac{1}{n})^{1/2}$. Dann ist $(f_n)_{n\in\mathbb{N}}$ eine Cauchyfolge, aber sie konvergiert nicht, das heißt der Raum $(C^1([-1,1]), \|\cdot\|_\infty)$ ist kein Banachraum. Man beweist diese Aussagen am Besten dadurch, dass man sieht, dass

$$\lim_{n\to\infty} f_n(x) = |x|$$

   bzgl $\|\cdot\|_\infty$ konvergiert (also gleichmäßig konvergiert), denn es gilt

$$\begin{aligned} \|f_n - |\cdot|\|_\infty &= \sup_{|x|\leq 1} \left(\left(x^2 + \tfrac{1}{n}\right)^{1/2} - |x| \right) \\ &\leq \sup_{|x|\leq 1} \left(\left(|x|^2 + \tfrac{2|x|}{\sqrt{n}} + \tfrac{1}{n} \right)^{1/2} - |x| \right) = \tfrac{1}{\sqrt{n}} \to 0. \end{aligned}$$

   Die Folge (f_n) konvgergiert also in dem größeren Raum $C([-1,1]) \supset C^1([-1,1])$. Wie Sie aus Analysis 1 wissen, ist $(C([-1,1]), \|\cdot\|_\infty)$ ein Banachraum.

   Da der Grenzwert $f(x) := |x|$ nicht in $C^1([-1,1])$ liegt (nicht differenzierbar bei $x = 0$), ist $(C^1([-1,1]), \|\cdot\|_\infty)$ nicht abgeschlossen, also kein Banachraum.

Die Folge $(f_n) \subset C^1([-1,1])$ ist eine Cauchyfolge (da sie in dem größeren Raum $C([-1,1])$ konvergiert), aber sie konvergiert nicht in $C^1([-1,1])$, da der Grenzwert nicht in diesem Unterraum liegt.

Sei $c_{\mathbb{N}}$ der Raum aller konvergenten komplexen Folgen, und $d_{\mathbb{N}}$ der Raum aller abbrechenden komplexen Folgen, d.h.

$$\begin{aligned} c_{\mathbb{N}} &:= \{x : \mathbb{N} \to \mathbb{C} : \ x \ \text{konvergiert}\}, \\ d_{\mathbb{N}} &:= \{x : \mathbb{N} \to \mathbb{C} : \ x_n \neq 0 \ \text{für höchstens endlich viele } n \in \mathbb{N}\}. \end{aligned}$$

Zeigen Sie: $c_{\mathbb{N}} \subset \ell^\infty_{\mathbb{N}}$ ist abgeschlossen, also ein Banachraum, und $d_{\mathbb{N}} \subset \ell^\infty_{\mathbb{N}}$ ist nicht abgeschlossen, also kein Banachraum.

Wir geben noch einige weitere wichtige Beispiele von Banachräumen.

Satz 1.18. $(\ell^p(\mathbb{N}), \|\cdot\|_p)$ *ist für jedes* $p \in [1,\infty]$ *ein Banachraum.*

Beweis. Für $p = \infty$ haben wir dies bereits in Satz 1.15 bewiesen. Sei also $1 \leq p < \infty$.

Sei (x_n) eine Cauchy-Folge in $\ell^p(\mathbb{N})$. Mittels der umgekehrten Dreiecksungleichung

$$\big|\|x_n\|_p - \|x_m\|_p\big| \leq \|x_n - x_m\|_p \to 0$$

sehen wir, dass die Folge der Normen $(\|x_n\|_p)_n$ konvergiert. Sei $C := \lim_n \|x_n\|_p$.

Wir notieren die x_n als Funktionen $\mathbb{N} \to \mathbb{K}$, dh schreiben $x_n(k)$ für $(x_n)_k$, und bemerken $|x_n(k) - x_m(k)| \leq \left(\sum_{j=1}^\infty |x_n(j) - x_m(j)|^p\right)^{1/p} = \|x_n - x_m\|_p \to 0$ für jedes $k \in \mathbb{N}$. Aufgrund der Vollständigkeit von $\mathbb{K}$ existieren also die Grenzwerte $x(k) := \lim_n x_n(k)$. Wir behaupten $x \in \ell^p(\mathbb{N})$ und $x_n \to x$.

Dazu zeigen wir

$$\sum_{k=1}^K |x(k)|^p = \lim_{n\to\infty} \sum_{k=1}^K |x_n(k)|^p \leq \lim_{n\to\infty} \sum_{k=1}^\infty |x_n(k)|^p = \lim_{n\to\infty} \|x_n\|_p^p = C^p.$$

Da die linke Seite mit K monoton wächst und die Schranke C^p von K unabhängig ist, erhalten wir $\sum_{k=1}^\infty |x(k)|^p < \infty$, also $x \in \ell^p(\mathbb{N})$.

Um $x_n \to x$ zu zeigen, argumentieren wir zuerst analog, nämlich

$$\begin{aligned} \sum_{k=1}^K |x_n(k) - x(k)|^p &= \lim_{m\to\infty} \sum_{k=1}^K |x_n(k) - x_m(k)|^p \leq \lim_{m\to\infty} \sum_{k=1}^\infty |x_n(k) - x_m(k)|^p \\ &= \lim_{m\to\infty} \|x_n - x_m\|_p^p, \end{aligned}$$

und damit $\|x_n - x\|_p \leq \lim_{m\to\infty} \|x_n - x_m\|_p$ im Grenzwert $K \to \infty$. (Mit der umgekehrten Dreiecksungleichung sieht man wie oben, dass $\lim_{m\to\infty} \|x_n - x_m\|_p$ existiert.)

Da (x_n) eine Cauchyfolge ist, finden wir zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ mit $\|x_n - x_m\|_p < \varepsilon$ für $n, m \geq N$. Also gilt $\|x_n - x\|_p < \varepsilon$ für $n \geq N$, dh $x_n \to x$. □

Auch die auf Grundlage eines Maßraumes (Ω, μ) definierten Räume $L^p(\Omega, \mu)$ aus Beispiel 1.6 sind für alle $p \in [1, \infty]$ Banachräume (Analysis 3).

Satz 1.19. *Sei X ein normierter Raum, Y ein Banachraum, und $\mathcal{B}(X, Y)$ der Raum aller beschränkten linearen Operatoren $X \to Y$ (mit der Operatornorm, siehe Def. 1.11). Dann ist $\mathcal{B}(X, Y)$ ein Banachraum.*

Beachten Sie, dass eventuelle Vollständigkeit von X hier irrelevant, die Vollständigkeit von Y aber wesentlich ist.

Beweis. Wir wissen bereits, dass $\mathcal{B}(X, Y)$ ein normierter Raum ist und müssen seine Vollständigkeit zeigen. Sei also $(A_n)_n \in \mathcal{B}(X, Y)$ eine Cauchyfolge, und $x \in X$. Dann gilt $\|A_n x - A_m x\| \leq \|A_n - A_m\| \|x\| \to 0$, so dass wir $(A_n x)$ als Cauchyfolge in Y erkennen. Aufgrund der Vollständigkeit von Y ist diese Folge konvergent. Wir definieren

$$A : X \to Y, \qquad Ax := \lim_{n\to\infty} A_n x.$$

Man prüft leicht nach, dass A linear ist (Lemma 1.10). Wir behaupten $A \in \mathcal{B}(X, Y)$ und $A_n \to A$. Dazu betrachten wir $x \in X$ und $\|Ax\| = \lim_n \|A_n x\| \leq \lim_n \|A_n\| \cdot \|x\|$. Da dieser Grenzwert existiert (wieder umgekehrte Dreiecksungleichung), erhalten wir $\|A\| \leq \lim_n \|A_n\| < \infty$, also $A \in \mathcal{B}(X, Y)$. Außerdem gilt

$$\|A_n - A\| = \sup_{\|x\|=1} \|A_n x - Ax\| = \sup_{\|x\|=1} \lim_{m\to\infty} \|A_n x - A_m x\| \leq \lim_{m\to\infty} \|A_n - A_m\|,$$

was wie am Ende des Beweises von Satz 1.18 $\|A_n - A\| \to 0$, also $A_n \to A$, impliziert. □

Beschränkte Operatoren sind ein zentrales Thema dieser Vorlesung, dem wir uns später im Detail zuwenden werden. Hier notieren wir nur das folgende Korollar über den Dualraum $X' = \mathcal{B}(X, \mathbb{K})$ eines normierten Raumes X.

Korollar 1.20. *Der Dualraum X' eines normierten Raumes X ist ein Banachraum.*

Beweis. Klar wegen der Vollständigkeit von $\mathbb{K}$. □

Nach diesen Beispielen betrachten wir Banachräume aus einer abstrakteren Sichtweise. Zunächst klären wir, was wir unter isomorphen Banachräumen verstehen möchten. In der linearen Algebra nennt man zwei Vektorräume X, Y isomorph, wenn es eine bijektive lineare Abbildung zwischen ihnen gibt. Isomorphe Vektorräume haben also dieselben algebraischen Eigenschaften. Für Banachräume möchte man zusätzlich, dass sich die Normen der beiden Räume kompatibel verhalten. Dazu führen wir die folgenden Begriffe ein.

Definition 1.21. Seien X, Y normierte Räume.

a) Ein *Isomorphismus* zwischen X und Y ist eine bijektive Abbildung $A \in \mathcal{B}(X, Y)$ mit $A^{-1} \in \mathcal{B}(Y, X)$.

b) Eine *Isometrie* ist ein Operator $A \in \mathcal{B}(X, Y)$ mit $\|Ax\| = \|x\|$ für alle $x \in X$.

c) X und Y sind *(isometrisch) isomorph*, wenn es einen (isometrischen) Isomorphismus zwischen ihnen gibt.

Diese Begriffe bedürfen einiger Erläuterungen. Zum ersten Punkt bemerken wir, dass eine bijektive stetige lineare Abbildung $A \in \mathcal{B}(X, Y)$ stets eine lineare Umkehrabbildung $A^{-1} : Y \to X$ besitzt. Wir werden später sehen, dass, falls X und Y Banachräume sind, A^{-1} ebenfalls beschränkt (stetig) ist. Ist Y aber kein Banachraum, so ist diese Aussage falsch, wie wir jetzt erläutern.

Satz 1.22. *Seien X, Y normierte Räume und $A : X \to Y$ eine bijektive lineare Abbildung. Dann ist A^{-1} genau dann stetig, wenn es eine Konstante $c > 0$ gibt, so dass*

$$\|Ax\| \geq c\|x\|, \qquad x \in X,$$

gilt.

Beweis. Falls A^{-1} stetig ist, gilt $\|A^{-1}y\| \leq \|A^{-1}\|\|y\|$ (mit $\|A^{-1}\| < \infty$) für alle $y \in Y$. Für $x \in X$ und $y := Ax$ ergibt das $\|x\| \leq \|A^{-1}\|\|Ax\|$, also die Behauptung mit $c := \|A^{-1}\|^{-1} < \infty$.

Gilt hingegen $\|Ax\| \geq c\|x\|$ für ein $c > 0$ und alle $x \in X$, so gilt für $x := A^{-1}y$ die Abschätzung

$$\|A^{-1}y\| = \|x\| \leq c^{-1}\|Ax\| = c^{-1}\|y\|,$$

also $\|A^{-1}\| \leq c^{-1} < \infty$. □

Beispiel 1.23. Sei $X = \ell^\infty(\mathbb{N})$. Wir betrachten die Abbildung

$$A : \ell^\infty_{\mathbb{N}} \to \ell^\infty_{\mathbb{N}}, \qquad (Ax)_n := \frac{1}{n} \cdot x_n.$$

Dann gilt für $x \in \ell^\infty_{\mathbb{N}}$

$$\|Ax\|_\infty = \sup_n \frac{|x_n|}{n} \leq \sup_n |x_n| = \|x\|_\infty < \infty,$$

was $Ax \in \ell^\infty_{\mathbb{N}}$ und $\|A\| \leq 1 < \infty$ zeigt. Also $A \in \mathcal{B}(\ell^\infty_{\mathbb{N}})$. Weiterhin ist A injektiv, denn $\ker A = \{x \in X : n^{-1}x_n = 0 \, \forall n \in \mathbb{N}\} = \{0\}$. Betrachten wir nun das Bild von A als normierten Raum $Y := A(\ell^\infty_{\mathbb{N}})$ (mit Norm $\|\cdot\|_\infty$) und die koeingeschränkte Abbildung $B : X \to Y, x \mapsto Ax$. Dann ist B bijektiv und beschränkt. Aber die Umkehrabbildung B^{-1} ist nach Satz 1.22 unbeschränkt: Auf den Einheitsvektoren $(e_k)_n := \delta_{k,n}$ gilt $Ae_k = k^{-1}e_{(k)}$, also $\|Ae_{(k)}\|_\infty = k^{-1}\|e_{(k)}\|_\infty$ für alle $k \in \mathbb{N}$. Demnach gibt es kein $c > 0$ mit $\|Ax\|_\infty \geq c\|x\|_\infty$ für alle $x \in X$.

Zum zweiten und dritten Punkt von Definition 1.21 bemerken wir, dass eine Isometrie immer injektiv ist (klar), aber nicht bijektiv zu sein braucht.

Beispiel 1.24. Wir betrachten $X = \ell^p_{\mathbb{N}}$ mit $p \in [1,\infty]$ und den **Shift-Operator**

$$S : X \to X, \qquad (Sx)_n := \begin{cases} x_{n-1} & n \geq 2 \\ 0 & n = 1 \end{cases}.$$

Dann gilt $\|Sx\|_p^p = \sum_{n=2}^{\infty} |x_{n-1}|^p = \sum_{n=1}^{\infty} |x_n|^p = \|x\|_p^p$. Also ist S eine Isometrie. Aber wegen $(Sx)_1 = 0$ für alle x liegt der Vektor $(1, 0, 0, \ldots)$ nicht im Bild von S, dh S ist nicht surjektiv.

Ist A eine bijektive Isometrie, so ist A^{-1} auch eine Isometrie (einfache Übung).

Zeigen Sie, dass der Raum $c_{\mathbb{N}}$ aller konvergenten Folgen und der Raum c_0 aller Nullfolgen (jeweils mit $\|\cdot\|_\infty$) isomorphe Banachräume sind.
Tipp: Betrachte $A : c \to c_0$, $(Ax)_1 := \lim_k x_k$, $(Ax)_n := x_{n-1} - \lim_k x_k$, $n \geq 2$.

Um die Begriffe von Isomorphismus und isometrischem Isomorphismus abzugrenzen, betrachten wir zwei Banachräume $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$. Sind X, Y isomorph mit Isomorphismus $A : X \to Y$, so können wir auf X die neue Norm

$$\|x\|_X^A := \|Ax\|_Y$$

einführen (dies ist tatsächlich eine Norm, wie man leicht nachprüft). Mittels Satz 1.22 zeigt sich dann, dass $\|\cdot\|_X$ und $\|\cdot\|_X^A$ äquivalente Normen sind. Ist A sogar ein *isometrischer* Isomorphimus, so sind die Normen $\|\cdot\|_X$ und $\|\cdot\|_X^A$ gleich.

Der Unterschied zwischen Isomorphismus und isometrischem Isomorphismus ist auch schon für endlichdimensionale Banachräume sichtbar: Für $d \in \mathbb{N}$ und $p, q \in [1,\infty]$ sind $(\mathbb{R}^n, \|\cdot\|_p)$ und $(\mathbb{R}^n, \|\cdot\|_q)$ isomorph, aber isometrisch isomorph nur falls $p = q$ oder falls $n = 2$ und $p, q \in \{1, \infty\}$. Dies können Sie in [SC95] nachlesen.

Das folgende Lemma illustriert, wie Isomorphismen Eigenschaften eines normierten Raumes auf das isomorphe Bild übertragen. Es gilt insbesondere für surjektive Isometrien (=isometrische Isomorphismen).

Lemma 1.25. *Seien X, Y normierte Räume und $A : X \to Y$ ein Isomorphismus. Dann ist X genau dann vollständig, wenn Y vollständig ist.*

Beweis. Angenommen, X is vollständig und $y_n = Ax_n$ eine Cauchyfolge in Y. Dann ist wegen $\|x_n - x_m\| = \|A^{-1}y_n - A^{-1}y_m\| \leq \|A^{-1}\|\|y_n - y_m\| \to 0$ die Folge $(x_n)_n$ eine Cauchyfolge in X. Aufgrund der Vollständigkeit von X gibt es einen Grenzwert $x_n \to x \in X$. Dann gilt mit $y := Ax$ auch $\|y_n - y\| = \|Ax_n - Ax\| \leq \|A\|\|x_n - x\| \to 0$, dh $(y_n)_n$ konvergiert. Also ist Y vollständig.

Für die andere Richtung verwendet man A^{-1} statt A. □

1.4 Direkte Summen und Quotienten von Banachräumen

In diesem Abschnitt wollen wir zwei Konstruktionen diskutieren, die aus gegebenen Banachräumen neue Banachräume machen. Die einfachste Konstruktion ist die direkte Summe.

Definition 1.26. Seien $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ normierte Räume. Die direkte Summe von X und Y ist die Menge

$$X \oplus Y = \{(x, y) : \ x \in X, \ y \in Y\} = X \times Y,$$

mit der Vektorraumstruktur $(x, y) + \lambda(x', y') = (x + \lambda x', y + \lambda y')$ (mit $x, x' \in X$, $y, y' \in Y$, $\lambda \in \mathbb{K}$).

Für $p \in [1, \infty]$ definieren wir weiterhin

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

Satz 1.27.

a) Die direkte Summe $X \oplus Y$ zweier normierter Räume X, Y mit einer der Normen $\|\cdot\|_p$, $p \in [1, \infty]$, ist ein normierter Raum.

b) Die Normen $\|\cdot\|_p$ sind alle äquivalent und es gilt $(x_n, y_n) \to (x, y)$ bzgl einer dieser Normen genau dann wenn $x_n \to x$ und $y_n \to y$.

c) Sind X und Y vollständig, so ist auch $(X \oplus Y, \|\cdot\|_p)$ vollständig.

Der Beweis wird in den Übungen geführt, wo sich auch Beispiele finden.

Eine zweite wichtige Konstruktion ist die des Quotienten bzgl eines Unterraums. Dazu erinnern wir zunächst an das rein algebraische Konzept: Ist $U \subset V$ ein Untervektorraum eines Vektorraums V, so können wir auf V die Äquivalenzrelation

$$v \sim v' :\Leftrightarrow v - v' \in U$$

einführen. Der Quotient, dh die Menge $V/U := V/\sim$ der Äquivalenzklassen $[v] = v+U$, ist ein Vektorraum bzgl der (wohldefinierten) Operationen $[v]+[v'] := [v+v']$ und $\lambda[v] := [\lambda v]$.

Im Zusammenhang mit *normierten* Vektorräumen brauchen wir auch eine Norm auf V/U. Dazu verabreden wir zuerst, in einem normierten Raum V den Abstand eines Punktes $v \in V$ zu einer Teilmenge $A \subset V$ als

$$d(v, A) := \inf\{\|v - a\| : \ a \in A\}$$

zu definieren. Mit dieser Definition gilt $\overline{A} = \{v \in V : \ d(v, A) = 0\}$ (Übung).

Satz 1.28. *Sei V ein normierter Raum und $U \subset V$ ein Unterraum.*

a) $\|[v]\| := d(v, U)$ ist eine Halbnorm auf V/U. Ist U abgeschlossen, so ist dies eine Norm.

b) Ist V vollständig und U abgeschlossen, so ist V/U mit der in a) eingeführten Norm ein Banachraum.

> *c) Die Quotientenabbildung (für U abgeschlossen)*
>
> $$\pi : V \to V/U, \qquad \pi(v) := [v],$$
>
> *ist linear und stetig.*

Beweis. a) Wir prüfen zuerst, dass $\|\cdot\|$ wohldefiniert ist, also nicht von der Wahl eines Repräsentanten v der Klasse $[v]$ abhängt. Ist $v' \sim v$, also $v' - v \in U$, so gilt

$$d(v', U) = \inf\{\|v' - u\| : u \in U\} = \inf\{\|v + (v' - v) - u\| : u \in U\} = d(v, U).$$

Wir prüfen nun die (Halb-)normeigenschaften nach. Es gilt sicher $\|[v]\| \geq 0$ für alle $v \in V$ und $\|\lambda[v]\| = \inf\{\|\lambda v - u\| : u \in U\} = |\lambda| \inf\{\|v - \lambda^{-1}u\| : u \in U\} = |\lambda|\|[v]\|$ für $\lambda \neq 0$. Für $\lambda = 0$ ist $\|0[v]\| = 0$ trivial.

Für die Dreiecksungleichung betrachten wir $v_1, v_2 \in V$. Aufgrund der Definition von $\|[v_i]\|$ als Infimum gibt es Vektoren $u_1, u_2 \in U$ mit $\|v_i - u_i\| \leq \|[v_i]\| + \varepsilon$, $i = 1, 2$. Damit erhalten wir

$$\begin{aligned}
\|[v_1] + [v_2]\| &= \|[v_1 + v_2]\| \\
&\leq \|v_1 + v_2 - (u_1 + u_2)\| \\
&\leq \|[v_1]\| + \|[v_2]\| + 2\varepsilon,
\end{aligned}$$

was die Dreiecksungleichung zeigt.

Klassen mit $\|[v]\| = 0$ erfüllen $d(v, U) = 0$, also $v \in \overline{U}$. Ist $U = \overline{U}$ abgeschlossen, folgt $v \in U$, also $[v] = 0$. Damit ist $\|\cdot\|$ in diesem Fall eine Norm.

b) Wir verwenden folgendes

> **Lemma 1.29.** *Ein normierter Raum V ist vollständig genau dann, wenn jede absolut summierbare Folge $(v_n) \subset V$ (dh jede Folge mit $\sum_{n=1}^{\infty} \|v_n\| < \infty$) summierbar ist (dh $\sum_{n=1}^{N} v_n$ konvergiert in V für $N \to \infty$).*

Der Beweis dieses Lemmas wird in den Übungen geführt.

Wir zeigen nun b) mit Hilfe dieses Lemmas. Sei also eine Folge $(v_n) \subset V$ gegeben, so dass $\sum_n \|[v_n]\| < \infty$. Wir müssen zeigen, dass $(\sum_{n=1}^{N} [v_n])_N$ in V/U konvergiert. Nach Definition der Quotientennorm als Infimum finden wir Vektoren $u_n \in U$, so dass $\|v_n - u_n\| \leq \|[v_n]\| + 2^{-n}$ annehmen. Dann gilt auch $\sum_{n=1}^{\infty} \|v_n - u_n\| < \infty$, und die Vollständigkeit von V impliziert die Existenz von $v := \sum_{n=1}^{\infty} (v_n - u_n) \in V$. Wir erhalten

$$\left\| [v] - \sum_{n=1}^{N} [v_n - u_n] \right\| = \left\| \left[v - \sum_{n=1}^{N} (v_n - u_n) \right] \right\| \leq \left\| v - \sum_{n=1}^{N} (v_n - u_n) \right\| \to 0$$

für $N \to \infty$. Dies beweist die Behauptung.

c) π ist offensichtlich linear und hat wegen $\|[v]\| = \inf\{\|v - u\| : u \in U\} \leq \|v\|$ Norm $\|\pi\| \leq 1$, ist also stetig. □

Beispiel 1.30. Als Beispiel betrachten wir $V = C([0,1])$ mit Supremumsnorm, $U = \{f \in V : f(0) = 0\}$. Dann ist V/U isometrisch isomorph zu $(\mathbb{C}, |\cdot|)$.
Der Beweis erfolgt in den Übungen.

Ein weiteres Beispiel von Quotientenräumen ist die Konstruktion, die wir schon im Zusammenhang mit den L^p- bzw. $\mathcal{L}^p$-Räumen gesehen haben: Ist X ein Raum mit einer Halbnorm p, dann ist ist $\ker p = \{x \in X : p(x) = 0\}$ ein abgeschlossener Unterraum und auf $X/\ker p$ ist die Quotientennorm $\|[x]\| := d(x, \ker p)$ eine Norm.

1.5 Vervollständigung

Wie wir bereits in vielen Situationen gesehen haben, ist die Vollständigkeit von normierten Räumen eine wichtige Eigenschaft. Wir diskutieren deshalb hier eine Methode, einen nicht vollständigen Raum so zu vergrößern, dass er vollständig wird (Vervollständigung).

Ist V ein Banachraum und $U \subset V$ ein nicht notwendigerweise vollständiger Unterraum, so wissen wir bereits aus Lemma 1.16, dass die Vollständigkeit von U (eine interne Eigenschaft von U, die nur durch Folgen innerhalb von U definiert ist) zu der Abgeschlossenheit von U (eine externe Eigenschaft von U, die über die Inklusion $U \subset V$ definiert ist) äquivalent ist. In dieser Situation ist die Vervollständigung von U naheliegend: Wir betrachten den Abschluss $\overline{U}$.

Lemma 1.31. *Ist $U \subset V$ ein Unterraum eines Banachraums V, so ist der Abschluss $\overline{U} \subset V$ ein Banachraum.*

Beweis. Per Konstruktion ist $\overline{U}$ vollständig und in V enthalten. Wir müssen nur zeigen, dass es auch ein Vektorraum ist. Für $x, y \in \overline{U}$ gibt es Folgen $(x_n), (y_n) \subset U$ mit $x_n \to x$ und $y_n \to y$. Nach Lemma 1.10 gilt $x_n + y_n \to x + y$, also $x + y \in \overline{U}$. Analog sieht man für $\lambda \in \mathbb{K}$ die Konvergenz $\lambda x_n \to \lambda x$, also $\lambda x \in \overline{U}$, was die Unterraumeigenschaften von $\overline{U}$ beweist. □

Per Konstruktion ist $\overline{U}$ der kleinste Banachraum, der U enthält. Weiterhin liegt U *dicht* in $\overline{U}$ (Erinnerung Analysis: Eine Teilmenge A eines metrischen Raums B liegt dicht in B, falls $\overline{A} = B$) und die Inklusionsabbildung $\Phi : U \to \overline{U}$, $\Phi(u) := u$, ist eine Isometrie.

Wir möchten diese Konstruktion jetzt auf die allgemeine Situation verallgemeinern, dass wir einen nicht notwendigerweise vollständigen normierten Raum $(X, \|\cdot\|)$ vorliegen haben, der nicht als Unterraum eines Banachraumes gegeben ist. In diesem Fall muss die Vervollständigung von X konstruiert werden. Die folgenden Ergebnisse gelten (in geeigneter Form) für jeden metrischen Raum. Wir konzentrieren uns hier aber auf den Fall normierter Räume.

Definition 1.32. Sei X ein normierter Raum. Eine *Vervollständigung* von X ist ein Paar (Y, Φ) bestehend aus einem Banachraum Y und einer Isometrie $\Phi : X \to Y$ mit dichtem Bild (dh $\overline{\Phi(X)} = Y$).

Satz 1.33. *Sei X ein normierter Raum.*

a) (Existenz) X besitzt eine Vervollständigung (Y, Φ).

b) (Fortsetzung stetiger Abbildungen auf die Vervollständigung) Ist Z ein Banachraum und $A : X \to Z$ linear und stetig, so existiert eine eindeutige stetige lineare Abbildung $\widehat{A} : Y \to Z$ mit $\widehat{A} \circ \Phi = A$. Es gilt $\|\widehat{A}\| = \|A\|$.

c) (Eindeutigkeit) Je zwei Vervollständigungen (Y, Φ), $(\tilde{Y}, \tilde{\Phi})$ von X sind isometrisch isomorph.

Beweis. a) Eine direkte Methode zur Konstruktion einer Vervollständigung verläuft analog zur Konstruktion von $\mathbb{R}$ aus $\mathbb{Q}$. Wir skizzieren diese Methode hier nur kurz, weil wir später einen eleganteren Beweis führen werden.

Man betrachtet den Raum $CF(X)$ aller Cauchyfolgen in X und darauf die Äquivalenzrelation $(x_n) \sim (y_n) :\Leftrightarrow \|x_n - y_n\| \to 0$. Wir definieren $Y := CF(X)/\sim$, was ein Vektorraum bzgl der (wohldefinierten) Operationen $[(x_n)]+[(y_n)] := [(x_n+y_n)]$ und $\lambda[(x_n)] := [(\lambda x_n)]$ ist. Wir definieren auf Y eine Norm durch

$$\|[(x_n)]\| := \lim_{n\to\infty} \|x_n\|.$$

Dieser Limes existiert, da $\big|\|x_n\| - \|x_m\|\big| \leq \|x_n - x_m\| \to 0$, ist unabhängig von der Wahl eines Repräsentanten, und eine Norm. Dann ist Y zusammen mit $\Phi : X \to Y$, $x \mapsto [(x)]$ (hier bezeichnet (x) die konstante Folge $(x, x, \ldots)$) eine Vervollständigung von X (hier ist einiges zu beweisen, was in dieser Beweisskizze aber auslassen).

b) Nach Teil a) dürfen wir X als einen dichten Unterraum eines Banachraums Y auffassen. Sei $A : X \to Z$ stetig und linear (mit Z Banachraum). Da $X \subset Y$ dicht ist, gibt es zu jedem $y \in Y$ eine Folge $(x_n)_n \subset X$ mit $x_n \to y$ in Y. Wir wählen für jedes $x \in X$ eine solche Folge aus und definieren

$$\widehat{A} : Y \to Z, \qquad \widehat{A}y := \lim_{n\to\infty} Ax_n \quad \text{für } \ x_n \to y.$$

$(x_n)_n$ ist eine Cauchyfolge, und aufgrund der Stetigkeit von A gilt $\|Ax_n - Ax_m\| \leq \|A\|\|x_n - x_m\| \to 0$. Da Z vollständig ist, konvergiert $(Ax_n)_n$ in Z, dh der Limes in der Definition von $\widehat{A}$ existiert. Weiterhin ist $\widehat{A}$ wohldefiniert, hängt also nicht von der Wahl der Folge $(x_n)_n$ mit $x_n \to y$ ab: Ist (x'_n) eine weitere solche Folge, so gilt

$$\|Ax_n - Ax'_n\| \leq \|A\|\|x_n - x'_n\| \leq \|A\|\|x_n - y\| + \|A\|\|y - x'_n\| \to 0.$$

Man prüft leicht nach, dass $\widehat{A}$ linear ist. Offenbar gilt $\widehat{A}|_X = A$ (wähle die konstante approximierende Folge), dh $\widehat{A} \circ \Phi = A$ in der Sprache der Vervollständigungsabbildung. Die Stetigkeit von $\widehat{A}$ folgt aus

$$\|\widehat{A}y\| = \|\lim_{n\to\infty} Ax_n\| = \lim_{n\to\infty} \|Ax_n\| \leq \|A\| \lim_{n\to\infty} \|x_n\| = \|A\|\|y\|,$$

nämlich $\|\widehat{A}\| \leq \|A\|$ und mit $\widehat{A}|_X = A$ auch $\|\widehat{A}\| = \|A\|$.

Die Eindeutigkeit der Fortsetzung folgt, da Stetigkeit von $\widehat{A}$ die definierende Formel von $\widehat{A}$ erzwingt.

c) Seien (Y, Φ), $(\tilde{Y}, \tilde{\Phi})$ zwei Vervollständigungen von X. Dann ist $\Phi : X \to Y$ eine lineare stetige (da isometrische) Abbildung in einen Banachraum, die nach b) eine Fortsetzung $\widehat{\Phi}$ bzgl der Vervollständigung $(\tilde{Y}, \tilde{\Phi})$ hat, nämlich eine lineare stetige Abbildung $\widehat{\Phi} : \tilde{Y} \to Y$ mit $\widehat{\Phi}|X = \Phi$. Diese Abbildung ist isometrisch: $\|\widehat{\Phi}(y)\| = \|\lim_n \Phi x_n\| = \lim_n \|\Phi x_n\| = \lim_n \|x_n\| = \|y\|$ für eine Folge $x_n \to y$. Weiterhin ist $\widehat{\Phi}$ surjektiv: Nach Lemma 1.25, angewendet auf $\widehat{\Phi} : \tilde{Y} \to \widehat{\Phi}(\tilde{Y})$, ist das Bild von $\widehat{\Phi}$ vollständig (abgeschlossen). Andererseits enthält dieses Bild den Unterraum $\Phi(X)$, der in Y dicht liegt. Also ist $\widehat{\Phi}(\tilde{Y})$ sowohl abgeschlossen als dicht und damit gleich Y. $\square$

Aufgrund der in Teil c) etablierten Eindeutigkeit der Vervollständigung von X bis auf isometrischen Isomorphismus ist auch eine Notation wie $\widehat{X}$ für die Vervollständigung gerechtfertigt. Beachten Sie aber, dass aufgrund der abstrakten Konstruktion der Vervollständigung wenig konkrete Information über $\widehat{X}$ zur Verfügung steht. Zum Beispiel können wir die Vervollständigung von $C([0,1])$ bzgl der L^1-Norm betrachten. Aus der Maßtheorie wissen wir, dass $C([0,1]) \subset L^1([0,1])$ ein dichter Unterraum ist - also ist die Vervollständigung von $C([0,1])$ bzgl der L^1-Norm gleich $L^1([0,1])$, ein Raum von Äquivalenzklassen von Lebesgue-integrierbaren Funktionen. Wenn uns Maßtheorie nicht bekannt wäre, könnten wir trotzdem die abstrakte Vervollständigung betrachten. In diesem Fall würden wir wir aber nichts über die Elemente der Vervollständigung wissen und könnten sie nicht ohne weiteres als (Äquivalenzklassen von) integrierbaren Funktionen erkennen.

Als Anwendungsbeispiel von Vervollständigungen betrachten wir Integraloperatoren.

Beispiel 1.34. Sei $I := [a, b] \subset \mathbb{R}$ ein kompaktes Intervall und $K : I \times I \to \mathbb{C}$ eine stetige Funktion (Integralkern). Dann ist der Integraloperator

$$(T_K f)(x) := \int_a^b K(x, y) f(y)\, dy$$

eine lineare Abbildung $C(I) \to C(I)$ (für die Stetigkeit von $T_K f$ siehe [Ana3, Satz 2.27]). Wir betrachten die L^p-Norm und sehen mit Hilfe der Hölderschen Ungleichung mit $\frac{1}{p} + \frac{1}{q} = 1$

$$\begin{aligned}\|T_K f\|_{L^p}^p &= \int_a^b \left| \int_a^b K(x, y) f(y)\, dy \right|^p dx \leq \int_a^b \|K(x, \cdot\,)\|_{L^q}^p \, \|f\|_{L^p}^p \, dx \\ &\leq (b-a) \sup_{x \in I} \|K(x, \cdot\,)\|_{L^q}^p \cdot \|f\|_{L^p}^p .\end{aligned}$$

Das Supremum lässt sich wie folgt abschätzen: Für jedes $x \in I$ gilt

$$\|K(x, \cdot\,)\|_{L^q}^q = \int_a^b |K(x, y)|^q dy \leq \|K\|_\infty^q \cdot (b-a),$$

was $\|T_K\| \leq (b-a)^{1/p}(b-a)^{1/q}\|K\|_\infty = (b-a)\|K\|_\infty < \infty$ ergibt. Fassen wir T_K als Abbildung von $C(I)$ in den größeren Banachraum $L^p(I)$ auf, so können wir mit Satz 1.33 b) schließen, dass der Integraloperator eine stetige Ausdehnung zu einem Operator $L^p(I) \to L^p(I)$ hat.

2 Hilberträume

In diesem Kapitel beschäftigen wir uns mit Hilberträumen. Hilberträume sind spezielle Banachräume, in denen die Norm durch ein Skalarprodukt induziert wird.

2.1 Skalarprodukte und Hilberträume

Auf dem endlichdimensionalen Raum $\mathbb{K}^d$ (wieder mit $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$) haben wir das Euklidische Skalarprodukt $\langle x, y\rangle := \sum_{k=1}^d \overline{x_k} y_k$, dass die Euklidische Norm $\|\cdot\|_2$ durch $\|x\|_2 = \sqrt{\langle x, x\rangle}$ festlegt. Wir verallgemeinern diese Situation nun auf andere Räume.

Definition 2.1. Sei V ein Vektorraum über $\mathbb{K}$.

a) Eine *Sesquilinearform* ist eine Abbildung $s : V \times V \to \mathbb{K}$, die im zweiten Argument linear und im ersten Argument antilinear[a] ist, dh

$$\begin{aligned} s(v, \cdot\,) &: V \to \mathbb{K}, \qquad w \mapsto s(v, w) \qquad \text{ist linear,} \\ s(\,\cdot\,, v) &: V \to \mathbb{K}, \qquad w \mapsto s(w, v) \qquad \text{ist antilinear.} \end{aligned}$$

b) Eine *hermitische Form* ist eine Sesquilinearform $s : V \times V \to \mathbb{K}$, die

$$s(v, w) = \overline{s(w, v)}, \qquad v, w \in V,$$

erfüllt.

c) Eine hermitische Form heißt *positiv semidefinit* (bzw *positiv definit*), wenn $s(v, v) \geq 0$ (bzw $s(v, v) > 0$) für alle $v \in V \setminus \{0\}$ gilt.

d) Ein *Skalarprodukt* ist eine positiv definite hermitische Form.

e) Ein *Prä-Hilbertraum* ist ein Vektorraum mit einem Skalarprodukt.

[a] Eine antilineare Abbildung $A : V \to W$ zwischen komplexen Räumen erfüllt per Definition $A(v_1 + v_2) = Av_1 + Av_2$ und $A(\lambda v) = \overline{\lambda} A(v)$ für $v, v_1, v_2 \in V$, $\lambda \in \mathbb{C}$. Für reelle Räume ist Antilinearität das gleiche wie Linearität.

Beachten Sie, dass eine Sesquilinarform s stets $s(v, v) = 0$ für $v = 0$ erfüllt. Im Falle eines Skalarproduktes gilt per Definition auch $s(v, v) = 0 \Rightarrow v = 0$. Skalarprodukte werden meist als $\langle v, w\rangle = s(v, w)$ notiert.

Beispiel 2.2. Auf $V = C([a, b])$ ist

$$\langle f, g\rangle := \int_a^b \overline{f(x)} g(x)\, dx$$

ein Skalarprodukt, das sogenannte L^2-Skalarprodukt. Beweis: Übung.

Definition 2.3. Zwei Vektoren x, y in einem Prähilbertraum V heißen *orthogonal*, geschrieben $x \perp y$, wenn $\langle x, y\rangle = 0$. Eine *orthogonale Menge* ist eine Teilmenge $M \subset V$ mit der Eigenschaft $x \perp y$ für alle $x, y \in M$, $x \neq y$. Eine *orthonormale Menge* ist

> eine orthogonale Menge $M \subset V$, für die $\langle x, x\rangle = 1$ für alle $x \in M$ gilt. Die selben Bezeichnungen werden auch für eine positiv semidefinite hermitische Form s verwendet.

Für eine orthonormale Menge M gilt also

$$\langle x, y\rangle = \delta_{x,y} = \begin{cases} 1 & x = y \\ 0 & x \neq y \end{cases}.$$

Dies gilt entsprechend für nur positiv *semi*definite hermitische Formen. Im Kontext eines Prähilbertraums $(V, \langle\,\cdot\,,\,\cdot\,\rangle)$ führen wir die Abkürzung

$$\|v\| := \sqrt{\langle v, v\rangle} \in [0, \infty)$$

ein. Wir werden bald sehen, dass dies eine Norm ist (die Dreiecksungleichung ist nicht offensichtlich). Im Falle einer positiv semidefiniten Form s ist $v \mapsto \sqrt{s(v,v)}$ nur eine Halbnorm, da es Vektoren $v \neq 0$ geben kann, die $s(v, v) = 0$ erfüllen.

Der aus der Schule bekannte Satz von Pythagoras besagt, dass für zwei bzgl. des Euklidischen Skalarproduktes auf $\mathbb{R}^2$ aufeinander senkrecht stehende Vektoren $x, y \in \mathbb{R}^2$

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

gilt, was wir sofort mit Hilfe der Skalarprodukteigenschaften nachrechnen können:

$$\|x - y\|^2 = \langle x - y, x - y\rangle = \langle x, x\rangle - \langle x, y\rangle - \langle y, x\rangle + \langle y, y\rangle = \|x\|^2 + \|y\|^2.$$

Der folgende Satz verallgemeinert dies. Es genügt hier, positive Semidefinitheit anzunehmen.

> **Satz 2.4 (Satz von Pythagoras).** *Sei $\{v_1, \ldots, v_N\}$ eine orthonormale Menge bzgl. einer positiv semidefiniten hermiteschen Form auf einem Vektorraum V. Dann gilt für alle $x \in V$*
>
> $$s(x, x) = \sum_{n=1}^{N} |s(x, v_n)|^2 + s\left(x - \sum_{n=1}^{N} s(v_n, x)v_n,\ x - \sum_{n=1}^{N} s(v_n, x)v_n\right).$$
>
> *Insbesondere gilt die* Besselsche Ungleichung
>
> $$s(x, x) \geq \sum_{n=1}^{N} |s(x, v_n)|^2$$
>
> *für alle $x \in V$.*

Beweis. Sei $x \in V$. Wir setzen $u := \sum_{n=1}^{N} s(v_n, x)v_n$ und $w := x - u$. Dann sind u und w

orthogonal, denn aufgrund der Eigenschaften von s gilt

$$\begin{aligned}
s(u,w) &= s\left(\sum_{n=1}^{N} s(v_n,x)v_n, x - \sum_{m=1}^{N} s(v_m,x)v_m\right) \\
&= \sum_{n=1}^{N} \overline{s(v_n,x)}s(v_n,x) - \sum_{n,m=1}^{N} \overline{s(v_n,x)}s(v_m,x)\underbrace{s(v_n,v_m)}_{=\delta_{nm}} \\
&= \sum_{n=1}^{N} |s(v_n,x)|^2 - \sum_{n=1}^{N} |s(v_n,x)|^2 \\
&= 0.
\end{aligned}$$

Das impliziert

$$s(x,x) = s(u+w,u+w) = s(u,u) + s(u,w) + s(w,u) + s(w,w) = s(u,u) + s(w,w),$$

die behauptete Gleichung. Die Besselsche Ungleichung folgt nun aus $s(w,w) \geq 0$, da s positiv semidefinit ist. $\square$

Die folgende Ungleichung ist im Kontext von Hilberträumen von überragender Bedeutung und wird sehr häufig benutzt.

Lemma 2.5. *Sei $s : V \times V \to \mathbb{K}$ eine positiv semidefinite hermitesche Form auf einem $\mathbb{K}$-Vektorraum V. Dann gilt die* Cauchy-Schwarzsche Ungleichung

$$|s(v,w)|^2 \leq s(v,v) \cdot s(w,w), \qquad v,w \in V.$$

Beweis. Wir betrachten $v,w \in V$ und Skalare $\lambda,\mu \in \mathbb{K}$. Dann gilt aufgrund der positiven Semidefinitheit

$$\begin{aligned}
0 \leq s(\lambda v + \mu w, \lambda v + \mu w) &= |\lambda|^2 s(v,v) + \overline{\lambda}\mu s(v,w) + \lambda\overline{\mu} s(w,v) + |\mu|^2 s(w,w) \\
&= \left\langle \begin{pmatrix} \lambda \\ \mu \end{pmatrix}, A \begin{pmatrix} \lambda \\ \mu \end{pmatrix} \right\rangle, \qquad A := \begin{pmatrix} s(v,v) & s(v,w) \\ s(w,v) & s(w,w) \end{pmatrix},
\end{aligned}$$

wobei $\langle \cdot\,, \cdot \rangle$ das Euklidische Skalarprodukt auf $\mathbb{C}^2$ bezeichnet. Also ist die Matrix A positiv semidefinit und hat deshalb nicht-negative Determinante,

$$0 \leq \det A = s(v,v)s(w,w) - s(v,w)s(w,v) = s(v,v)s(w,w) - |s(v,w)|^2.$$

Durch Umstellen folgt die Cauchy-Schwarz Ungleichung. $\square$

Im Falle eines Skalarproduktes gibt es einen einfacheren Beweis der Cauchy-Schwarz Ungleichung: Wenden Sie die Bessel-Ungleichung auf die orthonormale Familie $\{\frac{v}{\|v\|}\}$ (mit $v \neq 0$ an.)

Satz 2.6. *Sei V ein Prähilbertraum mit Skalarprodukt $\langle\,\cdot\,,\,\cdot\,\rangle$. Dann ist $\|v\| := \sqrt{\langle v, v\rangle}$ eine Norm auf V, die*

$$|\langle v, w\rangle| \leq \|v\|\|w\|$$

(Cauchy-Schwarz Ungleichung) und die Paralellogramm-Identität[a]

$$\|v+w\|^2 + \|v-w\|^2 = 2\|v\|^2 + 2\|w\|^2, \qquad v, w \in V,$$

erfüllt. Ist stattdessen $\langle\,\cdot\,,\,\cdot\,\rangle$ nur eine positiv semidefinite hermitische Form, so ist $\|\cdot\|$ nur eine Halbnorm.

[a] Um den Namen dieser Identität zu verstehen, zeichnen Sie ein Parallelogramm in $\mathbb{R}^2$, das von zwei Vektoren v, w aufgespannt wird.

Beweis. Offenbar gilt $\|v\| \geq 0$ für alle $v \in V$, und im Falle eines Skalarproduktes gilt auch $\|v\| = 0 \Leftrightarrow v = 0$. Weiterhin haben wir für $\lambda \in \mathbb{K}$ auch $\|\lambda v\| = |\lambda|\|v\|$ aufgrund der Sesquilinearität des Skalarproduktes. Für den Beweis der Dreiecksungleichung benutzen wir die Cauchy-Schwarz Ungleichung $|\langle v, w\rangle| \leq \|v\|\|w\|$:

$$\begin{aligned}
\|v+w\|^2 &= \langle v, v\rangle + \langle v, w\rangle + \langle w, v\rangle + \langle w, w\rangle \\
&= \|v\|^2 + 2\mathrm{Re}\,\langle v, w\rangle + \|w\|^2 \\
&\leq \|v\|^2 + 2|\langle v, w\rangle| + \|w\|^2 \\
&\leq \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 \\
&= (\|v\| + \|w\|)^2.
\end{aligned}$$

Die Parallelogramm-Ungleichung folgt durch einfaches Ausmultiplizieren:

$$\begin{aligned}
\|v+w\|^2 + \|v-w\|^2 &= \langle v, v\rangle + \langle v, w\rangle + \langle w, v\rangle + \langle w, w\rangle \\
&\quad + \langle v, v\rangle - \langle v, w\rangle - \langle w, v\rangle + \langle w, w\rangle \\
&= 2\|v\|^2 + 2\|w\|^2.
\end{aligned}$$

□

Sei $\mathcal{H}$ ein Prähilbertraum und $\psi, \varphi \in \mathcal{H}$. Unter welchen Bedingungen gilt Gleichheit in der Cauchy-Schwarz Ungleichung, dh wann gilt $|\langle \psi, \varphi\rangle| = \|\psi\|\|\varphi\|$?

Wir sehen also, dass jeder Prähilbertraum ein normierter Raum mit einer durch das Skalarprodukt definierten Norm ist. Dies führt uns auf den Begriff eines Hilbertraums.

Definition 2.7. Ein *Hilbertraum* ist ein Prähilbertraum, der in der durch sein Skalarprodukt induzierten Norm vollständig ist.

Insbesondere ist jeder Hilbertraum ein Banachraum. Die Umkehrung dieser Aussage ist falsch: Nicht jeder Banachraum ist ein Hilbertraum. Wie oben gezeigt, muss die Norm eines Banachraums mindestens die Parallelogrammgleichung erfüllen, um ein Hilbertraum sein zu können. Tatsächlich ist dies die einzige Bedingung.

Satz 2.8. *Ein Banachraum $(V, \|\cdot\|)$ ist ein Hilbertraum genau dann, wenn $\|\cdot\|$ die Parallelogramm-Identität erfüllt. In diesem Fall ist*

$$\langle v, w\rangle := \frac{1}{4}\left(\|v+w\|^2 - \|v-w\|^2 + i\|v-iw\|^2 - i\|v+iw\|^2\right) \qquad (\textit{für } \mathbb{K}=\mathbb{C}),$$
$$\langle v, w\rangle := \frac{1}{4}\left(\|v+w\|^2 - \|v-w\|^2\right) \qquad (\textit{für } \mathbb{K}=\mathbb{R})$$

ein Skalarprodukt auf V, das die Norm $\|v\| = \sqrt{\langle v, v\rangle}$ induziert.

Wir sehen also insbesondere, dass sich in einem Hilbertraum das Skalarprodukt durch die Norm ausdrücken lässt. Diese Gleichung heißt *Polarisationsidentität.* Andererseits definiert nicht jede Norm ein Skalarprodukt durch Polarisation, Voraussetzung ist die Parallelogramm-Identität.

Beweis. Wir zeigen zuerst, dass in einem Hilbertraum die Polarisationsidentität gilt. Dies ist ein reines Ausmultiplizieren: Für $\mathbb{K}=\mathbb{C}$ gilt

$$\begin{aligned}\|v+w\|^2 - \|v-w\|^2 + i\|v-iw\|^2 - i\|v+iw\|^2 &= \langle v+w, v+w\rangle - \langle v-w, v-w\rangle \\ &\quad + i\langle v-iw, v-iw\rangle - i\langle v+iw, v+iw\rangle \\ &= 4\langle v, w\rangle.\end{aligned}$$

Die Rechnung im Fall $\mathbb{K}=\mathbb{R}$ ist analog und einfacher.[2]

Wir haben bereits gesehen, dass die Norm eines Hilbertraumes die Parallelogramm-Identität erfüllt. Der Beweis, dass die durch die Polarisationsidentität definierte Abbildung $\langle\cdot,\cdot\rangle$ ein Skalarprodukt ist, wird in den Übungen geführt. □

Beispiel 2.9.

a) $\mathbb{K}^d$ ist ein Hilbertraum mit dem Euklidischen Skalarprodukt

$$\langle z, w\rangle_2 := \sum_{k=1}^{d} \overline{z_k} w_k \qquad z = (z_1, \dots, z_d),\ w = (w_1, \dots, w_d) \in \mathbb{K}^d.$$

Die Skalarprodukteigenschaften lassen sich leicht prüfen. Die induzierte Norm ist $\|z\|_2 = \sqrt{\langle z, z\rangle_2} = \sum_{k=1}^d |z_k|^2$.

Da $\mathbb{K}^d$ endlichdimensional ist, ist er automatisch vollständig bzgl $\|\cdot\|_2$.

b) Der Banachraum $(\mathbb{K}^d, \|\cdot\|_p)$ (siehe Beispiel 1.3) ist für $p \neq 2$ kein Hilbertraum. Um dies einzusehen, müssen wir zeigen, dass die Parallelogrammgleichung nicht erfüllt ist (Übung).

[2]Die Rechnung zum Nachprüfen der Polarisationsidentität zeigt auch, dass eine positiv semi-definite Sesquilinearform automatisch hermitesch ist, was zum Nachprüfen der Skalarprodukteigenschaften hilfreich ist.

c) $\ell^2_{\mathbb{N}}$ (siehe Beispiel 1.4) ist ein Hilbertraum mit dem Skalarprodukt

$$\langle x, y\rangle := \sum_{n=1}^{\infty} \overline{x_n} y_n, \qquad x = (x_n)_{n\in\mathbb{N}},\ y = (y_n)_{n\in\mathbb{N}} \in \ell^2_{\mathbb{N}}.$$

Wir müssen zuerst zeigen, dass das Skalarprodukt wohldefiniert ist, dh dass die Reihe konvergiert. Für $x \in \ell^2_{\mathbb{N}}$ gilt $\|x\|_2^2 = \sum_{n=1}^{\infty} |x_n|^2 < \infty$.

Dazu verwenden wir die Cauchy-Schwarz Ungleichung in $\mathbb{C}^N$: Für $N \in \mathbb{N}$ gilt

$$\left|\sum_{n=1}^{N} \overline{x_n} y_n\right| \le \left(\sum_{n=1}^{N} |x_n|^2\right)^{1/2} \left(\sum_{n=1}^{N} |y_n|^2\right)^{1/2} \le \|x\|_2 \|y\|_2 \,.$$

Da N beliebig und die rechte Seite von N unabhängig ist, ergibt sich $|\langle x, y\rangle| \le \|x\|_2\|y\|_2 < \infty$. Also ist das Skalarprodukt wohldefiniert. Die Skalarprodukteigenschaften lassen sich einfach nachprüfen.

d) Der Banachraum $(\ell^p_{\mathbb{N}}, \|\cdot\|_p)$ ist für $p \neq 2$ kein Hilbertraum (analog zu b)).

e) Für einen Maßraum (Ω, Σ, μ) ist $L^2(\Omega, \mu)$ (siehe Beispiel 1.6) ein Hilbertraum mit dem Skalarprodukt

$$\langle f, g\rangle = \int_{\Omega} \overline{f(x)} g(x) d\mu(x),$$

wie in Analysis 3 gezeigt wird. Für $p \neq 2$ ist der Banachraum $(L^p(\Omega, \mu), \|\cdot\|_{L^p})$ kein Hilbertraum.

Diese Beispiele zeigen, dass Hilberträume spezielle Banachräume sind – die meisten Banachräume sind keine Hilberträume. Wir werden sehen, dass sich in Hilberträumen aufgrund des Skalarproduktes in vielerlei Hinsicht einfacher arbeiten lässt als in allgemeinen Banachräumen. Auch in Anwendungen spielen Hilberträume eine hervorragende Rolle.

Eine wichtige Konsequenz der Cauchy-Schwarz Ungleichung ist die Stetigkeit des Skalarproduktes.

Lemma 2.10. *Das Skalarprodukt eines Prähilbertraumes $\mathcal{H}$ ist eine stetige Abbildung*

$$\langle\,\cdot\,,\,\cdot\,\rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{K}.$$

Beweis. Die Norm eines normierten Raumes ist offenbar stetig: Aus $v_n \to v$ folgt $\|v_n\| \to \|v\|$ (siehe Lemma 1.10 c)). Nach der Polarisationsidentität in Satz 2.8 lässt sich das Skalarprodukt $\langle v, w\rangle$ als Polynom in Normen, ausgewertet auf Linearkombinationen der beiden Vektoren v, w, ausdrücken. Da auch die linearen Operationen stetig sind (Lemma 1.10)), folgt die Behauptung. $\square$

Wir betrachten kurz zwei Konzepte, die wir bereits im Kontext von Banachräumen kennengelernt haben: Isomorphie und direkte Summen von Hilberträumen. Da Hilberträume durch ein Skalarprodukt und nicht durch eine Norm charakterisiert sind, definieren wir Isomorphie wie folgt.

Definition 2.11. Zwei Hilberträume $(\mathcal{H}_1, \langle\,\cdot\,,\,\cdot\,\rangle_1)$ und $(\mathcal{H}_2, \langle\,\cdot\,,\,\cdot\,\rangle_2)$ heißen isomorph (als Hilberträume), falls es eine lineare bijektive Abbildung $U : \mathcal{H}_1 \to \mathcal{H}_2$ gibt, so dass

$$\langle Ux, Uy\rangle_2 = \langle x, y\rangle_1, \qquad x, y \in \mathcal{H}_1.$$

Eine solche Abbildung U heißt *unitär* oder *unitärer Operator*. In diesem Fall schreiben wir auch

$$\mathcal{H}_1 \cong \mathcal{H}_2.$$

Wir bemerken, dass eine unitäre Abbildung $U : \mathcal{H}_1 \to \mathcal{H}_2$ eine Isometrie für die von den Skalarprodukten induzierten Normen ist: Für alle $x \in \mathcal{H}_1$ gilt

$$\|Ux\|_2 = \sqrt{\langle Ux, Ux\rangle_2} = \sqrt{\langle x, x\rangle_1} = \|x\|_1.$$

Betrachten wir die Hilberträume $\mathcal{H}_1, \mathcal{H}_2$ als Banachräume mit den von ihren Skalarprodukten induzierten Normen, so ist ein unitärer Operator also nichts anderes als ein isometrischer Isomorphismus.

Wir betrachten nun direkte Summen von Hilberträumen. Als Vektorraum ist $\mathcal{H}_1 \oplus \mathcal{H}_2$ wieder als kartesisches Produkt $\mathcal{H}_1 \times \mathcal{H}_2$ genau wie für Banachräume definiert. Aber wir müssen nun ein Skalarprodukt statt einer Norm definieren, um $\mathcal{H}_1 \times \mathcal{H}_2$ zu einem Hilbertraum zu machen.

Definition 2.12. Seien $(\mathcal{H}_1, \langle\,\cdot\,,\,\cdot\,\rangle_1)$ und $(\mathcal{H}_2, \langle\,\cdot\,,\,\cdot\,\rangle_2)$ Hilberträume. Dann ist die *direkte Summe* $\mathcal{H}_1 \oplus \mathcal{H}_2$ der Vektorraum $\mathcal{H}_1 \times \mathcal{H}_2$ zusammen mit dem Skalarprodukt

$$\langle (x_1, x_2), (y_1, y_2)\rangle = \langle x_1, y_1\rangle_1 + \langle x_2, y_2\rangle_2.$$

Aus der Definition ist direkt ersichtlich, dass $\langle\,\cdot\,,\,\cdot\,\rangle$ ein Skalarprodukt auf $\mathcal{H}_1 \times \mathcal{H}_2$ ist. Da die induzierte Norm die 2-Norm ist, dh $\langle (x_1, x_2), (x_1, x_2)\rangle^{1/2} = (\|x_1\|_1^2 + \|x_2\|_2^2)^{1/2}$ und wir bereits wissen, dass $\mathcal{H}_1 \times \mathcal{H}_2$ vollständig ist in bzgl dieser Norm, sehen wir auch, dass die direkte Summe von zwei Hilberträumen wieder ein Hilbertraum ist.

Bemerken Sie, dass bzgl. des Skalarproduktes von $\mathcal{H}_1 \oplus \mathcal{H}_2$ die Unterräume $\mathcal{H}_1 \oplus \{0\}$ und $\{0\} \oplus \mathcal{H}_2$ senkrecht aufeinander stehen; wir haben hier also eine orthogonale direkte Summe definiert.

2.2 Orthogonalität und adjungierte Operatoren

Wir befassen uns nun eingehender mit Orthogonalität in Hilberträumen. Dazu definieren wir das orthogonale Komplement einer Menge.

Definition 2.13. Das *orthogonale Komplement* $M^\perp$ einer Teilmenge M eines Prähilbertraums $\mathcal{H}$ ist

$$M^\perp := \{v \in \mathcal{H} : \langle v, m\rangle = 0\ \forall m \in M\}.$$

Das orthogonale Komplement $M^\perp$ von M besteht also aus allen Vektoren, die auf allen Vektoren aus M senkrecht stehen. Daraus ergibt sich direkt:

Lemma 2.14. *Sei $\mathcal{H}$ ein Prähilbertraum und $M \subset \mathcal{H}$ eine Teilmenge. Dann gilt*

a) $M^\perp$ *ist ein abgeschlossener Unterraum von $\mathcal{H}$,*

b) $M \subset (M^\perp)^\perp$,

c) Ist $N \subset \mathcal{H}$ eine weitere Teilmenge, so gilt $N \subset M \Rightarrow M^\perp \subset N^\perp$,

d) $M^\perp = \overline{\operatorname{span}(M)}^\perp$ *(hier bezeichnet $\overline{\cdot}$ den Abschluss in $\mathcal{H}$ in der Norm).*

Beweis. a) Sei $v, w \in M^\perp$, $\lambda \in \mathbb{K}$, und $m \in M$. Dann gilt

$$\langle \lambda v + w, m\rangle = \bar{\lambda}\langle v, m\rangle + \langle w, m\rangle = 0,$$

also $\lambda v + w \in M^\perp$. Dies zeigt, dass $M^\perp$ ein Unterraum ist. Sei nun $(m_n^\perp)_{n\in\mathbb{N}} \subset M^\perp$ eine Folge in $M^\perp$, die in $\mathcal{H}$ konvergiert, $m_n^\perp \to x \in \mathcal{H}$. Dann gilt aufgrund der Stetigkeit des Skalarproduktes für alle $m \in M$

$$\langle x, m\rangle = \langle \lim_{n\to\infty} m_n^\perp, m\rangle = \lim_{n\to\infty} \langle m_n^\perp, m\rangle = \lim_{n\to\infty} 0 = 0.$$

Also gilt $x \in M^\perp$, dh $M^\perp$ ist abgeschlossen.

b) Sei $m \in M$ und $w \in M^\perp$. Dann gilt $\langle m, w\rangle = 0$, also $m \in (M^\perp)^\perp$.

c) Falls ein Vektor auf allen Vektoren in M senkrecht steht, so steht er insbesondere auf allen Vektoren in der kleineren Menge N senkrecht.

d) Wegen $M \subset \overline{\operatorname{span}(M)}$ gilt nach c) $M^\perp \supset \overline{\operatorname{span}(M)}^\perp$. Wir müssen also nur $M^\perp \subset \overline{\operatorname{span}(M)}^\perp$ zeigen. Sei also $m^\perp \in M^\perp$ und $x \in \overline{\operatorname{span}(M)}$. Dann gibt es eine Folge $(x_n) \subset \operatorname{span}(M)$ mit $x_n \to x$. Da $m^\perp$ auf M senkrecht steht, steht es auch auf den Linearkombinationen x_n in der linearen Hülle von M senkrecht, dh $0 = \langle m^\perp, x_n\rangle$ für alle n. Aufgrund der Stetigkeit von $\langle \cdot, \cdot \rangle$ folgt nun $0 = \langle m^\perp, x\rangle$, also $m^\perp \in \overline{\operatorname{span}(M)}^\perp$. □

Ein fundamentales Resultat, in dem sich Hilberträume besser als Banachräume verhalten, ist eine natürliche Zerlegung in direkte orthogonale Summen, die wir nun besprechen. Wir beginnen mit einem Lemma.

Lemma 2.15. *Sei $\mathcal{H}$ ein Hilbertraum, $\mathcal{K} \subset \mathcal{H}$ ein abgeschlossener Unterraum, und $v \in \mathcal{H}$ ein Vektor. Dann gibt es einen eindeutigen Vektor $k_0 \in \mathcal{K}$ mit der Eigenschaft*

$$\|v - k_0\| = d(v, \mathcal{K}) := \inf\{\|v - k\| : k \in \mathcal{K}\}.$$

Beweis. Nach Definition von $d(v, \mathcal{K}) = \inf\{\|v - k\| : k \in \mathcal{K}\}$ gibt es eine Folge $k_n \in \mathcal{K}$, so dass $\|v - k_n\| \to d(v, \mathcal{K})$ für $n \to \infty$. Dies ist eine Cauchyfolge, denn

$$\begin{aligned}
\|k_n - k_m\|^2 &= \|(k_n - v) - (k_m - v)\|^2 \\
&= 2\|k_n - v\|^2 + 2\|k_m - v\|^2 - \|(k_n - v) + (k_m - v)\|^2 \quad \text{(Parallelogramm-Gl.)} \\
&= 2\|k_n - v\|^2 + 2\|k_m - v\|^2 - 4\|v - \tfrac{1}{2}(k_n - k_m)\|^2 \\
&\leq 2\|k_n - v\|^2 + 2\|k_m - v\|^2 - 4d(v, \mathcal{K})^2 \\
&\longrightarrow 2d(v, \mathcal{K})^2 + 2d(v, \mathcal{K})^2 - 4d(v, \mathcal{K})^2 \quad \text{für } n, m \to \infty \\
&= 0.
\end{aligned}$$

Da $\mathcal{H}$ nach Voraussetzung vollständig ist, existiert der Limes $k_0 := \lim_n k_n$ und liegt in $\mathcal{K}$, da $\mathcal{K}$ abgeschlossen ist. Es gilt

$$\|v - k_0\| = \lim_{n\to\infty} \|v - k_n\| = d(v, \mathcal{K}),$$

was die Existenzaussage beweist. Die Eindeutigkeitsaussage beweist sich analog: Falls k_0 und k_0' in $\mathcal{K}$ liegen und $\|v - k_0\| = \|v - k_0'\| = d(v, \mathcal{K})$ erfüllen, so liefert obige Rechnung mit k_0 anstelle von k_n und k_0' anstelle von k_m

$$\|k_0 - k_0'\|^2 \leq 2\|k_0 - v\|^2 + 2\|k_0' - v\|^2 - 4d(v, \mathcal{K})^2 = 0,$$

also $k_0 = k_0'$. □

Unsere geometrische Anschauung in endlicher Dimension legt die Aussage dieses Lemmas nahe. Beachten Sie aber, dass es in allgemeinen Banachräumen *nicht* gilt.

Wir kommen nun zu der versprochenen Aufteilung eines Hilbertraums in orthogonale Unterräume.

Satz 2.16 (Projektionstheorem). *Sei $\mathcal{K} \subset \mathcal{H}$ ein abgeschlossener Unterraum eines Hilbertraums $\mathcal{H}$. Dann gilt*

$$\mathcal{H} \cong \mathcal{K} \oplus \mathcal{K}^\perp.$$

Insbesondere kann jeder Vektor $v \in \mathcal{H}$ auf eindeutige Weise als $v = k + k^\perp$ mit $k \in \mathcal{K}$ und $k^\perp \in \mathcal{K}^\perp$ geschrieben werden, und es gilt dann $\|v\|^2 = \|k\|^2 + \|k^\perp\|^2$.

Beweis. Sei $v \in \mathcal{H}$. Nach dem vorausgehenden Lemma gibt es einen eindeutigen Vektor $k \in \mathcal{K}$ mit $\|v - k\| = d(v, \mathcal{K})$. Wir setzen $k^\perp := v - k$ und behaupten $k^\perp \in \mathcal{K}^\perp$. Für alle $y \in \mathcal{K}$ und alle $t \in \mathbb{R}$ gilt

$$d(v, \mathcal{K})^2 \leq \|v - (k + t \cdot y)\|^2 = \|k^\perp - t \cdot y\|^2 = d(v, \mathcal{K})^2 - 2t \operatorname{Re} \langle k^\perp, y\rangle + t^2 \|y\|^2,$$

also $-2t \operatorname{Re} \langle k^\perp, y\rangle + t^2 \|y\|^2 \geq 0$ für alle $t \in \mathbb{R}$. Dies impliziert $\operatorname{Re} \langle k^\perp, y\rangle = 0$ (betrachte Graph dieser Parabel). Führen wir das Argument analog mit it statt t durch, so erhalten wir $\operatorname{Im} \langle k^\perp, y\rangle = 0$, insgesamt also $k^\perp \in \mathcal{K}^\perp$.

Um die Eindeutigkeit der Zerlegung zu zeigen, nehmen wir $v = k_1 + k_1^\perp = k_2 + k_2^\perp$ mit $k_1, k_2 \in \mathcal{K}$ und $k_1^\perp, k_2^\perp \in \mathcal{K}^\perp$ an. Dann gilt $k_1 - k_2 = k_2^\perp - k_1^\perp \in \mathcal{K} \cap \mathcal{K}^\perp$. Aber $\mathcal{K} \cap \mathcal{K}^\perp = \{0\}$, denn jedes x in diesem Raum erfüllt $0 = \langle x, x\rangle = \|x\|^2$.

Die Normgleichung $\|v\|^2 = \|k\|^2 + \|k^\perp\|^2$ ist eine direkte Konsequenz der Orthogonalität von k und $k^\perp$.

Der gesuchte unitäre Operator, der den Hilbertraumisomorphismus $\mathcal{H} \cong \mathcal{K} \oplus \mathcal{K}^\perp$ vermittelt, ist

$$U : \mathcal{K} \oplus \mathcal{K}^\perp \to \mathcal{H}, \qquad U(k, k^\perp) := k + k^\perp.$$

Er ist nach oben Gesagtem surjektiv und aufgrund der Eindeutigkeit der Zerlegung injektiv. Die Unitarität folgt aus

$$\begin{aligned}\langle U(k_1, k_1^\perp), U(k_2, k_2^\perp)\rangle_{\mathcal{H}} &= \langle k_1 + k_1^\perp, k_2 + k_2^\perp\rangle_{\mathcal{H}} = \langle k_1, k_2\rangle_{\mathcal{H}} + \langle k_1^\perp, k_2^\perp\rangle_{\mathcal{H}} \\ &= \langle (k_1, k_1^\perp), (k_2, k_2^\perp)\rangle_{\mathcal{K}\oplus\mathcal{K}^\perp}.\end{aligned}$$

□

Dieser Satz heißt das Projektionstheorem (oder der Satz von der orthogonalen Projektion), weil wir einen Hilbertraum in eine direkte orthogonale Summe zerlegen können. Dies definiert auch eine Projektion auf den abgeschlossenen Unterraum: Die lineare Abbildung

$$P_{\mathcal{K}} : \mathcal{H} \to \mathcal{K} \oplus \mathcal{K}^{\perp}, \qquad P_{\mathcal{K}}(k + k^{\perp}) := k,$$

heißt *die orthogonale Projektion auf* $\mathcal{K}$, wobei $k \in \mathcal{K}$, $k^{\perp} \in \mathcal{K}^{\perp}$. Es ist klar, dass $P_{\mathcal{K}}^2 = P_{\mathcal{K}}$ gilt. Weiterhin beobachten wir ohne Schwierigkeiten

$$\operatorname{Ran} P_{\mathcal{K}} = \mathcal{K}, \qquad \ker P_{\mathcal{K}} = \mathcal{K}^{\perp},$$

und $P_{\mathcal{K}}$ ist eine orthogonale Projektion in dem Sinne, dass $\operatorname{Ran} P_{\mathcal{K}}$ und $\ker P_{\mathcal{K}}$ aufeinander senkrecht stehen.

a) Definieren wir eine orthogonale Projektion als einen Operator $P \in \mathcal{B}(\mathcal{H})$, der $P^2 = P$ und $(\ker P)^{\perp} = \operatorname{Ran} P$ erfüllt, so erhalten wir eine Bijektion zwischen abgeschlossenen Unterräumen von $\mathcal{H}$ und orthogonalen Projektionen (s.o.). Diese erlaubt es, geometrische Fragen über die relative Postion von Unterräumen (zB: ein Unterraum enthält einen anderen, Schnitt von Unterräumen, zwei Unterräume stehen senkrecht, etc) in Operatorensprache zu übersetzen, wo sie oft effektiver bearbeitet werden können. Wir werden zu diesen Fragen zurückkommen, wenn wir weitere Werkzeuge zur Diskussion von Hilbertraumoperatoren zur Verfügung haben.

b) Orthogonale Projektionen sind wesentliche elementare Operatoren, aus denen kompliziertere Operatoren durch Linearkombination und Integration zusammengesetzt werden; denken Sie zB an die Projektionen auf die Eigenräume einer Matrix im endlichdimensionalen Fall. Wir werden diese Betrachtungen im Kapitel zur Spektraltheorie weiterführen.

c) Abgeschlossene Unterräume treten in Paaren $(\mathcal{K}, \mathcal{K}^{\perp})$ auf, was auf der Ebene der Projektionen durch die Paare $(P_{\mathcal{K}}, 1 - P_{\mathcal{K}})$ reflektiert ist (s.u.). In Anwendungen in der Quantenphysik modellieren orthogonale Projektionen Ja/Nein-Aussagen, und $1 - P$ entspricht der Negation der zu P gehörigen Aussage (Quantenlogik).

Korollar 2.17. *Sei $\mathcal{K} \subset \mathcal{H}$ ein abgeschlossener Unterraum eines Hilbertraums $\mathcal{H}$. Dann gilt*

$$P_{\mathcal{K}^{\perp}} = 1 - P_{\mathcal{K}}$$

(mit $1 = id_{\mathcal{H}}$). Ist $U \subset \mathcal{H}$ ein nicht notwendigerweise abgeschlossener Unterraum von $\mathcal{H}$, so gilt

$$\overline{U} = (U^{\perp})^{\perp}.$$

Beweis. Da auch das orthogonale Komplement $\mathcal{K}^{\perp}$ ein abgeschlossener Unterraum von $\mathcal{H}$ ist, ist die orthogonale Projektion $P_{\mathcal{K}^{\perp}}$ definiert. Auf $v = k + k^{\perp}$ gilt

$$P_{\mathcal{K}^{\perp}}(k + k^{\perp}) = k^{\perp} = (k + k^{\perp}) - k = (k + k^{\perp}) - P_{\mathcal{K}}(k + k^{\perp}),$$

also $P_{\mathcal{K}^{\perp}} = 1 - P_{\mathcal{K}}$ wie behauptet.

Für beliebige Unterräume U wissen wir bereits aus Lemma 2.14 d) $U^\perp = \overline{U}^\perp$. Damit erhalten wir

$$P_{\overline{U}} = 1 - P_{\overline{U}^\perp} = 1 - P_{U^\perp} = P_{(U^\perp)^\perp},$$

also $P_{\overline{U}} = P_{(U^\perp)^\perp}$. Damit folgt $\overline{U} = P_{\overline{U}}\mathcal{H} = P_{(U^\perp)^\perp}\mathcal{H} = (U^\perp)^\perp$. □

Es gibt zwei spezielle (triviale) Projektionen, die den trivialen Unterräumen $\mathcal{K} = \mathcal{H}$ und $\mathcal{K} = \{0\}$ entsprechen, nämlich

$$P_{\mathcal{H}} = 1, \qquad P_{\{0\}} = 0.$$

Beispiele zu Orthogonalzerlegungen finden sich in den Übungen.

Wir kommen nun zu einer wichtigen Anwendung des Projektionssatzes, dem sogenannten Lemma von Riesz, das den Dualraum eines Hilbertraums charakterisiert. Erinnern Sie sich, dass der Dualraum eines Banachraums (insbesondere also eines Hilbertraums) als $\mathcal{H}' := \mathcal{B}(\mathcal{H}, \mathbb{K})$ definiert ist, dh alle stetigen linearen Funktionale $\mathcal{H} \to \mathbb{K}$ enthält.

Satz 2.18 (Das Riesz'sche Lemma). *Sei $\mathcal{H}$ ein Hilbertraum. Für jeden Vektor $y \in \mathcal{H}$ ist das Funktional*

$$y^* = \langle y| : \mathcal{H} \to \mathbb{K}, \qquad y^*(x) := \langle y, x\rangle$$

stetig, und für jedes stetige lineare Funktional $\varphi \in \mathcal{H}'$ gibt es einen eindeutigen Vektor $y \in \mathcal{H}$, so dass $\varphi = y^*$. In diesem Fall gilt $\|\varphi\|_{\mathcal{H}'} = \|y\|_{\mathcal{H}}$.

Kurz gesagt: Die Abbildung

$$\Phi : \mathcal{H} \to \mathcal{H}', \qquad \Phi(y) = y^*$$

ist eine antilineare bijektive Isometrie.

Die Notation $\langle y|$ ist angesichts der Abbildungsvorschrift $\langle y| : x \mapsto \langle y, x\rangle$ sehr suggestiv – noch mehr, wenn die Vektoren $x \in \mathcal{H}$ als $|x\rangle$ notiert werden. Sie geht auf Dirac zurück und wird "Bra-Ket Notation"genannt, da die Klammer (bracket) des Skalarproduktes in den linken "Braünd den rechten "Ket"Teil zerlegt wird. Beachten Sie, dass hier die Konvention, das Skalarprodukt im rechten Eintrag linear zu wählen, wesentlich ist.

Das Riesz-Lemma wird auch als Rieszscher Darstellungssatz bezeichnet, da es die abstrakten stetigen linearen Funktionale konkret durch Vektoren darstellt.

Beweis. Für $y \in \mathcal{H}$ ist $y^* : x \mapsto \langle y, x\rangle$ ein lineares Funktional, dass aufgrund der Cauchy-Schwarz Ungleichung $|y^*(x)| \leq \|y\|\|x\|$ die Abschätzung $\|y^*\|_{\mathcal{H}'} \leq \|y\|_{\mathcal{H}}$ erfüllt. Einsetzen von $x = y$ zeigt $\|y^*\|_{\mathcal{H}'} = \|y\|_{\mathcal{H}}$.

Sei nun $\varphi \in \mathcal{H}'$. Um zu zeigen, dass $\varphi = y^*$ mit einem eindeutigen $y \in \mathcal{H}$ gilt, betrachten wir $\mathcal{K} := \ker\varphi$, einen Unterraum von $\mathcal{H}$. Aufgrund der Stetigkeit von φ ist $\mathcal{K}$ abgeschlossen.

Falls $\mathcal{K} = \mathcal{H}$, so ist $\varphi = 0$ eindeutig durch 0^* gegeben. Sei also $\mathcal{K} \neq \mathcal{H}$ und damit $\mathcal{K}^\perp \neq \{0\}$. Wähle $x_0 \in \mathcal{K}^\perp \backslash \{0\}$. Wir behaupten $\mathcal{H} = \mathcal{K} \oplus \mathbb{C}x_0$, dh $\mathcal{K}^\perp = \mathbb{C}x_0$ ist eindimensional.

Um dies zu zeigen, schreiben wir einen beliebigen Vektor $z \in \mathcal{H}$ als

$$z = \left(z - \frac{\varphi(z)}{\varphi(x_0)} x_0 \right) + \frac{\varphi(z)}{\varphi(x_0)} x_0 \in \mathcal{K} \oplus \mathbb{C}x_0.$$

Also stehen $\mathcal{K}$ und $\mathbb{C}x_0$ aufeinander senkrecht und spannen ganz $\mathcal{H}$ auf, dh $\mathcal{H} = \mathcal{K} \oplus \mathbb{C}x_0$.

Wir definieren nun

$$y := \overline{\varphi(x_0)} \|x_0\|^{-2} \cdot x_0$$

und behaupten $\varphi = y^*$. Wegen der orthogonalen Zerlegung $\mathcal{H} = \mathcal{K} \oplus \mathbb{C}x_0$ genügt es, diese Gleichung auf $\mathcal{K}$ und x_0 separat zu prüfen. Für $x \in \mathcal{K}$ gilt $\varphi(x) = 0$ per Definition von $\mathcal{K}$ und $y^*(x) = \varphi(x_0)\|x_0\|^{-2}\langle x_0, x\rangle = 0$ wegen $x_0 \in \mathcal{K}^\perp$. Falls $x = x_0$, so gilt

$$y^*(x_0) = \varphi(x_0)\|x_0\|^{-2} x_0, x_0\rangle = \varphi(x_0) = \varphi(x).$$

Damit ist $\varphi = y^*$ gezeigt.

Es bleibt die Eindeutigkeit zu zeigen. Angenommen $y_1^* = \varphi = y_2^*$. Dann gilt

$$\begin{aligned} \|y_1 - y_2\|^2 &= \langle y_1, y_1 - y_2\rangle - \langle y_2, y_1 - y_2\rangle \\ &= y_1^*(y_1 - y_2) - y_2^*(y_1 - y_2) = \varphi(y_1 - y_2) - \varphi(y_1 - y_2) = 0. \end{aligned}$$

□

Das Riesz'sche Lemma besagt, dass die stetigen Funktional auf einem Hilbertraum ganz einfach aussehen: Sie sind einfach durch Skalarproduktnehmen mit einem Vektor gegeben, $x \mapsto \langle y, x\rangle$. Dies steht in großem Kontrast zu der viel komplizierteren Situation bei Banachräumen, und wird sich in vielen Untersuchungen als sehr vorteilhaft erweisen.

Wir besprechen zwei Anwendungen des Riesz-Lemmas auf Hilbertraumoperatoren. Die erste Anwendung betrifft eine Charakterisierung von Operatoren $A : \mathcal{H} \to \mathcal{H}$ durch ihre Matrixelemente[3] $\langle x, Ay\rangle$.

Lemma 2.19. *Sei $\mathcal{H}$ ein Hilbertraum und $s : \mathcal{H} \to \mathcal{H}$ eine Sesquilinearform. Dann gibt es einen beschränkten Operator $A \in \mathcal{B}(\mathcal{H})$, so dass $\langle Ax, y\rangle = s(x, y)$, genau dann wenn eine Konstante $c > 0$ existiert, so dass*

$$|s(x, y)| \le c\|x\|\|y\|$$

für alle $x, y \in \mathcal{H}$ gilt. In diesem Fall bestimmen sich A und s gegenseitig eindeutig, und es gilt

$$\|A\| = \sup_{v,w\neq 0} \frac{|s(v, w)|}{\|v\|\|w\|} = \sup_{v,w\neq 0} \frac{|\langle v, Aw\rangle|}{\|v\|\|w\|}$$

[3] Für einen linearen Operator $A : \mathcal{H} \to \mathcal{H}$ heißen die Zahlen $\langle x, Ay\rangle$ Matrixelemente von A in Analogie zum endlichdimensionalen Fall: Ist A eine $(n \times n)$-Matrix und $(e_k)_{k=1..n}$ die Standardbasis von $\mathbb{C}^n$, so ist $\langle e_k, Ae_l\rangle = A_{kl}$ der Eintrag von A in Zeile k und Spalte l.

Beweis. Ist $A \in \mathcal{B}(\mathcal{H})$, gegeben, gilt $|\langle Ax, A\rangle| \leq \|Ax\|\|y\| \leq \|A\|\|x\|\|y\|$, dh die Abschätzung gilt mit $c = \|A\| < \infty$ und offenbar ist s eindeutig durch A fixiert.

Ist s gegeben und erfüllt die Abschätzung, so ist bei festem $x \in \mathcal{H}$ die Abbildung $y \mapsto s(x, y)$ ein stetiges lineares Funktional auf $\mathcal{H}$. Nach dem Riesz Lemma gibt es also einen Vektor $\psi_x \in \mathcal{H}$ mit $s(x, y) = \langle\psi_x, y\rangle$. Es ist klar, dass ψ_x linear von x abhängt, und wir definieren $Ax := \psi_x$. Betrachten wir nun die Ungleichung $|s(x, y)| = |\langle Ax, y\rangle| \leq c\|x\|\|y\|$ für $y := Ax$, so erhalten wir $\|Ax\|^2 \leq c\|x\|\|Ax\|$, und somit $\|Ax\| \leq c\|x\|$, dh $\|A\| \leq c < \infty$. Die Eindeutigkeit von A folgt aus der Eindeutigkeitsaussage des Riesz Lemmas.

Für die Normgleichung bemerken wir zunächst folgende Konsequenz der Cauchy-Schwarz Ungleichung: Für $\xi \in \mathcal{H}$ gilt

$$\|\xi\| = \sup_{\eta \neq 0} \frac{|\langle\xi, \eta\rangle|}{\|\eta\|}.$$

Der Beweis von "$\geq$" ist die Cauchy-Schwarz Ungleichung, und der Beweis von "$\leq$" erfolgt durch Einsetzen von $\eta = \xi$.

Damit erhalten wir für die behauptete Normgleichung

$$\sup_{v,w \neq 0} \frac{|\langle v, Aw\rangle|}{\|v\|\|w\|} = \sup_{w \neq 0} \sup_{v \neq 0} \frac{|\langle v, Aw\rangle|}{\|v\|\|w\|} = \sup_{w \neq 0} \frac{\|Aw\|}{\|w\|} = \|A\|.$$

□

Wir bemerken nebenbei, dass als Konsequenz dieses Resultates zwei beschränkte Operatoren A, B genau dann gleich sind, wenn $\langle v, Aw\rangle = \langle v, Bw\rangle$ für alle $v, w \in \mathcal{H}$ gilt. Durch Ausnutzen der Stetigkeit genügt es auch, wenn $\langle v, Aw\rangle = \langle v, Bw\rangle$ für alle v, w in einer in $\mathcal{H}$ dichten Teilmenge gilt, was manchmal einfacher zu zeigen ist.

Beispiel 2.20.

a) Für $a \in \mathbb{R}$ gibt es einen eindeutigen Operator $T_a \in \mathcal{B}(L^2(\mathbb{R}))$ mit

$$\langle\psi, T_a\varphi\rangle = \int_{\mathbb{R}} \overline{\psi(x-a)}\varphi(x+a)dx.$$

Um dies zu zeigen, genügt die Abschätzung

$$\begin{aligned}\left|\int_{\mathbb{R}} \overline{\psi(x-a)}\varphi(x+a)dx\right| &\leq \int_{\mathbb{R}} |\psi(x-a)||\varphi(x+a)|dx \\ &\leq \left(\int_{\mathbb{R}} |\psi(x-a)|^2 dx\right)^{1/2} \left(\int_{\mathbb{R}} |\varphi(y+a)|^2 dy\right)^{1/2} = \|\psi\|\|\varphi\|.\end{aligned}$$

Es gilt $\|T_a\| \leq 1$ (tatsächlich $\|T_a\| = 1$). In diesem Fall lässt sich der Operator aber auch direkt ablesen, nämlich $(T_a f)(x) = f(x + 2a)$.

b) Für $K \in L^2(\mathbb{R}^2)$ gibt es einen eindeutigen Operator $I_K \in \mathcal{B}(L^2(\mathbb{R}))$ mit

$$\langle\psi, I_K\varphi\rangle = \int_{\mathbb{R}} \overline{\psi(x)} \int_{\mathbb{R}} K(x, y)\varphi(y)dy\, dx.$$

In diesem Fall erfolgt der Beweis in den Übungen.

Eine weitere wichtige Anwendung des Riesz Lemmas ist die Konstruktion von adjungierten Operatoren. Dazu erinnern wir uns zunächst an das Konzept der zu einer komplexen $(n \times n)$-Matrix A adjungierten Matrix A^*, die als transponiert Konjugierte definiert ist, $(A^*)_{kl} = \overline{A_{lk}}$. Formuliert durch das Standardskalarprodukt auf $\mathbb{C}^n$ liest sich dies als $\langle v, Aw \rangle = \langle A^* v, w \rangle$. Diese Gleichung erheben wir nun zur Definition des adjungierten Operators im allgemeinen Fall.

Definition und Satz 2.21. Sei $\mathcal{H}$ ein Hilbertraum und $A \in \mathcal{B}(\mathcal{H})$ ein beschränkter Operator. Der *zu A adjungierte Operator* ist der eindeutige Operator $A^* \in \mathcal{B}(\mathcal{H})$, der

$$\langle v, Aw \rangle = \langle A^* v, w \rangle, \qquad v, w \in \mathcal{H},$$

erfüllt.

Beweis. Aufgrund der Beschränktheit von A erfüllt die Sesquilinearform $s(v, w) := \langle v, Aw \rangle$ die Abschätzung $|s(v, w)| \leq \|A\| \|v\| \|w\|$. Also existiert A^* eindeutig mit den gewünschten Eigenschaften nach Lemma 2.19. □

Die *Adjunktion* ist die Abbildung $\mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$, $A \mapsto A^*$. Das folgende Lemma fasst ihre wichtigsten Eigenschaften zusammen.

Lemma 2.22. *Sei $\mathcal{H}$ ein Hilbertraum und $A, B \in \mathcal{B}(\mathcal{H})$, $\lambda \in \mathbb{K}$. Dann gilt*

a) $(A^*)^* = A$ *und* $(A + \lambda B)^* = A^* + \overline{\lambda} B^*$, *dh die Adjunktion ist eine antilineare Involution.*

b) $(AB)^* = B^* A^*$.

c) $\|A^*\| = \|A\|$ *und* $\|A^* A\| = \|A\|^2$ *(C^*-Eigenschaft).*

Beweis. a) Für $v, w \in \mathcal{H}$ gilt

$$\langle (A^*)^* v, w \rangle = \langle v, A^* w \rangle = \overline{\langle A^* w, v \rangle} = \overline{\langle w, Av \rangle} = \langle Av, w \rangle,$$

und damit $(A^*)^* = A$. Weiterhin gilt

$$\langle (A + \lambda B)^* v, w \rangle = \langle v, (A + \lambda B) w \rangle = \langle v, Aw \rangle + \lambda \langle v, Bw \rangle = \langle (A^* + \overline{\lambda} B^*) v, w \rangle,$$

und damit $(A + \lambda B)^* = A^* + \overline{\lambda} B^*$.

b) Für $v, w \in \mathcal{H}$ gilt

$$\langle (AB)^* v, w \rangle = \langle v, ABw \rangle = \langle A^* v, Bw \rangle = \langle B^* A^* v, w \rangle,$$

und damit $(AB)^* = B^* A^*$.

c) Wir haben

$$\|A^*\| = \sup_{v, w \neq 0} \frac{|\langle v, A^* w \rangle|}{\|v\| \|w\|} = \sup_{v, w \neq 0} \frac{|\langle w, Av \rangle|}{\|v\| \|w\|} = \|A\|,$$

und für die C^*-Eigenschaft

$$\begin{aligned}\|A^*A\| &= \sup_{\|v\|=\|w\|=1} |\langle v, A^*Aw\rangle| \\ &= \sup_{\|v\|=\|w\|=1} |\langle Av, Aw\rangle| \\ &\geq \sup_{\|v\|=1} |\langle Av, Av\rangle| \\ &= \|A\|^2.\end{aligned}$$

Die umgekehrte Abschätzung $\|A^*A\| \leq \|A^*\|\|A\| = \|A\|^2$ folgt sofort aus $\|AB\| \leq \|A\|\|B\|$ und $\|A^*\| = \|A\|$. $\square$

Die Bedeutung der C^*-Eigenschaft ist an dieser Stelle noch nicht offensichtlich, wird aber später klar werden. Der Name "C^*-Eigenschaft" rührt daher, dass es eine Klasse von normierten Algebren namens C^*-Algebren gibt, für die diese Eigenschaft axiomatisch gefordert wird. $\mathcal{B}(\mathcal{H})$ ist eine C^*-Algebra – mehr dazu in Funktionalanalysis 2!

Wir illustrieren die Wirkung des adjungierten Operators in folgendem einfachen aber wichtigen Lemma.

Lemma 2.23. *Sei $\mathcal{H}$ ein Hilbertraum und $A \in \mathcal{B}(\mathcal{H})$. Dann gilt*

$$\ker(A^*) = \operatorname{Ran}(A)^\perp.$$

Insbesondere hat A genau dann dichtes Bild wenn A^* injektiv ist.

Beweis. "$\supset$": Sei $v \in (\operatorname{Ran} A)^\perp$, dh für alle $w \in \mathcal{H}$ gilt $0 = \langle v, Aw\rangle = \langle A^*v, w\rangle$. Also gilt $v \in \ker(A^*)$ (setze $w = A^*v$).

"$\subset$" Sei $A^*v = 0$. Dann gilt $0 = \langle A^*v, w\rangle = \langle v, Aw\rangle$, also $v \perp \operatorname{Ran} A$.

Durch Bilden von orthogonalen Komplementen haben wir

$$\ker(A^*)^\perp = \operatorname{Ran}(A)^{\perp\perp} = \overline{\operatorname{Ran} A}.$$

Injektivität von A^* ist äquivalent zu $\ker(A^*) = \{0\} \Leftrightarrow (\ker A^*)^\perp = \mathcal{H}$, und damit äquivalent zur Dichtheit von $\operatorname{Ran} A$. $\square$

Da es schwierig sein kann, über Dichtheit von Bildern zu entscheiden, ist die letzte Aussage dieses Lemma oft hilfreich. Wir illustrieren dies an einem einfachen Beispiel.

Beispiel 2.24.

a) Sei $g \in L^\infty(\mathbb{R})$ eine wesentlich beschränkte Funktion und $M_g : f \mapsto gf$ der zugehörige beschränkte Multiplikationsoperator auf $L^2(\mathbb{R})$. Wir berechnen ${M_g}^*$ wie folgt: Seien $f, h \in L^2(\mathbb{R})$. Dann gilt

$$\begin{aligned}\langle f, M_g h\rangle = \langle f, gh\rangle &= \int_{\mathbb{R}} \overline{f(x)}g(x)h(x)dx = \int_{\mathbb{R}} \overline{\overline{g(x)}f(x)}h(x)dx \\ &= \langle \bar{g}f, h\rangle = \langle M_{\bar{g}}f, h\rangle,\end{aligned}$$

wovon wir ${M_g}^* = M_{\bar{g}}$ ablesen.

Wann hat M_g dichtes Bild? Dazu müssen wir den Kern $\ker(M_{\bar{g}})$ bestimmen. Eine Funktion $f \in L^2(\mathbb{R})$ liegt in $\ker(M_{\bar{g}})$ genau dann, wenn $\overline{g(x)}f(x) = 0$ für fast alle $x \in \mathbb{R}$. Falls g nur auf einer Nullmenge verschwindet, so folgt $f(x) = 0$ fast überall, dh $f = 0$ in $L^2(\mathbb{R})$. Falls es eine Menge $S \subset \mathbb{R}$ von Maß $\mu(S) > 0$ gibt, so dass $g(x) = 0$ für alle $x \in S$, so liegt $f := \chi_S \neq 0$ in $\ker(M_{\bar{g}})$.

Zusammengefasst M_g hat genau dann dichtes Bild, wenn die Nullstellenmenge $\{x \in \mathbb{R} : \ g(x) = 0\}$ von g eine Nullmenge ist.

b) Als ein weiteres Beispiel für adjungierte Operatoren betrachten wir für $v, w \in \mathcal{H}$ den Operator

$$|v\rangle\langle w| : \mathcal{H} \to \mathcal{H}, \qquad x \mapsto \langle w, x\rangle v$$

und behaupten $(|v\rangle\langle w|)^* = |w\rangle\langle v|$. In der Tat: Für $x, y \in \mathcal{H}$ gilt

$$\langle x, |v\rangle\langle w|y\rangle = \langle x, \langle w, y\rangle v\rangle = \langle w, y\rangle\langle x, v\rangle = \langle\langle v, x\rangle w, y\rangle = \langle |w\rangle\langle v|x, y\rangle.$$

Hier sehen wir $\ker(|v\rangle\langle w|)^* = \ker |w\rangle\langle v| = \{v\}^\perp = (\operatorname{Ran} |v\rangle\langle w|)^\perp$ besonders einfach.

Weitere Beispiele von adjungierten Operatoren erfolgen in den Übungen.

Die Adjunktion definiert verschiedene wichtige Klassen von Operatoren:

Definition 2.25. Sei $\mathcal{H}$ ein Hilbertraum und $A \in \mathcal{B}(\mathcal{H})$.

a) A heißt *selbstadjungiert* falls $A = A^*$.

b) A heißt *unitär* falls A invertierbar ist und $A^* = A^{-1}$ gilt. Dies ist äquivalent zu $AA^* = A^*A = 1$.

c) A heißt *orthogonale Projektion* falls $A^2 = A$ und $A = A^*$ gilt.

d) A heißt *normal*, falls $AA^* = A^*A$ gilt.

Der Name Selbstadjungiertheit ist selbsterklärend. Die Identität ist selbstadjungiert. Selbstadjungierte Operatoren spielen in der Spektraltheorie und den Anwendungen eine zentrale Rolle.

Die Bezeichnung unitär muss gerechtfertigt werden, da wir sie bereits in Definition 2.11 vergeben haben. Spezialisieren wir Definition 2.11 auf einen einzelnen Hilbertraum, so besagt sie $\langle Av, Aw\rangle = \langle v, w\rangle$ und A bijektiv. Die Skalarproduktgleichung ist äquivalent zu $A^*A = 1$ (solche Operatoren sind genau die Isometrien). Ist eine Isometrie surjektiv (also bijektiv), so ist ihr Inverses auch eine Isometrie, also gilt auch $AA^* = 1$. Das zeigt, dass die neue Definition mit der alten übereinstimmt.

Die Bezeichnung orthogonale Projektion muss auch gerechtfertigt werden, da wir sie bereits im Zusammenhang von Projektionen auf abgeschlossene Unterräume fixiert haben – dort wurde ein Operator A orthogonale Projektion genannt, falls $A^2 = A$ und $\operatorname{Ran} A = (\ker A)^\perp$.

Dann können wir $\mathcal{H}$ gemäß $\mathcal{H} \cong \operatorname{Ran} A \oplus \ker A$ zerlegen und haben im Sinne dieser Zerlegung $A = 1 \oplus 0 = 1^* \oplus 0^* = A^*$, dh A ist selbstadjungiert im Sinne von Definition 2.25. Gilt hingegen $A^2 = A$ und $A = A^*$, so folgt mit Lemma 2.23 auch $\ker A = \ker A^* = (\operatorname{Ran} A)^\perp$. Also sind beide Definitionen äquivalent.

Der Begriff eines normalen Operators enthält insbesondere die selbstadjungierten und unitären Operatoren. Wir werden normale Operatoren in der Spektraltheorie eingehender betrachten.

Beispiel 2.26. Die Operatoren $A = |v\rangle\langle w|$ aus Beispiel 2.24 b) erfüllen $\|A\| = \|v\|\|w\|$ und $A^* = |w\rangle\langle v|$ – sie sind also genau dann orthogonale Projektionen, wenn $v = w$ Norm 1 hat. Dann ist $|v\rangle\langle v|$ eine orthogonale Projektion mit Rang ($= \dim \operatorname{Ran} A$) 1.

2.3 Orthonormalbasen

Der aus der linearen Algebra bekannte Begriff einer Orthonormalbasis hat sich dort oft als nützlich erwiesen. Zuerst erinnern wir uns kurz, dass jede orthonormale Teilmenge $B \subset \mathcal{H}$ eines Prä-Hilbertraums linear unabhängig ist: Für $b_1, \ldots, b_n \in B$ und $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ mit $\lambda_1 b_1 + \ldots \lambda_n b_n = 0$ folgt durch Anwendung von b_k^* sofort $\lambda_k = 0$ für jedes $k \in \{1, \ldots, n\}$. Liegt ein Vektor $v \in \mathcal{H}$ ausgedrückt als Linearkombination von Vektoren in B vor, dh $v = \lambda_1 b_1 + \ldots \lambda_n b_n$, so ergibt Anwendung von b_k^* sofort die Koeffizienten $\lambda_k = \langle b_k, v\rangle$.

Aufgrund dieser und weiterer nützlicher Eigenschaften wollen wir nun den Begriff einer Orthonormalbasis (ONB) in einem allgemeinen Hilbertraum entwickeln. In der Linearen Algebra definieren wir eine ONB eines endlichdimensionalen Hilbertraums als ein Orthonormalsystem (=orthonormale Menge) B, die ein Erzeugendensystem ist, dh $\mathcal{H} = \operatorname{span} B$. Im unendlichdimensionalen Fall wollen wir schwächer $\mathcal{H} = \overline{\operatorname{span} B}$ fordern, dh der Raum aller Linearkombinationen von Vektoren aus B muss nur ein dichter Unterraum, aber nicht der ganze Raum $\mathcal{H}$ sein. Dies wird im Endeffekt auf "unendliche Linearkombinationen" der Art $v = \sum_{b \in B} \langle b, v\rangle b$ hinauslaufen (die Details der Konvergenz dieser Reihen wird weiter unten geklärt). Wir betonen gleich zu Beginn dieses Abschnittes, dass eine so definierte Orthonormalbasis im unendlichdimensionalen Fall *keine* (Hamel-)Basis ist – nicht jeder Vektor $v \in \mathcal{H}$ wird eine (per Definition immer endliche) Linearkombination von Basisvektoren $b \in B$ sein. Nichtdestotrotz wird sich der neue ONB Begriff als viel praktischer als der Begriff einer Hamelbasis erweisen.

Definition 2.27. Eine *Orthonormalbasis* eines Hilbertraums ist ein Orthonormalsystem (=orthonormale Menge) $B \subset \mathcal{H}$ mit $\overline{\operatorname{span} B} = \mathcal{H}$.

Der folgende Satz gibt verschiedene äquivalente Formulierungen dieses Begriffs und zeigt insbesondere, dass jeder Vektor in $\mathcal{H}$ in B entwickelt werden kann.

Satz 2.28. *Sei $\mathcal{H}$ ein Hilbertraum und $B \subset \mathcal{H}$ ein Orthonormalsystem. Dann sind äquivalent:*

a) B ist eine Orthonormalbasis.

b) B ist ein maximales Orthonormalsystem: Das einzige Orthonormalsystem $B' \subset \mathcal{H}$, das $B \subset B'$ erfüllt, ist $B' = B$.

c) Erfüllt ein Vektor $v \in \mathcal{H}$ die Gleichung $\langle b, v\rangle = 0$ für alle $b \in B$, so folgt $v = 0$.

d) Es gilt $v = \sum_{b\in B}\langle b, v\rangle b$, wobei die Summe unabhängig von der Anordnung der Summanden in der Norm von $\mathcal{H}$ konvergiert (unbedingte Konvergenz). Höchstens abzählbar viele der Koeffizienten $\langle b, v\rangle$ sind von Null verschieden.

e) Für alle $v, w \in \mathcal{H}$ gilt $\langle v, w\rangle = \sum_{b\in B}\langle v, b\rangle\langle b, w\rangle$.

f) Für alle $v \in \mathcal{H}$ gilt $\|v\|^2 = \sum_{b\in B} |\langle v, b\rangle|^2$ (Parseval-Gleichung).

Beweis. a)⇒b) Sei $B' \supset B$ ein Orthonormalsystem und $\psi \in B'\backslash B$. Dann gilt $\psi \in (\operatorname{span}B)^\perp = (\overline{\operatorname{span}B})^\perp = \mathcal{H}^\perp = \{0\}$, also $\psi = 0$ im Widerspruch zur Orthonormalität von B'.

b)⇒c) Angenommen, v erfüllt die Bedingung und $v \neq 0$. Dann ist $B' = B \cup \{\frac{v}{\|v\|}\}$ ein B echt enthaltendes Orthonormalsystem, im Widerspruch zu b).

c)⇒d) Wir verwenden die Besselsche Ungleichung: Für jede endliche Menge $\{b_1, \ldots, b_N\} \subset B$ gilt $\|v\|^2 \geq \sum_{k=1}^N |\langle b_k, v\rangle|^2$. Also ist $B_{v,n} := \{b \in B : |\langle b, v\rangle| > \frac{1}{n}\}$ für jedes $n \in \mathbb{N}$ endlich. Da nach c) für $v \neq 0$ die Gleichung $\langle b, v\rangle = 0$ nicht für alle $b \in B$ gilt, ist $B_v := \{b \in B : \langle b, v\rangle \neq 0\} = \bigcup_{n\in\mathbb{N}} B_{v,n}$ als abzählbare Vereinigung endlicher Mengen abzählbar.

Sei nun $(b_n)_{n\in\mathbb{N}}$ eine beliebige Abzählung von B_v (der Fall $|B_v| < \infty$ ist trivial). Dann ist $\sum_{k=1}^N |\langle b_k, v\rangle|^2$ in N monoton wachsend und nach der Besselschen Ungleichung durch $\|v\|^2$ beschränkt, also konvergent. Wir definieren $v_N := \sum_{k=1}^N\langle b_k, v\rangle b_k$. Dies ist eine Cauchy-Folge: Für $N > M$ gilt

$$\|v_N - v_M\|^2 = \left|\sum_{k=M+1}^{N} \langle b_k, v\rangle b_k\right|^2 = \sum_{k=M+1}^{N} |\langle b_k, v\rangle|^2 \to 0$$

für $N, M \to \infty$, da $\sum_{k=1}^\infty |\langle b_k, v\rangle|^2 < \infty$. Wir behaupten, dass der Grenzwert w dieser Folge v ist. Um dies zu zeigen, genügt es nach c) zu zeigen, dass $v - w$ auf allen $b \in B$ senkrecht steht. Dies ist $b \notin B_v$ offensichtlich. Für $b = b_n \in B_v$ gilt aufgrund der Stetigkeit des Skalarprodukts

$$\langle v - w, b_n\rangle = \lim_{N\to\infty} \left\langle v - \sum_{k=1}^{N}\langle b_k, v\rangle b_k, b_n \right\rangle = \lim_{N\to\infty} \left(\langle v, b_n\rangle - \overline{\langle b_n, v\rangle}\right) = 0.$$

d)⇒e) Wir setzen die Basisentwicklungen aus d) ein:

$$\langle v, w\rangle = \left\langle v, \sum_{b\in B}\langle b, w\rangle b \right\rangle = \sum_{b\in B} \langle v, b\rangle \langle b, w\rangle.$$

e)⇒f) folgt direkt durch Spezialisieren auf $w = v$.

f)⇒a) Nach Satz 2.4 gilt für jede endliche Teilmenge $B_N \subset B$

$$\|v\|^2 = \sum_{b\in B_N} |\langle b, v\rangle|^2 + \left\| v - \sum_{b\in B_N} \langle b, v\rangle b \right\|^2.$$

Also existiert die Reihe $\sum_{b\in B} |\langle b, v\rangle|^2$ wie im Beweisteil c)⇒d). Wegen $\|v\|^2 = \sum_{b\in B} |\langle v, b\rangle|^2$ gilt $v = \sum_{b\in B}\langle b, v\rangle b$ im Sinne der Konvergenz von d). Da $v \in \mathcal{H}$ beliebig war, folgt $\overline{\text{span} B} = \mathcal{H}$. □

Eine ONB B ist also eine "Basis" (zwar nicht im Sinne der Definition einer Hamelbasis, aber wir werden die Elemente einer ONB trotzdem als "Basisvektoren" bezeichnen), dass jeder Vektor $v \in \mathcal{H}$ als "unendliche Linearkombination" geschrieben werden kann, und zwar auf eindeutige Art und Weise. Dabei gilt stets $v = \sum_{b\in B}\langle b, v\rangle b$ (unbedingte Konvergenz) mit nur abzählbar vielen von Null verschiedenen Koeffizienten $\langle b, v\rangle$, und $\|v\|^2 = \sum_{b\in B} |\langle b, v\rangle|^2$. Zwei Vektoren x, y stimmen genau dann überein, wenn $\langle b, x\rangle = \langle b, y\rangle$ für alle $b \in B$ gilt.

Betrachten wir nur die (endlichen, also eigentlichen) Linearkombinationen von Vektoren einer ONB, so erhalten wir immerhin einen dichten Unterraum $\text{span} B \subset \mathcal{H}$.

Beispiel 2.29.

a) Sei $\mathcal{H} = \ell^2_{\mathbb{N}}$ und $b_n \in \ell^2_{\mathbb{N}}$ die Folge $(b_n)_k = \delta_{n,k}$, $k, n \in \mathbb{N}$. Dann ist $B := \{b_n : n \in \mathbb{N}\}$ eine orthonormale Menge bzgl des Skalarproduktes von $\ell^2_{\mathbb{N}}$. Wir behaupten, dass B sogar eine Orthonormalbasis ist und verwenden dafür Charakterisierung c) aus dem vorhergehenden Satz.

   Sei also $x \in \ell^2_{\mathbb{N}}$ mit $0 = \langle b_n, x\rangle = \sum_{k=1}^{\infty}(b_n)_k x_k = x_n$ für alle n. Dann gilt $x = 0$. Also ist B eine ONB.

b) Sei $\mathcal{H} = L^2([0, 2\pi])$ und $f_n(x) := (2\pi)^{-1/2}e^{inx}$, $n \in \mathbb{Z}$. Wir behaupten, dass $B := \{f_n : n \in \mathbb{Z}\}$ eine ONB ist. Dazu prüfen wir zuerst, dass B ein Orthonormalsystem ist: Für $n \neq m$ gilt

$$\langle f_n, f_m\rangle = \frac{1}{2\pi}\int_0^{2\pi} e^{i(m-n)x}dx = \frac{1}{2\pi}\frac{1}{i(m-n)}\left[e^{i(m-n)x}\right]_0^{2\pi} = 0,$$

   und $\|f_n\|^2 = \frac{1}{2\pi}\int_0^{2\pi} 1dx = 1$.

   Um die Vollständigkeit von B zu zeigen, verwenden wir den zweiten Weierstraßschen Approximationssatz (siehe Buch von Werner Kor. IV.2.12), der besagt, dass der Raum $\text{span} B$ der trigonometrischen Polynome ein dichter Unterraum des Banachraums der stetigen periodischen Funktion $C_{\text{per}}([0, 2\pi]) = \{f \in C([0, 2\pi]) : f(0) = f(2\pi)\}$ bzgl. Supremumsnorm ist (dies ist ein Ergebnis aus der Theorie der Fourierreihen). Weiterhin ist $C_{\text{per}}([0, 2\pi]) \subset L^2([0, 2\pi])$ ein dichter Unterraum bzgl L^2-Norm: Dies folgt daraus, dass eine L^2-Funktion beliebig gut in L^2-Norm durch stetige Funktionen angenähert werden kann, und die Periodizität kann auch durch einen beliebig kleinen Fehler in L^2-Norm erreicht werden (Details: Übung).

   Da der Abschluss von $\text{span} B$ in L^2-Norm den Abschluss von $\text{span} B$ in Supremumsnorm enthält (eine Folge von Funktionen $g_n : [0, 2\pi] \to \mathbb{C}$, die gleichmäßig konvergiert, konvergiert erst recht in L^2-Norm), folgt die Behauptung $\overline{\text{span} B} = \mathcal{H}$.

Nachdem wir einige Beispiele von Orthonormalbasen gesehen haben, ist eine natürliche Frage, wie man sich Orthonormalbasen verschafft bzw ob sie überhaupt immer existieren.

Bzgl der ersten Frage erinnern wir an das *Gram-Schmidtsche Orthonormalisierungsverfahren* aus der Linearen Algebra:

Lemma 2.30 (Gram-Schmidt-Verfahren). *Sei $\mathcal{H}$ ein Hilbertraum und $A := \{v_n : n \in \mathbb{N}\} \neq \{0\} \subset \mathcal{H}$ eine abzählbare Teilmenge. Dann existiert ein Orthonormalsystem $S \subset \mathcal{H}$ mit* $\operatorname{span} S = \operatorname{span} A$.

Beweis. Als ersten Basisvektor definieren wir $b_1 := v_1/\|v_1\|$. Der zweite Basisvektor soll aus v_2 durch Subtraktion des zu b_1 parallelen Anteils und Normierung entstehen, $b_2' := v_2 - \langle b_1, v_2\rangle b_1$ und $b_2 := b_2'/\|b_2'\|$. Dies funktioniert genau dann, wenn $b_2' \neq 0$. Falls $b_2' = 0$, so ist v_2 linear abhängig von b_1. In diesem Fall überspringen wir v_2 und machen mit v_3 weiter, etc. Nach eventuellen Überspringungen gilt dann $b_2 \perp b_1$ und $\|b_2\| = 1$. Induktiv definieren wir nun

$$b_{n+1}' := v_{n+1} - \sum_{k=1}^{n} \langle b_k, v_{n+1}\rangle b_k, \qquad b_{n+1} := b_{n+1}'/\|b_{n+1}'\|,$$

falls $b_{n+1}' \neq 0$; anderenfalls lassen wir v_{n+1} aus.

Dann ist $S := \{b_n : n \in \mathbb{N}\}$ bzw $S := \{b_n : 1 \leq n \leq N\}$ ein Orthonormalsystem mit $\operatorname{span} S \subset \operatorname{span} A$. Andererseits sieht man auch leicht per Induktion, dass $v_n \in \operatorname{span} S$ für alle nicht ausgelassenen v_n gilt. Die ausgelassenen v_n sind aber linear abhängig von den nicht ausgelassenen, so dass $\operatorname{span} S = \operatorname{span} A$ gilt. □

Das Gram-Schmidt Verfahren produziert also ausgehend von einer abzählbaren Menge $A \subset \mathcal{H}$ eine orthonormale Menge S, die genau dann eine ONB ist, wenn $\operatorname{span} A$ ein dichter Unterraum von $\mathcal{H}$ ist.

Diese Beobachtung bringt uns in die Nähe der Topologie-Vorlesung. Wir definieren einen topologischen (oder auch nur: metrischen) Raum als *separabel*, wenn er eine abzählbare dichte Teilmenge besitzt[4].

Satz 2.31. *Ein Hilbertraum $\mathcal{H}$ ist genau dann separabel, wenn er eine abzählbare ONB B besitzt. Es gilt*

$$\mathcal{H} \cong \begin{cases} \mathbb{C}^{|B|} & |B| < \infty \\ \ell^2_{\mathbb{N}} & B \text{ abzählbar unendlich} \end{cases}.$$

Beweis. Angenommen, $\mathcal{H}$ ist separabel und $A \subset \mathcal{H}$ eine abzählbare dichte Teilmenge. Dann liefert das Gram-Schmidt Verfahren angewendet auf A eine abzählbare ONB.

Angenommen, $\mathcal{H}$ hat eine abzählbare Basis $B = \{b_n : n \in \mathbb{N}\}$. Wir definieren A als die Menge aller (endlichen) Linearkombinationen $\sum_n r_n \cdot b_n$ der b_n mit Koeffizienten $r_n \in \mathbb{Q} + i\mathbb{Q}$ mit rationalen Real- und Imaginärteilen (bzw. $r_n \in \mathbb{Q}$ im Falle eines reellen Hilbertraums). Dann ist A abzählbar. Um zu zeigen, dass A dicht ist, genügt es zu zeigen, dass jede endliche Linearkombination von Basisvektoren durch Vektoren in A approximiert werden kann. Sei

[4] Ein topologischer Raum, der dem 2. Abzählbarkeitsaxiom genügt, ist separabel. Ein metrischer Raum ist genau dann separabel, wenn er dem 2. Abzählbarkeitsaxiom genügt.

also $x = \sum_{n=1}^{N} c_n \cdot b_n$ und wähle Koeffizienten $r_n \in \mathbb{Q} + i\mathbb{Q}$ mit $|r_n - c_n| < \varepsilon$. Dann gilt mit $a := \sum_{n=1}^{N} r_n b_n$ die Abschätzung $\|x - a\|^2 = \sum_{n=1}^{N} |r_n - c_n|^2 \leq N\varepsilon^2$. Also ist $A \subset \mathcal{H}$ dicht.

Wir zeigen die Isomorphieaussage für den Fall einer abzählbar unendlichen Basis $B = \{b_n : n \in \mathbb{N}\}$, der endliche Fall ist analog und einfacher. Dazu definieren wir einen Operator

$$U : \mathcal{H} \to \ell^2_{\mathbb{N}}, \qquad Ux := (\langle b_n, x \rangle)_{n \in \mathbb{N}}$$

Nach Satz 2.28 ist U wohldefiniert, surjektiv, und isometrisch, also unitär. □

Die allermeisten praktisch auftretenden Hilberträume sind separabel und damit im unendlichdimensionalen Fall isometrisch isomorph. Es gibt allerdings auch nicht separable Hilberträume, wie man sich leicht an Beispielen überlegt: Sei S eine beliebige Menge und $\ell^2_S = \{f : S \to \mathbb{C} : \sum_{s \in S} |f(s)|^2 < \infty\}$ (mit dem unbedingten Konvergenzbegriff aus Satz 2.28). Dann bilden die Funktionen $f_t : s \mapsto \delta_{s,t}$ eine ONB mit der gleichen Kardinalität wie S.

Wir zeigen noch, dass auch nicht separable Hilberträume ONBs besitzen.

Satz 2.32. *Jeder Hilbertraum besitzt eine Orthonormalbasis.*

Beweis. Wir betrachten die Menge M aller orthonormaler Teilmengen unseres Hilbertraums $\mathcal{H}$. Die Menge M ist partiell durch Inklusion geordnet, dh $S \leq S' :\Leftrightarrow S \subset S'$ ist eine reflexive, transitive, antisymmetrische Relation (=partielle Ordnung) auf M. Dies unterscheidet sich von einer totalen Ordnung, für die zusätzlich gelten würde, dass für zwei Elemente S, S' stets $S \leq S'$ oder $S' \leq S$ gilt (zB die Ordnung auf $\mathbb{R}$). Eine *Kette* in einem partiell geordneten Raum ist eine total geordnete Teilmenge. Betrachten wir eine Kette $K \subset M$ und dazu $o(K) := \bigcup_{S \in K} k$; dann ist $o(K)$ eine obere Schranke von K in dem Sinne, dass $S \leq o(K)$ für alle $S \in K$. Weiterhin ist $o(K) \in M$ eine orthonormale Menge: Denn für $v, v' \in o(K)$ existieren $S, S' \in K$ mit $v \in S$, $v' \in S'$. Da K eine Kette ist, liegen beide Vektoren v und v' in S oder S', haben also Norm 1 und stehen senkrecht aufeinander falls $v \neq v'$.

Nun können wir das *Lemma von Zorn* anwenden, das besagt, dass jede nichtleere partiell geordnete Menge, in der jede Kette eine obere Schranke hat, ein maximales Element (bzgl der partiellen Ordnung) besitzt.

In unserem Hilbertraumkontext erhalten wir also eine orthonormale Menge die maximal ist in dem Sinne, dass sie in keiner anderen orthonormalen Menge außer sich selbst enthalten ist. Nach Satz 2.28 b) ist dies eine Orthonormalbasis von $\mathcal{H}$. □

Dieser Beweis war sehr abstrakt und beruht mit dem Lemma von Zorn auf dem Auswahlaxiom. Eine Konstruktionsanleitung für Orthonormalbasen erhalten wir so sicherlich nicht, nur eine Existenzaussage.

In Hilbertraumsituationen werden uns Orthonormalbasen noch häufig begegnen. Wir führen hier noch zwei Bemerkungen an, die ONBs benutzen.

Beispiel 2.33. *Die Einheitskugel* $\mathcal{H}_1 := \{v \in \mathcal{H} : \|v\| \leq 1\}$ *ist abgeschlossen und beschränkt, aber für* $\dim \mathcal{H} = \infty$ *nicht kompakt.*
Beweis: Die Beschränktheit ist offensichtlich und die Abgeschlossenheit nicht schwer zu

sehen (finde eine offene Kugel U um $x \notin \mathcal{H}_1$ mit $U \cap \mathcal{H}_1 = \emptyset$). Um zu zeigen, dass $\mathcal{H}_1$ für $\dim \mathcal{H} = \infty$ nicht kompakt ist, müssen wir zeigen, dass $\mathcal{H}_1$ eine Folge ohne konvergente Teilfolge enthält. Wir betrachten dazu ein abzählbares Orthonormalsystem $(b_n)_{n\in\mathbb{N}} \subset \mathcal{H}_1$. Dies existiert als ONB im separablen Fall und als Teilmenge einer ONB im nicht separablen Fall. Aufgrund der Orthonormalität gilt $\|b_n - b_m\| = \sqrt{2}$ für alle $n \neq m$. Also ist keine Teilfolge von (b_n) Cauchy. □

Das zweite Beispiel besagt, dass im Gegensatz zum endlichdimensionalen Fall im unendlichdimensionalen Fall lineare Abbildungen nicht so einfach durch (unendliche) Matrizen beschrieben werden können.

Beispiel 2.34. Gegeben einen unendlichdimensionalen separablen Hilbertraum $\mathcal{H}$ mit ONB $(b_n)_{n\in\mathbb{N}}$, so definiert jedes $A \in \mathcal{B}(\mathcal{H})$ eine "unendliche Matrix" $A_{nm} := \langle b_n, Ab_m\rangle$, $n, m \in \mathbb{N}$. Aber nicht jede solche "Matrix" definiert einen beschränkten Operator! Sei zB $A_{nm} = 1$ für alle $n, m \in \mathbb{N}$. Würden diese Koeffizienten von einem Operator $A \in \mathcal{B}(\mathcal{H})$ kommen, so hätten wir $Ab_m = \sum_n \langle b_n, Ab_m\rangle b_n = \sum_n b_n$, was wegen $\|\sum_n b_n\|^2 = \sum_n 1 = \infty$ kein Vektor in $\mathcal{H}$ ist.

Damit eine unendliche Matrix $(A_{nm})_{n,m\in\mathbb{N}}$ einen beschränkten Operator A darstellt, ist es *notwendig*, dass $\sum_n |A_{nm}|^2 < \infty$ für alle m und $\sum_m |A_{nm}|^2 < \infty$ für alle n gilt. Denn wir haben für $A \in \mathcal{B}(\mathcal{H})$

$$\begin{aligned}
\|Ab_m\|^2 &= \sum_n |\langle b_n, Ab_m\rangle|^2 = \sum_n |A_{nm}|^2 < \infty, \\
\|A^*b_n\|^2 &= \sum_m |\langle b_m, A^*b_n\rangle|^2 = \sum_m |\langle Ab_m, b_n\rangle|^2 = \sum_m |A_{nm}|^2 < \infty.
\end{aligned}$$

Weiterhin ist $\sum_{n,m} |A_{nm}|^2 < \infty$ *hinreichend* dafür, dass (A_{nm}) einen beschränkten Operator A darstellt: Definiere eine Sesquilinarform durch

$$s(x, y) := \sum_{n,m} \overline{\langle b_n, x\rangle} \langle b_m, y\rangle A_{nm}.$$

Dann gilt per Cauchy-Schwarz in $\ell^2_{\mathbb{N}^2}$

$$|s(x, y)|^2 \leq \sum_{n,m} |\langle b_n, x\rangle|^2 |\langle b_m, y\rangle|^2 \cdot \sum_{n,m} |A_{nm}|^2 = \|x\|^2\|y\|^2 \cdot \sum_{n,m} |A_{nm}|^2.$$

Also ist s wohldefiniert und es gibt nach Lemma 2.19 einen eindeutigen beschränkten Operator $A \in \mathcal{B}(\mathcal{H})$ mit $\|A\|^2 \leq \sum_{n,m} |A_{nm}|^2$, so dass $\langle x, Ay\rangle = s(x, y)$ und insbesondere $\langle b_n, Ab_m\rangle = A_{nm}$ gilt. Diese Operatoren heißen *Hilbert-Schmidt Operatoren*.

Sie bilden eine spezielle Klasse innerhalb der Menge aller beschränkter Operatoren. Zum Beispiel ist die Identität ein beschränkter Operator, aber ihre Matrixelement $1_{nm} = \delta_{nm}$ erfüllen *nicht* $\sum_{n,m} |1_{nm}|^2 < \infty$, dh die Identität (auf einem unendlichdimensionalen Hilbertraum) ist nicht Hilbert-Schmidt.

Hilbertraummethoden finden in vielen Gebieten Anwendungen, wie wir noch später in diesem Kurs sehen werden. Hier einige Beispiele (siehe auch die Wikipedia-Seite)

a) **Fourieranalysis.** Die Theorie der Fourierreihen (siehe Beispiel 0.1 und Beispiel 2.29 b)) und der Fouriertransformation (siehe Analysis 3, Kapitel 3) sind wichtige Anwendungsfelder von Hilbertraummethoden. Gegeben eine periodische Funktion $f : [0, 2\pi] \to \mathbb{C}$, ist ihre *Fourierreihe* durch

$$x \mapsto (2\pi)^{-1/2} \sum_{n \in \mathbb{Z}} f_n \cdot e^{2\pi i n x}$$

mit den *Fourierkoeffizienten* $f_n := (2\pi)^{-1/2} \int_0^{2\pi} e^{-2\pi i n x} f(x) dx$ definiert, was einer Zerlegung von f in harmonische Schwingungen entspricht (mit Anwendungen zB in der Signalverarbeitung und Datenkompression (MP3, JPEG)).

Es ist eine schwierige und nicht im Detail geklärte Frage, für genau welche f die Fourierreihe gleichmäßig gegen f konvergiert (hinreichende Bedingungen sind aber leicht anzugeben). Aus der Perspektive des Hilbertraums $L^2([0, 2\pi])$ ist hingegen klar, dass für jedes $f \in L^2([0, 2\pi])$ die Fourierreihe gegen f in der L^2-Norm konvergiert, da es sich um eine Entwicklung in eine ONB handelt.

Das kontinuierliche Analogon der Fourierreihen ist die Fouriertransformation, die zunächst auf L^1-Funktionen $f \in L^1(\mathbb{R})$ durch

$$(\mathcal{F} f)(k) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-ikx} f(x) dx$$

definiert ist. Es zeigt sich, dass $\mathcal{F}$ auf dem Schwartzraum $\mathscr{S}(\mathbb{R})$ eine lineare Bijektion ist, die die Plancherelgleichung $\langle (\mathcal{F} f), (\mathcal{F} g) \rangle = \langle f, g \rangle$ (mit dem L^2-Skalarprodukt $\langle \cdot \, , \, \cdot \rangle$) erfüllt (siehe Analyis 3).

Auch hier ist eine Hilbertraumperspektive hilfreich: Da $\mathscr{S}(\mathbb{R})$ ein dichter Unterraum des Hilbertraums $L^2(\mathbb{R})$ ist und $\mathcal{F}$ aufgrund der Plancherelgleichung in L^2-Norm beschränkt ist, können wir $\mathcal{F}$ zu einem beschränkten Operator $L^2(\mathbb{R}) \to L^2(\mathbb{R})$ ausdehnen. Wiederum aufgrund der Plancherelgleichung folgt, dass diese Ausdehnung unitär ist. Die Fouriertransformation wird oft dazu verwendet, zwischen einer Orts- und Impulsraumbeschreibung von physikalischen System hin- und herzuwechseln.

b) **Quantenmechanik.** In der Quantenmechanik werden *Observable* (dh messbare Größen wie zB Energie, Impuls, Ort, Drehimpuls, Spin) eines Systems durch selbstadjungierte Operatoren $A = A^*$ auf einem Hilbertraum $\mathcal{H}$ beschrieben. Diese Operatoren sind zwar in vielen Fällen nicht beschränkt, können aber stets durch beschränkte Funktionen von A beschrieben werden (mehr dazu in Funktionalanalysis 2). Zumeist interessiert man sich für mehrere Observable, die zusammen eine Teilmenge (genauer: C^*-Algebra) $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ bilden.

Die *Zustände*, in denen das System präpariert ist, werden durch bestimmte stetige lineare Funktionale $\omega : \mathcal{A} \to \mathbb{C}$ beschrieben, dh Elemente des Dualraums $\mathcal{A}'$. Die Eigenschaften der Funktionale sind dabei so gewählt, dass eine konsistente Wahrscheinlichkeitsinterpretation möglich ist, wobei die Zahl $\omega(A)$ als der Erwartungswert (statistischer Mittelwert) der durch A beschriebenen Observable in dem durch ω beschriebenen Zustand angesehen wird.

Ein einfaches Beispiel ist ein *Vektorzustand*: Dies ist das von einem Einheitsvektor $\psi \in \mathcal{H}$, $\|\psi\| = 1$, gegebene Funktional $\omega_\psi(A) := \langle \psi, A\psi \rangle$. Dann ergibt sich die *Unschärfe* (Standardabweichung) von A im Zustand ω_ψ als

$$\Delta_\psi(A) = \sqrt{\omega_\psi(A^2) - \omega_\psi(A)^2} = \|(A - \omega_\psi(A))\psi\| \geq 0.$$

Auf diese Weise bilden Hilberträume und lineare Operatoren auf Hilberträumen die natürliche Sprache der Quantenmechanik, in der zB Unschärferelationen bewiesen werden können.

Eine detaillierte Beschreibung der Wahrscheinlichkeitsinterpretation erfordert Maß- und Spektraltheorie und wird später in dieser Vorlesung besprochen.

3 Die Hauptsätze der Funktionalanalysis

In diesem Kapitel besprechen wir einige allgemeine Sätze zu Funktionalen und Operatoren, die das Fundament für viele kommende Entwicklungen bilden. Unser Hauptinteresse liegt zwar auf Hilberträumen und Operatoren auf ihnen. Da $\mathcal{B}(\mathcal{H})$ für einen Hilbertraum $\mathcal{H}$ aber nur ein Banachraum ist, arbeiten wir durchgehend mit den allgemeineren Banachräumen.

3.1 Der Satz von Hahn-Banach

Das Lemma von Riesz (Satz 2.18) gibt uns eine genaue Beschreibung des Dualraums eines Hilbertraums, die uns im Falle eines Banachraums nicht zur Verfügung steht. Gegeben einen Banachraum X und seinen Dualraum $X' = \mathcal{B}(X, \mathbb{K})$ (ebenfalls ein Banachraum, siehe Korollar 1.20), so wissen wir aktuell noch nicht einmal, ob es vom Nullfunktional verschiedene Elemente von X' gibt. Ein erster Versuch, Elemente in X' zu finden, wäre vermutlich, eine Hamelbasis $\{e_i\}_{i\in I}$ von X und die auf der Basis definierten Koordinatenfunktionale $\varphi_j(e_i) := \delta_{ij}$ zu betrachten. Aber während diese Funktionale im Falle einer *Orthonormal*basis eines *Hilbert*raums stetig sind (es gilt dann $\varphi_j = e_j^*$, dh $\|\varphi_j\| = 1$), sind sie im Falle eines allgemeinen Banachraums typischerweise unstetig[5].

Die Hahn-Banach Sätze werden insbesondere die Existenz vieler stetiger Funktionale auf allgemeinen Banachräumen implizieren.

Wir beginnen mit einer rein algebraischen Formulierung für reelle Vektorräume (dh es wird keine Norm vorausgesetzt).

Definition 3.1. Sei X ein Vektorraum. Eine Abbildung $p : X \to \mathbb{R}$ heißt *sublinear*, falls

a) $p(\lambda x) = \lambda p(x)$ für alle $\lambda \geq 0$, $x \in X$,

b) $p(x + y) \leq p(x) + p(y)$ für alle $x, y \in X$

gilt.

Die Definition der Sublinearität ist so gewählt, dass sie sowohl lineare Funktionale als auch Halbnormen (insbesondere Normen) umfasst. Beachten Sie, dass eine sublineare Abbildung im Gegensatz zu einer Halbnorm negative Werte annehmen kann. Ein komplizierteres Beispiel einer sublinearen Abbildung ist

$$p : \ell^\infty \to \mathbb{R}, \qquad p(x) := \limsup_n x_n.$$

Theorem 3.2 (Satz von Hahn-Banach, reelle Version). *Sei X ein reeller Vektorraum, $U \subset X$ ein Untervektorraum, $p : X \to \mathbb{R}$ sublinear, und $\varphi : U \to \mathbb{R}$ ein lineares Funktional mit*

$$\varphi(x) \leq p(x), \qquad x \in U.$$

[5]Beispiel dazu: Betrachten Sie den Raum P aller reellen Polynome in einer Variable mit der Norm $\|p\| := \|p|_{[0,1]}\|_{L^1}$. Dann bilden die Monome $\{x^0, x^1, x^2, \dots\}$ eine Hamelbasis. Um zu sehen, dass zB das zu dem konstanten Monom x^0 gehörige Koordinatenfunktional unstetig ist, werten Sie es auf den Polynomen $(x-1)^n$, $n \in \mathbb{N}$, aus.

Dann existiert eine lineare Fortsetzung von φ auf X, dh ein lineares Funktional $\psi : X \to \mathbb{R}$ mit $\psi|_U = \varphi$, das

$$\psi(x) \leq p(x), \qquad x \in X,$$

erfüllt.

Bevor wir den Beweis angehen, sollten Sie sich mittels des folgenden Beispiels vor Augen führen, wie der Satz von Hahn-Banach die Existenz von vielen stetigen Funktionalen auf normierten Vektorräumen impliziert.

Beispiel 3.3. Sei X ein normierter reeller Vektorraum und $x_0 \in X \backslash \{0\}$ ein Vektor. Dann existiert ein stetiges lineares Funktional $\psi \in X'$ mit $\psi(x_0) = \|x_0\|$ und $\|\psi\|_{X'} = 1$.
Beweis: Wir betrachten den Untervektorraum $U := \mathbb{R}x_0 = \operatorname{span}\{x_0\}$, das Funktional $\varphi : U \to \mathbb{R}$, $\varphi(\lambda x_0) := \lambda\|x_0\|$, und die Norm $p = \|\cdot\|$. Dann gilt $\lambda\|x_0\| = \varphi(\lambda x_0) \leq \|\lambda x_0\| = p(\lambda x_0)$. Die Voraussetzungen des Satzes von Hahn-Banach sind also erfüllt und wir erhalten ein lineares Funktional $\psi : X \to \mathbb{R}$ mit $\psi(x_0) = \varphi(x_0) = \|x_0\|$ und $\psi(x) \leq \|x\|$ und $-\psi(x) = \psi(-x) \leq \|-x\| = \|x\|$ für alle $x \in X$, dh $\|\psi\|_{X'} = 1$.□

Beachten Sie weiterhin, dass die Aussage des Satzes von Hahn-Banach eine Existenzaussage ist. Im Allgemeinen wird die Fortsetzung des Funktionals hochgrading nicht eindeutig sein, wie wir auch im Beweis sehen werden.

Beweis. Wir betrachten zunächst den Fall, dass X eine Dimension mehr als U hat, dh es gibt einen zu U linear unabhängigen Vektor x_0 mit $X = \operatorname{span}\{U, x_0\}$. Jeder Vektor $x \in X$ ist also eindeutig in der Form $x = u + \lambda x_0$ mit $u \in U$ und $\lambda \in \mathbb{R}$ darstellbar. Eine Fortsetzung ψ von φ auf X ist dann von der Form $\psi(x) = \varphi(u) + \lambda r$ mit dem unbestimmten Parameter $r = \psi(x_0)$. Wir müssen r so wählen, dass die behauptete Ungleichung $\psi \leq p$ gilt. Dazu müssen wir

$$\varphi(u) + \lambda r \leq p(u + \lambda x_0)$$

für alle $u \in U$ und alle $\lambda \in \mathbb{R}$ sicherstellen. Für $\lambda = 0$ gilt die Abschätzung per Voraussetzung an φ. Für $\lambda > 0$ ist sie äquivalent zu

$$\begin{aligned} r &\leq \inf_{u \in U} \frac{1}{\lambda}\left(p(u + \lambda x_0) - \varphi(u)\right) = \inf_{u \in U}\left(p\left(\frac{u}{\lambda} + x_0\right) - \varphi\left(\frac{u}{\lambda}\right)\right) \\ &= \inf_{v \in U}\left(p\left(v + x_0\right) - \varphi\left(v\right)\right), \end{aligned}$$

und für $\lambda < 0$ ist sie äquivalent zu

$$\begin{aligned} r &\geq \sup_{u \in U} \frac{1}{\lambda}\left(p(u + \lambda x_0) - \varphi(u)\right) = \sup_{u \in U}\left(-p\left(-\frac{u}{\lambda} - x_0\right) - \varphi\left(\frac{u}{\lambda}\right)\right) \\ &= \sup_{v \in U}\left(-p\left(v - x_0\right) + \varphi\left(v\right)\right). \end{aligned}$$

Es existiert ein Parameter $r \in \mathbb{R}$, so dass diese beiden Ungleichungen erfüllt sind, genau dann, wenn

$$\begin{aligned} &\varphi(v) - p(v - x_0) \le p(w + x_0) - \varphi(w), \qquad \forall v, w \in U \\ \iff \qquad &\varphi(v) + \varphi(w) \le p(v - x_0) + p(w + x_0) \qquad \forall v, w \in U. \end{aligned}$$

Die letzte Bedingung ist per Voraussetzung an φ erfüllt:

$$\begin{aligned} \varphi(v) + \varphi(w) = \varphi(v + w) \le p(v + w) &= p((v - x_0) + (w + x_0)) \\ &\le p(v - x_0) + p(w + x_0). \end{aligned}$$

Also existiert eine Fortsetzung ψ von φ auf X im Falle $X = \operatorname{span}\{U, x_0\}$.

Für den allgemeinen Fall benutzen wir das Lemma von Zorn: Jede nichtleere partiell geordnete Menge, in der jede Kette (= total geordnete Teilmenge) eine obere Schranke besitzt, hat maximale Elemente (siehe Satz 2.32).

Wir betrachten die Menge aller Unterräume $U \subset V \subset X$, auf die φ linear zu einem $\psi_V : V \to \mathbb{R}$ mit $\psi \le p$ fortgesetzt werden kann,

$$\mathcal{V} := \{(V, \psi_V) : \ U \subset V \subset X \text{ Unterraum}, \psi_V : V \to \mathbb{R} \text{ linear}, \psi_V|_U = \varphi, \ \psi_V \le p\},$$

mit der Ordnungsrelation

$$(V_1, \psi_{V_1}) \le (V_2, \psi_{V_2}) \ :\Leftrightarrow \ V_1 \subset V_2, \ \psi_{V_2}|_{V_1} = \psi_{V_1}.$$

Wegen $(U, \varphi) \in \mathcal{V}$ ist $\mathcal{V}$ nicht leer. Es ist klar, dass $\le$ eine partielle Ordnung auf $\mathcal{V}$ ist, dh dass $(V, \psi_V) \le (V, \psi_V)$ gilt (Reflexivität), dass $\le$ transitiv ist, und dass $\le$ antisymmetrisch ist $(V_1, \psi_{V_1}) \le (V_2, \psi_{V_2})$ und $(V_2, \psi_{V_2}) \le (V_1, \psi_{V_1})$ impliziert $(V_1, \psi_{V_1}) = (V_2, \psi_{V_2})$.

Sei also $K = \{(V_i, \psi_{V_i})\}_{i \in I} \subset \mathcal{V}$ eine Kette. Dann setzen wir

$$V := \bigcup_{i \in I} V_i, \qquad \psi_V(x) := \psi_{V_i}(x) \text{ für } x \in V_i.$$

Da K total geordnet ist, sind V und ψ_V wohldefiniert (insbesondere hängt die Definition von ψ_V nicht von einer Wahl von i ab). Offenbar gilt $(V_i, \psi_{V_i}) \le (V, \psi_V)$ für alle $i \in I$, dh $(V, \psi_V) \in \mathcal{V}$ ist eine obere Schranke.

Nach dem Lemma von Zorn gibt es also ein bzgl der Relation $\le$ maximales Element $(V, \psi_V) \in \mathcal{V}$. Angenommen, $V \ne X$. Dann gibt es einen von V linear unabhängigen Vektor $x \in X$, und nach dem ersten Beweisschritt eine Forsetzung von ψ_V (und damit von φ) auf $\operatorname{span}\{V, x_0\} \supsetneq V$, im Widerspruch zur Maximalität von V. Also gilt $V = X$, und der Beweis ist abgeschlossen. □

Diese Version des Satzes von Hahn-Banach gilt nur für reelle Vektorräume, da die Ungleichungen $\varphi \le p$ für komplexe Zahlen nicht definiert sind. Andererseits haben wir es sehr häufig mit komplexen Vektorräumen zu tun. Wir besprechen deshalb gleich eine komplexe Version, die wir mit einem Lemma über reell- und komplex lineare Funktionale auf komplexen Vektorräumen vorbereiten.

Lemma 3.4. *Sei X ein komplexer Vektorraum.*

a) Ist $\varphi : X \to \mathbb{R}$ ein $\mathbb{R}$-lineares Funktional, so ist

$$\varphi_{\mathbb{C}} : X \to \mathbb{C}, \qquad \varphi_{\mathbb{C}}(x) := \varphi(x) - i\varphi(ix)$$

ein $\mathbb{C}$-lineares Funktional, und es gilt $\varphi = \operatorname{Re} \varphi_{\mathbb{C}}$.

b) Ist $\gamma : X \to \mathbb{C}$ ein $\mathbb{C}$-lineares Funktional, so ist $\varphi := \operatorname{Re} \gamma$ ein $\mathbb{R}$-lineares Funktional und $\varphi_{\mathbb{C}} = \gamma$.

c) Sei p eine Halbnorm und φ ein $\mathbb{C}$-lineares Funktional auf X. Dann gilt die Äquivalenz

$$|\varphi| \le p \iff |\operatorname{Re} \varphi| \le p.$$

Dabei sind beide Ungleichungen punktweise asugewertet auf beliebigen $x \in X$ zu verstehen.

d) Ist X normiert und φ ein $\mathbb{C}$-lineares stetiges Funktional auf X, so gilt $\|\varphi\| = \|\operatorname{Re} \varphi\|$.

Beweis. a) $\varphi_{\mathbb{C}}$ ist offenbar $\mathbb{R}$-linear. Wegen $\varphi_{\mathbb{C}}(ix) = \varphi(ix) - i\varphi(iix) = i\varphi(x) + i(-i)\varphi(ix) = i\varphi_{\mathbb{C}}(x)$ ist $\varphi_{\mathbb{C}}$ auch $\mathbb{C}$-linear. Die Gleichung $\varphi = \operatorname{Re} \varphi_{\mathbb{C}}$ gilt per Konstruktion.

b) Die $\mathbb{R}$-Linearität von φ ist klar. Um $\gamma = \varphi_{\mathbb{C}}$ zu zeigen, berechnen wir auf beliebigem $x \in X$

$$\gamma(x) = \operatorname{Re} \gamma(x) + i\operatorname{Im} \gamma(x) = \varphi(x) - i\operatorname{Re} i\gamma(x) = \varphi(x) - i\operatorname{Re} \gamma(ix) = \varphi_{\mathbb{C}}(x).$$

c) Die Implikation $\Rightarrow$ folgt direkt aus $|\operatorname{Re} z| \le |z|$. Für die umgekehrte Implikation sei $x \in X$ mit $\varphi(x) \neq 0$. Dann gibt es eine komplexe Zahl z mit Betrag $|z| = 1$, so dass $\varphi(x) = z|\varphi(x)|$ (Polardarstellung). Damit gilt

$$|\varphi(x)| = z^{-1}\varphi(x) = \varphi(z^{-1}x) = |\operatorname{Re}(\varphi(z^{-1}x))| \le p(z^{-1}x) = p(x).$$

d) ergibt sich aus c) durch Betrachtung von $p(x) = \|\varphi\|\|x\|$. □

Theorem 3.5 (Satz von Hahn-Banach, komplexe Version). *Sei X ein komplexer Vektorraum, $U \subset X$ ein Unterraum, $p : X \to \mathbb{R}$ sublinear und $\varphi : U \to \mathbb{C}$ ein lineares Funktional mit $\operatorname{Re} \varphi \le p$. Dann existiert ein lineares Funktional $\psi : X \to \mathbb{C}$ mit $\psi|_U = \varphi$ und $\operatorname{Re} \psi \le p$.*

Beweis. Das reelle Funktional $\operatorname{Re} \varphi$ erfüllt die Voraussetzungen des reellen Hahn-Banach Theorems, also gibt es ein reelles Funktional $\psi_{\mathbb{R}} : X \to \mathbb{R}$ mit $\psi_{\mathbb{R}} \le p$. Nach Teil a) des Lemmas gibt es ein komplexes Funktional ψ mit $\psi_{\mathbb{R}} = \operatorname{Re} \psi$. Nach Teil b) des Lemmas stimmt $\psi|_U$ mit φ überein, ist also eine Fortsetzung. □

Die häufigste Anwendung des Satzes von Hahn-Banach bezieht sich auf die Fortsetzung von linearen Funktionalen, wie bereits in Beispiel 3.3 angedeutet. Die allgemeine Version halten wir noch einmal als Satz fest.

Satz 3.6 (Fortsetzungssatz von Hahn-Banach). *Sei X ein normierter Vektorraum (über $\mathbb{R}$ oder $\mathbb{C}$), $U \subset X$ ein Untervektorraum, und $\varphi : U \to \mathbb{K}$ ein stetiges lineares Funktional (dh $\varphi \in U'$). Dann existiert ein stetiges lineares Funktional $\psi \in X'$ mit*

$$\psi|_U = \varphi, \qquad \|\psi\|_{X'} = \|\varphi\|_{U'}.$$

Beweis. Wir betrachten die Halbnorm $p(x) = \|\varphi\|\|x\|$ und unterscheiden zwischen dem reellen und komplexen Fall.

Im reellen Fall betrachten wir die Die reelle Version des Satzes von Hahn-Banach liefert uns ein Funktional $\psi : X \to \mathbb{R}$ mit $\psi|_U = \varphi$ und $\psi(x) \le p(x)$ und $-\psi(x) = \psi(-x) \le p(x)$, also $|\psi(x)| \le \|\varphi\|\|x\|$ und damit $\|\psi\| = \|\varphi\|$.

Im komplexen Fall gilt $\operatorname{Re} \varphi(u) \le |\varphi(u)| \le \|\varphi\|\|u\| = p(u)$ für alle $u \in U$. Der komplexe Hahn-Banach Satz liefert uns also ein komplexes komplex lineares Funktional $\psi : X \to \mathbb{C}$ mit $\operatorname{Re} \psi \le p$. Nach Teil c) des Lemmas gilt dann auch $|\psi| \le p$, also $\|\psi\| \le \|\varphi\|$ und damit $\|\psi\| = \|\varphi\|$. □

Wir diskutieren nun einige Konsequenzen des Satzes von Hahn-Banach.

Korollar 3.7. *Sei X ein normierter Raum.*

a) Zu jedem $x \in X$, $x \neq 0$, existiert ein stetiges Funktional $\varphi \in X'$ mit $\|\varphi\| = 1$ und $\varphi(x) = \|x\|$. Insbesondere trennt X' die Punkte von X, dh zwei Vektoren $x, y \in X$ sind genau dann gleich, wenn sie $\varphi(x) = \varphi(y)$ für alle $\varphi \in X'$ erfüllen.

b) Für jeden Vektor $x \in X$ gilt $\|x\| = \sup\limits_{\varphi \in X', \|\varphi\|=1} |\varphi(x)|$.

Beweis. a) Die erste Aussage haben wir bereits in Beispiel 3.3 gezeigt. Die zweite Aussage folgt durch Betrachtung von $x - y$.

b) Per Definition der Norm in X' gilt $|\varphi(x)| \le \|\varphi\|\|x\|$ und damit die Ungleichung $\ge$ in der behaupteten Gleichung. Für die umgekehrte Ungleichung betrachten wir $x \neq 0$ (der Fall $x = 0$ ist trivial) und wählen ein Funktional $\varphi \in X'$ mit $\|\varphi\| = 1$ und $\varphi(x) = \|x\|$ wie in Teil a). Damit gilt $\|x\| = \varphi(x) \le \sup\limits_{\varphi \in X', \|\varphi\|=1} |\varphi(x)|$. □

Folgendes Beispiel illustriert, wie der Satz von Hahn-Banach Information über Dualräume geben kann. Es gibt auch einen guten Eindruck, was für Fortsetzungen von Funktionalen möglich sind.

Beispiel 3.8. Die Abbildung $T : \ell^1_{\mathbb{N}} \to (\ell^\infty_{\mathbb{N}})'$, $(Tx)(y) = \sum_{n=1}^\infty x_n y_n$ ist isometrisch aber nicht surjektiv.
Für $x \in \ell^1_{\mathbb{N}}$ und $y \in \ell^\infty_{\mathbb{N}}$ gilt $|(Tx)(y)| \le \|x\|_1 \|y\|_\infty$, also $\|Tx\| \le \|x\|_1$. Nun definieren wir zu $x \in \ell^1_{\mathbb{N}}$ die Folge $y_n := |x_n|/x_n$ falls $x_n \neq 0$ und $y_n = 0$ falls $x_n = 0$. Dann gilt $(Tx)(y) = \sum_n |x_n| = \|x\|_1$ und $\|y\|_\infty = 1$, also $\|Tx\| \ge \|x\|_1$. Also gilt $\|Tx\| = \|x\|_1$, dh T ist isometrisch.

Für die Nicht-Surjektivität führen wir einen Widerspruchsweis per Hahn-Banach. Dazu betrachten wir den Unterraum $c \subset \ell_\infty$ aller konvergenter Folgen und das Grenzwertfunktional $\lim : c \to \mathbb{K}$, $y \mapsto \lim_n y_n$. Wegen $|\lim_n y_n| \leq \|y\|_\infty$ ist dies ein stetiges Funktional auf c, und nach dem Fortsetzungssatz von Hahn-Banach gibt es ein stetiges Funktional $\mathrm{LIM} : \ell^\infty_\mathbb{N} \to \mathbb{K}$, dass lim fortsetzt. Falls T surjektiv wäre, so wäre $\mathrm{LIM} = Tx$ für ein $x \in \ell^1_\mathbb{N}$. Damit würde auf e_k (mit $(e_k)_n = \delta_{n,k}$)

$$x_k = (Tx)(e_k) = \mathrm{LIM}(e_k) = \lim_{n\to\infty} (e_k)_n = 0 \tag{3.1}$$

gelten, also $x = 0$. Aber dies steht im Widerspruch zu $\mathrm{LIM}|_c = \lim$.

Die in diesem Beispiel betrachteten Ausdehnungen des lim-Funktionals heißen *Banach-Limes*. Man kann zeigen, dass LIM positiv (dh $x \geq 0 \Rightarrow \mathrm{LIM}(x) \geq 0$) und shift-invariant gewählt werden kann.

In einer Übungsaufgabe auf Blatt 7 werden Sie zeigen, dass der Dualraum des Banachraums c_0 aller Nullfolgen isometrisch isomorph zu $\ell^1_\mathbb{N}$ ist, $(c_0)' \cong \ell^1_\mathbb{N}$. Dies geschieht mittels der gleichen Idee wie in Beispiel 3.8. Diese Abbildung stiftet auch einen isometrischen Isomorphimus $(\ell^p_\mathbb{N})' \cong \ell^q_\mathbb{N}$ für $1 \leq p < \infty$ und $\frac{1}{p} + \frac{1}{q} = 1$. Wir haben also insbesondere für $1 < p < \infty$

$$(\ell^p_\mathbb{N})' \cong \ell^q_\mathbb{N}, \qquad (\ell^q_\mathbb{N})' \cong \ell^p_\mathbb{N} \quad \Rightarrow \quad (\ell^p_\mathbb{N})'' \cong \ell^p_\mathbb{N}, \quad 1 < p < \infty,$$

wobei $(\ell^p_\mathbb{N})''$ den *Bidualraum* (=Dualraum des Dualraums) von $\ell^p_\mathbb{N}$ bezeichnet. Dies beinhaltet insbesondere den Hilbertraumfall $p = 2$.

Das Ergebnis der oben angesprochenen Aufgabe und Beispiel 3.8 sagen hingegen, dass diese Abbildungen *keinen* Isomorphimus $\ell^1 \cong (\ell^\infty)'$, wegen $\ell^\infty \cong (\ell^1)'$ also auch *keinen* Isomorphismus $\ell^1 \cong (\ell^1)''$. Man kann zeigen, dass es überhaupt keinen isometrischen Isomorphismus zwischen ℓ^1 und $(\ell^\infty)'$ gibt, da ℓ^1 separabel ist, $(\ell^\infty)'$ aber nicht (siehe Buch von Werner S. 112).

Für einen allgemeinen normierten Raum X gibt es eine kanonische Abbildung von X nach X''. Dazu betrachten wir zu einem Vektor $x \in X$ die Auswertungsbbildung

$$I(x) : X' \to \mathbb{K}, \qquad \varphi \mapsto \varphi(x).$$

Offenbar ist $I(x)$ linear, und wegen $|\varphi(x)| \leq \|\varphi\|_{X'}\|x\|_X$ ist $I(x)$ beschränkt, also ein Element des Bidualraums X''. Damit erhalten wir die kanonische Abbildung

$$I : X \to X'', \qquad I(x)(\varphi) := \varphi(x). \tag{3.2}$$

Satz 3.9. *Sei X ein normierter Raum. Dann ist die kanonische Abbildung $I : X \to X''$ eine Isometrie.*

Beweis. Nach Korollar 3.7 b) gilt $\|I(x)\|_{X''} = \sup_{\varphi\in X', \|\varphi\|_{X'}=1} |\varphi(x)| = \|x\|_X$. □

Die Isometrie I ist im Allgemeinen nicht surjektiv. Für nicht vollständiges X ist sie nie surjektiv, denn sonst hätten wir einen Isomorphimus zwischen dem nicht vollständigen Raum X und dem vollständigen Raum X'' (als Dualraum des normierten Raumes X' ist X'' immer vollständig), ein Widerspruch zu Lemma 1.25. Aber auch für vollständiges X muss I nicht surjektiv sein. Man nennt einen Banachraum X *reflexiv*, wenn I surjektiv ist. (Es gibt nicht reflexive Banachräume, die isometrisch isomorph zu ihrem Bidualraum sind.)

Alle Hilberträume sind reflexiv, wie man mit Hilfe des Lemmas von Riesz zeigen kann. Auch $\ell^p_{\mathbb{N}}$, $1 < p < \infty$, sind reflexiv, und c_0, ℓ^1, ℓ^∞ sind Beispiele nicht reflexiver Banachräume (ohne Beweis).

Wir schließen unsere Diskussion des Satzes von Hahn-Banach dadurch, dass wir die Lücke unseres Vervollständigungssatzes Satz 1.33 a) schließen. Dazu müssen wir zeigen, dass jeder normierte Raum X eine Vervollständigung (Y, Φ) besitzt, also einen Banachraum Y und eine Isometrie $\Phi : X \to Y$ mit dichtem Bild. Die kanonische Abbildung $I : X \to X''$ liefert schon eine Isometrie in einen Banachraum, nämlich X''. Definieren wir nun $Y := \overline{I(X)}^{\|\cdot\|_{X''}}$ und $\Phi : x \mapsto I(x) \in Y$, so sind wir fertig.

3.2 Gleichmäßige Beschränktheit

Nach dem Satz von Hahn-Banach wenden wir uns einem weiteren Theorem zu, das meist als Prinzip der gleichmäßigen Beschränktheit bekannt ist.

Als Vorbereitung dazu besprechen wir mit dem Satz von Baire einen wichtigen Satz über vollständige metrische Räume. Da es unterschiedliche Formulierungen gibt, erinnern wir zunächst an einige Begriffe.

Sei (X, d) ein metrischer Raum und $A \subset X$ eine Teilmenge.

- Ein Punkt $a \in A$ heißt *innerer Punkt von* A, wenn ein Radius $\varepsilon > 0$ existiert, so dass $U_\varepsilon(a) := \{x \in X : d(x, a) < \varepsilon\} \subset A$.
- A heißt *offen*, wenn jedes $a \in A$ ein innerer Punkt ist. Das Innere A° einer Menge ist die Menge aller inneren Punkte von A.
- A heißt *dicht*, wenn $\overline{A} = X$, dh wenn es zu jedem $x \in X$ eine Folge $(a_n) \subset A$ mit $a_n \to x$ gibt.
- A heißt *nirgends dicht*, wenn $\overline{A}$ keine inneren Punkte besitzt, also $(\overline{A})^\circ = \emptyset$. Dies ist äquivalent dazu, dass das Komplement $X \backslash \overline{A}$ dicht in X ist.

Zum Beispiel sind $\mathbb{Q}$ und $\mathbb{R} \backslash \mathbb{Q}$ dichte Teilmengen von $\mathbb{R}$. Jede endliche Teilmenge von $\mathbb{R}$ ist nirgends dicht, und auch zB $\mathbb{Z}$. Die Menge $A := (0, 1) \cap \mathbb{Q} \subset \mathbb{R}$ ist nicht dicht (denn der Abschluss ist $\overline{A} = [0, 1] \neq \mathbb{R}$, aber nicht nirgends dicht (denn der Abschluss enthält innere Punkte; sein Inneres ist $(0, 1)$)).

> Zeigen Sie, dass eine Menge $A \subset X$ genau dann nirgends dicht ist, wenn ihr Abschluss nirgends dicht ist. Folgern Sie: Ist $A \subset X$ eine nirgends dichte Menge und $B \subset X$ offen, so gilt stets $(X \backslash \overline{A}) \cap B \neq \emptyset$.

> **Satz 3.10 (Satz von Baire).** *Ein vollständiger metrischer Raum X ist nie eine abzählbare Vereinigung von nirgends dichten Mengen.*

Anders formuliert: Falls $(A_n)_n$ eine Folge von Teilmengen von X ist, so dass $X = \bigcup_n A_n$, so gibt es mindestens ein A_n, so dass $\overline{A}_n$ innere Punkte enthält.

Beweis. Wir betrachten eine Folge nirgends dichter Teilmengen $A_n \subset X$ und ihre Vereinigung $A := \bigcup_n A_n$. Wir werden eine Cauchyfolge (x_m) konstruieren, deren Grenzwert x in keinem der A_n liegt. Da x aber in X liegt, folgt dann $A \neq X$ und damit die Behauptung.

Gemäß der Übung vor dem Satz dürfen wir annehmen, dass die A_n abgeschlossen sind.

Zur Konstruktion dieser Folge: Da A_1 nirgends dicht ist, gilt insbesondere $A_1 \neq X$. Also finden wir einen Punkt $x_1 \notin A_1$. Da $X \backslash A_1$ offen ist, gibt es eine offene Kugel $B_1 := U_{\varepsilon_1}(x_1)$ mit Mittelpunkt x_1 und Radius $0 < \varepsilon_1 < 1$, so dass $B_1 \subset X \backslash A_1$, also $B_1 \cap A_1 = \emptyset$.

Da A_2 nirgends dicht ist, ist die offene Menge $B_1 \backslash A_2$ nicht leer (siehe Übung vor dem Satz), und wir wählen $x_1 \in B_1 \backslash A_2$ zusammen mit einer Kugel $B_2 = U_{\varepsilon_2}(x_2)$ mit $0 < \varepsilon_2 < \frac{1}{2}$ und $\overline{B}_2 \subset B_1$, $B_2 \cap A_2 = \emptyset$.

Induktiv wählen wir $x_n \in B_{n-1} \backslash A_n$ und wählen eine Kugel $B_n = U_{\varepsilon_n}(x_n)$ mit $0 < \varepsilon_n < 2^{1-n}$ und $\overline{B}_n \subset B_{n-1}$, $B_n \cap A_n = \emptyset$.

Die so konstruierte Folge ist eine Cauchyfolge: Für $n, m \in \mathbb{N}$, $n \geq m$, gilt $x_n, x_m \in B_{n-1}$, also $d(x_n, x_m) \leq 2 \cdot \varepsilon_{n-1} < 2^{1-n} \to 0$ für $n \to \infty$. Sei $x := \lim_n x_n$ der Grenzwert dieser Folge und $N \in \mathbb{N}$ beliebig. Da $x_n \in B_N$ für alle $n > N$, gilt $x \in \overline{B}_N \subset B_{N-1}$, was nach Konstruktion disjunkt zu A_{N-1} ist. □

Wir werden diesen Satz zwar meistens nur via gleich formulierte Konsequenzen anwenden, notieren hier aber zur Demonstration eine interessante direkte Anwendung.

Beispiel 3.11. Es gibt stetige Funktionen $f : [0, 1] \to \mathbb{R}$, die an keinem Punkt differenzierbar sind.
An dieser Aussage sieht zwar zuerst nichts nach dem Satz von Baire aus, aber ein Beweis mittels Baire ist trotzdem möglich. Dazu betrachten wir den Banachraum $C([0, 1])$ mit Supremumsnorm und die Mengen

$$A_n := \left\{ f \in C([0,1]) : \exists x : \forall x' \neq x : \left| \frac{f(x) - f(x')}{x - x'} \right| \leq n \right\}.$$

Wir werden zeigen, dass diese Mengen nirgends dicht in $C([0, 1])$ sind. Also gilt nach dem Satz von Baire $\bigcup_n A_n \neq C([0, 1])$. Sei f eine Funktion, die in keinem A_n liegt. Diese erfüllt dann: Für jedes $x \in [0, 1]$ gibt es $x' \neq x$ mit $\left| \frac{f(x)-f(x')}{x-x'} \right| > n$. Also ist der Differenzenquotient bei festem x nicht stetig in x' auf das ganze Intervall fortsetzbar, denn sonst wäre er beschränkt. Also ist f nirgends differenzierbar.

Es bleibt also zu zeigen, dass A_n nirgends dicht ist, dh dass $\overline{A}_n$ keine inneren Punkte hat. Dazu zeigen wir zunächst, dass A_n abgeschlossen ist.

Sei also $(f_k)_{k \in \mathbb{N}}$ eine bzgl Supremumsnorm konvergente Folge in A_n mit Grenzwert $f \in C([0, 1])$. Wir müssen $f \in A_n$ zeigen. Wegen $f_k \in A_n$ gibt es $x_k \in [0, 1]$, so dass

$$|f_k(x_k) - f_k(y)| \leq n|x_k - y|, \qquad x_k \in [0, 1].$$

Die Folge $(x_k)_{k\in\mathbb{N}}$ ist beschränkt, hat also eine konvergente Teilfolge. Nach Übergang zu einer Teilfolge $(k_l)_{l\in\mathbb{N}}$ haben wir weiterhin $f_{k_l} \to f$, und $x_{k_l} \to x$ für ein $x \in [0,1]$. Wir dürfen also oBdA annehmen, dass die Folge $(x_k)_{k\in\mathbb{N}}$ konvergiert. Dann erhalten wir

$$|f_k(x_k) - f(x)| \le |f_k(x_k) - f(x_k)| + |f(x_k) - f(x)| \le \|f_k - f\|_\infty + |f(x_k) - f(x)| \to 0$$

für $k \to \infty$ aufgrund der Konvergenz von (f_k) in der Stetigkeit von f.

Also gilt für alle $y \in [0,1]$

$$|f(x) - f(y)| = \lim_{k\to\infty} |f_k(x_k) - f_k(y)| \le \lim_{k\to\infty} n|x_k - y| = n|x-y|.$$

Also gilt $f \in A_n$, dh A_n ist abgeschlossen.

Wir müssen nun noch zeigen, dass A_n keine inneren Punkte hat. Dazu führen wir einen Widerspruchsweis und nehmen an, dass A_n einen inneren Punkt hat (für irgendein n). Nach dem ersten Weierstraßschen Approximationssatz (die Polynome liegen dicht in $C([0,1])$) kann der innere Punkt durch eine Folge von Polynomen in Supremumsnorm angenähert werden, dh A_n° enthält ein Polynom p. Wir betrachten nun eine Funktion mit kleiner Supremumsnorm aber großen Differenzenquotienten, um einen Widerspruch zu erzeugen. Sei dazu $m \in \mathbb{N}$, $\varepsilon > 0$ und $S_{\varepsilon,m}$ eine "Sägezahnfunktion", die auf den Teilintervallen $(k\varepsilon/m, (k+1)\varepsilon/m)$ eine Gerade mit Steigung $\pm m$ ist und Bild $[0,\varepsilon]$ hat (siehe Tafelbild). Dann gilt per Konstruktion a) $\|S_{\varepsilon,m}\|_\infty = \frac{1}{\varepsilon}$ und b) für jedes $x \in [0,1]$ gibt es $x' \in [0,1]$, so dass $\left|\frac{S_{\varepsilon,m}(x) - S_{\varepsilon,m}(x')}{x-x'}\right| = m$.

Aufgrund von a) gilt $p + S_{\varepsilon,m} \in A_n$ für ε klein genug. Es gibt also ein x, so dass für alle $x' \neq x$

$$n \ge \left|\frac{(p+S_{\varepsilon,m})(x) - (p+S_{\varepsilon,m})(x')}{x-x'}\right| \ge \left|\frac{S_{\varepsilon,m}(x) - S_{\varepsilon,m}(x')}{x-x'}\right| - \|p'\|_\infty$$

gilt, wobei wir die umgekehrte Dreiecksungleichung und den Mittelwertsatz verwendet haben. Wählen wir jetzt x' nach b) so, dass der erste Term $= m$ ist, so erhalten wir $n \ge m - \|p\|_\infty$, was für groß genügendes m einen Widerspruch ergibt.

Die am häufigsten verwendete Konsequenz des Satzes von Baire ist das Prinzip der gleichmäßigen Beschränktheit in seinen verschiedenen Formen.

Theorem 3.12 (Prinzip der gleichmäßigen Beschränktheit). *Sei X ein Banachraum, Y ein normierter Raum, und $\mathcal{T} \subset \mathcal{B}(X,Y)$ eine Menge von beschränkten Operatoren. Falls $\mathcal{T}$ punktweise beschränkt ist, dh für jedes $x \in X$ ist die Menge $\{Tx : T \in \mathcal{T}\} \subset Y$ beschränkt, so ist $\mathcal{T}$ gleichmäßig beschränkt, dh $\mathcal{T} \subset \mathcal{B}(X,Y)$ ist beschränkt.*

Die Annahme dieses Satzes sagt also:

- Für jedes $x \in X$ gibt es ein $c(x) > 0$, so dass $\|Tx\|_Y \le c(x)$ für alle $T \in \mathcal{T}$ gilt, dh

$$\forall x \in X \;:\; \sup_{T\in\mathcal{T}} \|Tx\|_Y < \infty,$$

und die Aussage des Satzes sagt:

- Es gibt es ein $C > 0$, so dass $\|T\|_{\mathcal{B}(X,Y)} \leq C$ für alle $T \in \mathcal{T}$ gilt, dh $\|Tx\|_Y \leq C\|x\|_X$ für alle $x \in X$ und alle $T \in \mathcal{T}$, dh

$$\sup_{T\in\mathcal{T}} \|T\|_{\mathcal{B}(X,Y)} < \infty.$$

Während in der ersten Abschätzung die Schranken von x abhängen, sind sie in der zweiten Abschätzung gleichmäßig in x, was den Namen "Prinzip der gleichmäßigen Beschränktheit" erklärt.

Beweis. Wir betrachten die Mengen $B_n := \{x : \|Tx\|_Y \leq n \forall T \in \mathcal{T}\}$, die aufgrund der Stetigkeit der T abgeschlossen sind. Gemäß unserer Annahme gilt $X = \bigcup_n B_n$. Nach dem Satz von Baire gibt es also ein n, so dass $\overline{B_n} = B_n$ innere Punkte hat.

Sei $x \in B_n$ ein innerer Punkt und $\varepsilon > 0$ so, dass $U_\varepsilon(x) \subset B_n$. Sei nun $T \in \mathcal{T}$ beliebig und $y \in X$ ein Vektor mit Norm $\|y\| < \varepsilon$. Dann gilt $x + y \in B_n$, also

$$\|Ty\| \leq \|T(x+y)\| + \|Tx\| \leq 2n.$$

Für einen beliebigen Vektor $z \in X$ gilt $\|\frac{\varepsilon}{2\|z\|}z\| < \varepsilon$ und damit $\|Tz\| = \frac{2\|z\|}{\varepsilon} \cdot 2n$, also $\|T\| \leq 4n/\varepsilon$. □

Dieses praktische Resultat hat viele Anwendungen. Die erste bezieht sich auf einen schwächeren Konvergenzbegriff für Folgen von Operatoren (punktweise Konvergenz), vergleiche Übung H6.1 c),d). Wir sagen, eine Folge $(A_n)_{n\in\mathbb{N}}$ stetiger linearer Operatoren zwischen zwei normierten Räumen X und Y konvergiere *punktweise*, falls die Folgen $(A_n x)_n$ für alle $x \in X$ konvergieren.

Satz 3.13. *Sei X ein Banachraum, Y ein normierter Raum und $A_n \in \mathcal{B}(X,Y)$, $n \in \mathbb{N}$, eine Folge von beschränkten Operatoren. Falls (A_n) punktweise konvergiert, so definiert $Ax := \lim_n Ax_n$ einen beschränkten Operator $X \to Y$, und die Folge (A_n) ist in $\mathcal{B}(X,Y)$ beschränkt.*

Beweis. Sei $x \in X$. Da die Folge $(A_n x)_n$ nach Voraussetzung in Y konvergiert, ist sie beschränkt. Nach dem Prinzip der gleichmäßigen Beschränktheit gibt es also eine Konstante $C > 0$ mit $\|A_n\| \leq C$ für alle n. Demnach gilt

$$\|Ax\| = \|\lim_n A_n x\| = \lim_n \|A_n x\| \leq C\|x\|, \qquad x \in X,$$

also $\|A\| \leq C$. □

Beachten Sie, dass *nicht* behauptet wird, dass die Folge A_n gegen A *in Operatornorm* konvergiert - das ist im Allgemeinen falsch!

Beispiel 3.14. Nehmen Sie zB die Operatoren $A_n \in \mathcal{B}(\ell^2_{\mathbb{N}})$,

$$A_n(x_1, x_2, x_3, \ldots) = (0, \ldots 0, x_{n+1}, x_n, \ldots).$$

Dann gilt $A_n x \to 0$ für alle $x \in \ell^2_{\mathbb{N}}$. Wegen $A_n e_{n+1} = e_{n+1}$ (mit $(e_n)_k = \delta_{nk}$) gilt aber $\|A_n\| \geq 1$ für alle n, und damit geht $\|A_n - 0\|$ nicht gegen Null.

Die Aussage von Satz 3.13 ist vielmehr, dass der Banachraum $\mathcal{B}(X, Y)$ (für X, Y Banach) nicht nur unter Limiten in der Operatornorm abgeschlossen (vollständig) ist, sondern auch unter den "schwächeren" punktweisen Limiten.

In Satz 3.13 haben wir nur X aber nicht Y als vollständig vorausgesetzt. Wenn auch Y vollständig ist, gilt folgender Satz.

Satz 3.15 (Satz von Banach-Steinhaus). *Seien X, Y Banachräume und $(A_n)_{n\in\mathbb{N}} \subset \mathcal{B}(X, Y)$ eine Folge beschränkter linearer Operatoren. Die Folge (A_n) konvergiert punktweise gegen einen beschränkten Operator $A \in \mathcal{B}(X, Y)$ genau dann, wenn sie a) beschränkt ist (in Operatornorm) und b) eine dichte Teilmenge $D \subset X$ existiert, so dass $(A_n x)_{n\in\mathbb{N}}$ für alle $x \in D$ konvergiert.*

Beweis. Die eine Implikation wurde bereits gezeigt.

Sei $x \in X$ und $\varepsilon > 0$. Da $D \subset X$ dicht ist, gibt es $\tilde{x} \in D$ mit $\|x - \tilde{x}\| < \varepsilon$. Wir wissen, dass $(A_n\tilde{x})_n$ konvergiert. Also gibt es $N \in \mathbb{N}$, so dass $\|A_n\tilde{x} - A_m\tilde{x}\| < \varepsilon$ für $n, m \geq N$. Weiterhin ist (A_n) beschränkt, dh es gibt $C > 0$, so dass $\|A_n\| \leq C$ für alle n. Damit erhalten wir

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

Also ist $(A_n x)_n$ eine Cauchyfolge, die aufgrund der Vollständigkeit von Y konvergiert. □

Beispiel 3.16. Als ein Beispiel betrachten wir $\mathcal{H} = L^2(\mathbb{R})$, eine beschränkte Folge von Funktionen $(g_n)_{n\in\mathbb{N}} \in L^\infty(\mathbb{R})$ mit Träger $\operatorname{supp}(g_n) \subset [0, 1]$, und die Operatoren

$$(A_n f)(x) := g_n(x - n) f(x).$$

Gemäß früherer Übungen gilt $\|A_n\| \leq \|g_n\|_{L^\infty}$, dh die (A_n) sind beschränkte Operatoren und ihre Operatornormen sind uniform beschränkt. Sei nun f eine stetige Funktion mit kompaktem Träger. Dann gilt $A_n f = 0$ für groß genügendes n, denn wegen $\operatorname{supp} g_n(\,\cdot\, - n) \subset [n, n+1]$ sind die Träger von f und $g_n(\,\cdot\, - n)$ für großes n disjunkt. Insbesondere konvergiert $A_n f$ für $f \in C_c(\mathbb{R})$ in $L^2(\mathbb{R})$ gegen Null. Also konvergiert $A_n f$ für beliebiges $f \in L^2(\mathbb{R})$ gegen Null.

Dieses Ergebnis hätten Sie auch recht leicht durch direkte Abschätzung erhalten können, das Beispiel dient also mehr der Illustration des Satzes von Banach-Steinhaus.

Diese punktweisen Limiten vertragen sich auch gut mit Operatorprodukten:

Lemma 3.17. *Sei X ein Banachraum und (A_n), (B_n) zwei Folgen in $\mathcal{B}(X)$, die punktweise gegen A bzw B konvergieren. Dann konvergiert $(A_n B_n)_n$ punktweise gegen AB.*

Beweis. Wir wissen bereits, dass beide Folgen beschränkt sind, dh es gibt $C > 0$ mit $\|A_n\|, \|B_n\| \leq C$ für alle n. Damit erhalten wir auf einem beliebigen Vektor $x \in X$

$$\begin{aligned}\|A_nB_nx - ABx\| &= \|A_n(B_nx - Bx) + (A_n - A)Bx\| \\ &\leq \|A_n(B_nx - Bx)\| + \|(A_n - A)Bx\| \\ &\leq C\|B_nx - Bx\| + \|(A_n - A)Bx\| \to 0.\end{aligned}$$

□

Als eine weitere Anwendung des Prinzips der gleichmäßigen Beschränktheit betrachten wir auch noch schwache Folgenkonvergenz für Folgen von Vektoren in Banachräumen (nicht Operatoren), siehe Hausaufgabe H6.2.

Definition 3.18. Sei X ein normierter Vektorraum. Eine Folge $(x_n) \subset X$ heißt *schwach konvergent*, falls für alle $\varphi \in X'$ die Folge $(\varphi(x_n))_n \subset \mathbb{K}$ konvergent ist.

Sie haben in H6.2 gezeigt: Konvergente Folgen sind auch schwach konvergent, aber es gibt schwach konvergente Folgen, die nicht konvergent (in Norm) sind. Außerdem haben Sie gezeigt, dass eine schwach konvergente Folge einen eindeutigen Limesvektor hat, dh $\varphi(x_n) \to \varphi(x)$ für alle $\varphi \in X'$ für ein eindeutiges $x \in X$.

Wir zeigen hier zusätzlich:

Lemma 3.19. *Eine Teilmenge M eines normierten Vektorraums X ist genau dann beschränkt, wenn für jedes $\varphi \in X'$ die Menge $\varphi(M) \subset \mathbb{K}$ beschränkt ist.*

Insbesondere ist in einem normierten Vektorraum jede schwach konvergente Folge beschränkt.

Beweis. Ist $M \subset X$ beschränkt, dh gibt es $C > 0$ mit $\|m\| \leq C$ für alle $m \in M$, so gilt natürlich $|\varphi(m)| \leq C\|\varphi\|$ für alle $m \in M$.

Für die Umkehrung benutzen wir das Prinzip der gleichmäßigen Beschränktheit und die Auswertungsfunktionale $I(x) : X' \to \mathbb{K}$, $\varphi \mapsto \varphi(x)$. Nach Voraussetzung gilt

$$\sup_{m\in M} |\varphi(m)| = \sup_{m\in M} |I(m)(\varphi)| < \infty.$$

Da X' ein Banachraum ist, können wir Theorem 3.12 verwenden und erhalten $\sup_{m\in M} \|I(m)\| < \infty$. Aber da $I : X \to X''$ isometrisch ist, zeigt dies $\sup_{m\in M} \|m\| < \infty$, also die Beschränktheit von M.

Die Aussage über Folgen ist klar, da für eine schwach konvergente Folge (x_n) die Menge $\varphi(x_n)_n$ aufgrund der Konvergenz beschränkt ist. □

Natürlich können Sie noch einen Schritt weiter gehen und Folgen von Operatorn $A_n \in \mathcal{B}(X, Y)$ betrachten, für die $\varphi(A_nx)_n$ für alle $x \in X$ und alle $\varphi \in Y'$ konvergieren. Wir werden diese "schwachen Konvergenzbegriffe" und die zugrundeliegenden schwachen Topologien später systematisch betrachten.

3.3 Der Satz von der offenen Abbildung

Erinnern Sie sich, dass eine Abbildung $f : X \to Y$ zwischen zwei metrischen (oder topologischen) Räumen stetig ist genau dann, wenn Urbilder offener Mengen offen sind, dh wenn für jedes offene $O \subset Y$ auch $f^{-1}(O) \subset X$ offen ist. Bei einer offenen Abbildung geht es stattdessen um die Bilder:

Definition 3.20. Seien X, Y metrische Räume. Eine Abbildung $f : X \to Y$ heißt *offen*, falls sie offene Mengen auf offene Mengen abbildet.

Als kleine Warnung vermerken wir, dass Sie hier nicht offene durch abgeschlossene Mengen ersetzen dürfen: Zum Beispiel ist die Projektionsabbildung $p : \mathbb{R}^2 \to \mathbb{R}$, $p(x, y) := x$, offen, bildet aber die abgeschlossene Menge $\{(x, y) : x \geq 0, xy \geq 1\}$ auf die nicht abgeschlossene Menge $(0, \infty)$ ab.

Wieso sollten wir uns für offene Abbildungen interessieren? Nun, per Definition ist eine bijektive Abbildung genau dann offen, wenn ihre Umkehrabbildung stetig ist. Offene Abbildungen werden also zur Untersuchung der Stetigkeit von Umkehrabbildungen benutzt.

Im folgenden Lemma betrachten wir den für die Funktionalanalysis wichtigsten Fall von *linearen* Abbildungen, in dem sich die Offenheit etwas vereinfachen lässt. Wir schreiben wie üblich $U_r(x) = \{z \in X : \|z - x\| < r\}$ für offene Kugeln mit Mittelpunkt x und Radius r in einem normierten Raum. Als Abkürzung schreiben wir im Folgenden $U_r := U_r(0)$, und manchmal $U_r^X(x)$ um anzudeuten, dass die Kugel $U_r(x)$ in dem Raum X gemeint ist.

Lemma 3.21. *Seien X, Y normierte Räume und $A : X \to Y$ linear. Dann sind äquivalent:*

a) A ist offen.

b) Es gibt $r > 0$, so dass $A(U_r^X)$ innere Punkte hat.

c) Es gibt $\varepsilon > 0$, so dass $U_\varepsilon^Y \subset A(U_1^X)$.

Weiterhin ist jede lineare offene Abbildung surjektiv.

Beweis. $a) \Rightarrow b)$ ist trivial.

$b) \Rightarrow c)$ Wenn $A(U_r^X)$ den inneren Punkt y hat, so gibt es $\delta > 0$ mit $y + U_\delta^Y \subset A(U_r^X)$. Da die Kugeln punktsymmetrisch (invariant unter $x \mapsto -x$) sind und A linear ist, haben wir dann auch $-y + U_\delta^Y \subset A(U_r^X)$. Für einen Vektor $\tilde{y} \in Y$ mit $\|\tilde{y}\| < 2\delta$ gilt also

$$\tilde{y} = (y + \tfrac{1}{2}\tilde{y}) + (-y + \tfrac{1}{2}\tilde{y}) \subset A(O_r^X) + A(O_r^X) \subset 2A(O_r^X),$$

also $U_\delta^Y \subset A(O_r^X)$. Skalierung mit $\frac{1}{r}$ liefert nun $U_{\delta/r}^Y \subset A(U_1^X)$, dh c) gilt mit $\varepsilon = \frac{\delta}{r}$.

$c) \Rightarrow a)$ Sei $O \subset X$ offen und $x \in O$. Wir müssen zeigen, dass Ax ein innerer Punkt von $A(O)$ ist. Da O offen ist, gibt es $r > 0$ mit $U_r^X(x) \subset O$, und nach Skalierung mit r^{-1} entsprechend $U_1^X(r^{-1}x) \subset r^{-1}O$. Anwendung von A und c) liefern nun $\varepsilon > 0$ mit

$$r^{-1}Ax + U_\varepsilon^Y \subset r^{-1}Ax + A(U_1^X) \subset r^{-1}A(O),$$

also $Ax + U_{r\varepsilon}^Y \subset A(O)$. Also ist Ax ein innerer Punkt von $A(O)$, dh $A(O)$ ist offen.

Zur behaupteten Surjektivität offener linearer Abbildungen: Sei $A : X \to Y$ offen und linear. Dann enthält das Bild von A eine offene Kugel U_ε^Y für ein $\varepsilon > 0$ (siehe c)). Ist $y \in Y \backslash \{0\}$ beliebig, so liegt $\frac{\varepsilon y}{2\|y\|}$ in dieser Kugel und damit im Bild. Da A linear ist, liegt dann auch y im Bild von A, dh A ist surjektiv. $\square$

Beispiel 3.22.

a) Sei $\mathcal{H}$ ein Hilbertraum und $P_\mathcal{K}$ die orthogonale Projektion auf einen abgeschlossenen Unterraum $\mathcal{K} \subset \mathcal{H}$. Dann ist $Q_\mathcal{K} : \mathcal{H} \to \mathcal{K}$, $v \mapsto P_\mathcal{K} v$, offen.
Beweis: Wir müssen eine offene Kugel $U_\varepsilon^\mathcal{K}(0)$ in $\mathcal{K}$ mit Mittelpunkt 0 finden, so dass $U_\varepsilon^\mathcal{K}(0) \subset P_\mathcal{K}(U_1^\mathcal{H}(0))$. Gegeben einen Vektor $k \in \mathcal{K}$ mit $\|k\| < 1$, so hat der Vektor $v := k \oplus 0 \in \mathcal{K} \oplus \mathcal{K}^\perp$ ebenfalls Norm $\|v\| = \|k\| < 1$, und $P_\mathcal{K} v = k$. Also gilt $U_\varepsilon^\mathcal{K}(0) \subset P_\mathcal{K}(U_1^\mathcal{H}(0))$ mit $\varepsilon = 1$. $\square$

b) Die Abbildung $A : \ell^\infty_\mathbb{N} \to c_0$, $(x_n) \mapsto (\frac{1}{n} x_n)$, ist *nicht* offen.
Beweis: Wir betrachten $r = 1$, also $A(U_1(0)) = \{(x_n) : |x_n| < \frac{1}{n} \ \forall n\}$. Für jedes $\varepsilon > 0$ und jedes $m \in \mathbb{N}$ enthält die offene Kugel $U_\varepsilon^{c_0}(0)$ die Folge $x_k^{\varepsilon,m} := \delta_{k,m}\varepsilon$, $k \in \mathbb{N}$. Für $\varepsilon m \geq 1$ liegt diese Folge nicht in $A(U_1(0))$.

Im ersten Beispiel haben wir die auf $\mathcal{K}$ koeingeschränkte Projektion und nicht die Projektion selbst, also $P_\mathcal{K} : \mathcal{H} \to \mathcal{H}$ betrachtet. Letztere ist nicht surjektiv (für $\mathcal{K} \neq \{0\}$) und damit nicht offen.

Der folgende Satz charakterisiert lineare offene Abbildungen auf erstaunlich einfache Art und Weise.

Theorem 3.23 (Satz von der offenen Abbildung). *Seien X, Y Banachräume und $A \in \mathcal{B}(X, Y)$ eine surjektive stetige lineare Abbildung. Dann ist A offen.*

Beweis. Die Idee ist, den Satz von Baire zu verwenden. Da A surjektiv ist, gilt

$$Y = \bigcup_{n \in \mathbb{N}} A(U_n^X).$$

Nach dem Satz von Baire gibt es also $n \in \mathbb{N}$, so dass $\overline{A(U_n^X)}$ innere Punkte hat. Wüssten wir sogar, dass $A(U_n^X)$ (ohne Abschluss) innere Punkte hat, so liefert Lemma 3.21 b) sofort, dass A offen ist, wie behauptet.

Da $\overline{A(U_n^X)}$ innere Punkte hat, hat auch $\overline{A(U_1^X)}$ innere Punkte (Skalierung um $\frac{1}{n}$). Wie im Beweis von Lemma 3.21 folgt, dass es $\varepsilon > 0$ gibt mit $U_\varepsilon^Y \subset \overline{A(U_1^X)}$. Wir zeigen nun als wesentlichen Beweisschritt $\overline{A(U_1^X)} \subset A(U_2^X)$. Das impliziert dann Existenz innerer Punkte von $A(U_2^X)$ und beendet den Beweis.

Sei also $y \in \overline{A(U_1^X)}$. Wir müssen $x \in U_2^X$ finden, so dass $y = Ax$ gilt. Der Vektor x wird durch einen Grenzwert definiert werden.

Zuerst bemerken wir, dass $y \in \overline{A(U_1^X)}$ Grenzwert einer Folge der Form $(A\xi_k)_k$ mit $\|\xi_k\| < 1$ für alle k. Wir finden also einen Vektor $x_1 \in U_1^X$ (Auswahl eines der Folgeglieder ξ_k), so dass $\|y - Ax_1\| < \frac{\varepsilon}{2}$ mit dem ε von $U_\varepsilon^Y \subset \overline{A(U_1^X)}$. Also $y - x_1 \in U_{\varepsilon/2}^X \subset \overline{A(U_{1/2}^X)}$.

Nun wiederholen wir diese Konstruktion für $y - Ax_1$ anstelle von y und $\varepsilon/2$ anstelle von ε, dh wir wählen $x_2 \in U^X_{1/2}$ mit $y - Ax_1 - Ax_2 \in U^X_{\varepsilon/4}$. Wir gehen induktiv vor und wählen $x_n \in U^X_{2^{1-n}}$ mit

$$y - \sum_{j=1}^{n} Ax_j \in U^X_{2^{-n}}.$$

Aufgrund von $\|x_n\| < 2^{1-n}$ existiert $x := \sum_{n=1}^{\infty} x_n$ und hat Norm $\|x\| < \sum_{n=1}^{\infty} 2^{1-n} = 2$ (geometrische Reihe), also $x \in U^X_2$. Da A stetig ist, konvergiert auch $\sum_{n=1}^{\infty} Ax_n = A\sum_{n=1}^{\infty} x_n = Ax$. Der fragliche Punkt y erfüllt per Konstruktion $y = Ax$. Also gilt $y \in A(U^X_2)$. □

Eine sehr wichtige Anwendung des Satzes von der offenen Abbildung ist das folgende Korollar.

Korollar 3.24 (Satz vom stetigen Inversen). *Seien X, Y Banachräume und $A \in \mathcal{B}(X, Y)$ bijektiv. Dann $A^{-1} \in \mathcal{B}(Y, X)$, dh $A^{-1} : Y \to X$ ist beschränkt.*

Beweis. Die Abbildung $A : X \to Y$ ist nach Voraussetzung surjektiv, stetig, und linear, nach dem Satz von der offenen Abbildung also offen. Das ist äquivalent zur Stetigkeit von A^{-1}. □

In Satz 1.22 hatten wir bereits gesehen, dass die Beschränktheit der Umkehrabbildung eines bijektiven beschränkten Operators $A \in \mathcal{B}(X, Y)$ äquivalent ist zur Existenz einer Konstante $c > 0$, so dass $\|Ax\| \geq c\|x\|$ für alle $x \in X$ gilt. Diese Abschätzung gilt also automatisch, wenn A bijektiv und beschränkt ist, dh wenn $\|Ax\| \leq C\|x\|$ für ein geeignetes $C > 0$ für alle $x \in X$ gilt.

Anders formuliert: Wenn Sie eine Gleichung der Form $Ax = y$ mit beschränktem bijektiven A zwischen Banachräumen X, Y betrachten (dies könnte zB eine Integralgleichung sein), so hängt die Lösung $x = A^{-1}y$ stetig von den gegebenen Daten y ab.

Beachten Sie aber, dass es wichtig ist, dass X, Y Banachräume, also vollständig sind. Wir geben ein Beispiel, um diesen Punkt zu betonen.

Beispiel 3.25. Sei $X = Y = L^\infty(\mathbb{R})_c$ der Raum der L^∞-Funktionen mit kompaktem Träger, versehen mit der wesentlichen Supremumsnorm (*kein* Banachraum), und $A : X \to X$ die Abbildung

$$(Af)(x) = \frac{f(x)}{1 + x^2}, \qquad x \in \mathbb{R}.$$

Dies ist eine lineare bijektive Abbildung. Wegen $\|Af\|_{L^\infty} = \sup_x \left|\frac{f(x)}{1+x^2}\right| \leq \|f\|_{L^\infty}$, also $\|A\| \leq 1$, ist A stetig. Die Umkehrabbildung ist offenbar

$$(A^{-1}f)(x) = (1 + x^2)f(x), \qquad x \in \mathbb{R}.$$

Dieser Operator ist *nicht* beschränkt: Sei χ_n die charakteristische Funktion von $[n, n+1]$. Dann gilt

$$\|A^{-1}\chi_n\|_{L^\infty} = \operatorname{esssup}_{x\in\mathbb{R}}|(1 + x^2)\chi_n(x)| \geq (1 + n^2) = (1 + n^2)\|\chi_n\|_{L^\infty},$$

also $\|A^{-1}\| \geq (1+n^2)$ für alle $n \in \mathbb{N}$.

Wir fügen noch zwei weitere Korollare hinzu, die den Satz von der offenen Abbildung illustrieren.

Korollar 3.26. *Sei X ein $\mathbb{K}$-Vektorraum und $\|\cdot\|$ und $\|\cdot\|'$ zwei Normen auf X, so dass X bzgl jeder dieser zwei Normen ein Banachraum ist. Dann gilt*

$$(\exists C > 0 : \; \|x\| \leq C\|x\|' \; \forall x \in X) \;\Rightarrow\; (\exists C' > 0 : \; \|x\|' \leq C'\|x\| \; \forall x \in X) \, .$$

Das heißt: Wenn $\|\cdot\|$ durch $\|\cdot\|'$ abgeschätzt werden kann, dann kann auch automatisch $\|\cdot\|'$ durch $\|\cdot\|$ abgeschätzt werden. Beachten Sie aber die starke Voraussetzung, dass X bzgl beider Normen vollständig sein muss.

Beweis. Die Identität $\mathrm{id} : (X, \|\cdot\|') \to (X, \|\cdot\|)$ ist linear, bijektiv, und nach Voraussetzung stetig. Also ist die Umkehrabbildung $\mathrm{id} : (X, \|\cdot\|) \to (X, \|\cdot\|')$ auch stetig. $\square$

Injektivität ist meist eine einfacher zu prüfende Eigenschaft als Surjektivität, und wir haben es oft mit Operatoren $A : X \to Y$ zu tun, die wir als injektiv aber nicht notwendigerweise surjektiv erkennen. In diesem Fall können wir A auf sein Bild koeinschränken und die bijektive Abbildung $A : X \to \operatorname{Ran} A$ betrachten. Das foldende Korollar sagt uns, wann diese Abbildung ein stetiges Inverses hat.

Korollar 3.27. *Seien X, Y Banachräume und $A \in \mathcal{B}(X, Y)$ injektiv. Die koeingeschränkte Abbildung $A : X \to \operatorname{Ran} A$ hat genau dann ein stetiges Inverses $A^{-1} : \operatorname{Ran} A \to X$, wenn $\operatorname{Ran} A$ abgeschlossen ist.*

Beweis. Falls $\operatorname{Ran} A$ abgeschlossen ist, so ist es ein Banachraum und der Satz vom stetigen Inversen sagt uns, dass $A^{-1} : \operatorname{Ran} A \to X$ stetig ist.

Ist andererseits $A^{-1} : \operatorname{Ran} A \to X$ stetig, so ist es als stetige bijektive Abbildung ein Isomorphismus zwischen den normierten Räumen $\operatorname{Ran} A$ und X. Da X vollständig ist, ist dann auch $\operatorname{Ran} A$ vollständig (Lemma 1.25), also ist $\operatorname{Ran} A$ abgeschlossen (Lemma 1.16). $\square$

Das folgende einfache Beispiel illustriert dieses Ergebnis.

Beispiel 3.28. Sei $X = Y = L^2(\mathbb{R})$ und $A \in \mathcal{B}(L^2(\mathbb{R}))$ der beschränkte Multiplikationsoperator $(Af)(x) = (1+x^2)^{-1} f(x)$. Dann ist A injektiv, denn $f \in \ker A$ erfüllt $(1+x^2)^{-1} f(x) = 0$ für fast alle $x \in \mathbb{R}$, also $f = 0$ in $L^2(\mathbb{R})$.

Wir behaupten, dass A kein abgeschlossenes Bild hat. Dazu bemerken wir, dass jede L^2-Funktion f mit kompaktem Träger in $\operatorname{Ran} A$ liegt, denn $f(x) = (1+x^2)^{-1} \cdot (1+x^2) f(x)$, und $x \mapsto (1+x^2) f(x)$ liegt in L^2 wegen dem kompakten Träger. Auf diesen Funktionen ist $A^{-1} : \operatorname{Ran} A \to L^2(\mathbb{R})$ offenbar durch $(A^{-1} f)(x) = (1+x^2) f(x)$ gegeben. Aber für $f = \chi_{[n,n+1]}$ sehen wir

$$\frac{\|A^{-1}\chi_{[n,n+1]}\|^2}{\|\chi_{[n,n+1]}\|^2} = \frac{\int_n^{n+1} (1+x^2)^2 dx}{\int_n^{n+1} dx} = \int_n^{n+1} (1+x^2)^2 dx \to \infty$$

für $n \to \infty$. Also ist $A^{-1} : \operatorname{Ran} A \to L^2$ nicht beschränkt, dh $\operatorname{Ran} A$ ist nicht abgeschlossen.

3.4 Der Satz vom abgeschlossenen Graphen

Das Konzept eines Graphen haben Sie bereits in der Analysis 1 kennengelernt: Sind N, M Mengen und $f : N \to M$ eine Abbildung, so ist

$$\Gamma(f) = \{(x, f(x)) : \ x \in N\} \subset N \times M$$

der Graph von f. Es lässt sich insbesondere auf lineare Abbildungen zwischen normierten Räumen anwenden. Wir wählen eine Formulierung, die für viele Anwendungen geeignet ist.

Definition 3.29. Seien X, Y Banachräume, $D \subset X$ ein Untervektorraum, und $A : D \to Y$ linear. Der *Graph* von A ist die Menge

$$\Gamma(A) := \{(x, Ax) : \ x \in D\} \subset X \oplus Y.$$

Hierbei steht "D" für Definitionsbereich (von A), man schreibt oft auch $D = \operatorname{dom} A$. Sehr oft ist nämlich ein Operator gegeben, der nur auf einem nicht vollständigen Untervektorraum eines Banachraums definiert ist, zB

$$P : \mathscr{S}(\mathbb{R}) \subset L^2(\mathbb{R}) \to L^2(\mathbb{R}), \qquad Pf := -if',$$

der in der Quantenmechanik wichtige Impulsoperator. Wir betrachten D aber stets als eine Teilmenge eines Banachraums X, den wir insbesondere als Vervollständigung von D wählen könnten. Auf diese Weise können wir den Graphen $\Gamma(A)$ als Teilmenge des Banachraums $X \oplus Y$ (siehe Def. 1.26 für die Norm auf diesem Raum) auffassen.

Beachten Sie, dass $\Gamma(A) \subset X \oplus Y$ ein Untervektorraum ist. Wir untersuchen nun, unter welchen Bedingungen dieser Unterraum abgeschlossen ist.

Lemma 3.30. *Sei $A : D \subset X \to Y$ ein auf einem Unterraum D eines Banachraums X definierter Operator in einen Banachraum Y. Dann ist $\Gamma(A)$ genau dann abgeschlossen, wenn gilt: Konvergiert eine Folge $(x_n)_n \in D$ gegen einen Punkt $x \in X$* und *konvergiert die Bildfolge $(Ax_n)_n$ gegen einen Punkt $y \in Y$, so folgt $x \in D$ und $Ax = y$.*

Beweis. Abgeschlossenheit von $\Gamma(A)$ in $X \oplus Y$ bedeutet: Konvergiert eine Folge $(x_n, Ax_n)_n \in \Gamma(A)$ gegen einen Punkt (x, y) (das heißt, konvergiert $x_n \to x$ und $Ax_n \to y$), so gilt $(x, y) \in \Gamma(A)$ (das heißt, $x \in D$ und $y = Ax$). □

Definition 3.31. Eine lineare Abbildung $A : D \subset X \to Y$ von einem Unterraum D eines Banachraums X in einen Banachraum Y heißt *abgeschlossen*, wenn $\Gamma(A) \subset X \oplus Y$ abgeschlossen ist.

Um diesen Begriff näher zu erläutern, betrachten wir den Fall $D = X$ und die drei Aussagen

a) $x_n \to x$,

b) $Ax_n \to y$,

c) $Ax = y$.

Falls A *stetig* ist, gilt: $a) \Rightarrow b), c)$. Falls A nur *abgeschlossen* ist, gilt $a), b) \Rightarrow c)$. In diesem Sinne sind abgeschlossene Operatoren also eine Verallgemeinerung von stetigen Operatoren.

Beispiel 3.32.

a) Wir betrachten den Hilbertraum $\ell^2_{\mathbb{N}}$, den dichten Unterraum

$$D := \{x \in \ell^2_{\mathbb{N}} : \sum_n n^2 |x_n|^2 < \infty\},$$

und den Operator $A : D \to \ell^2_{\mathbb{N}}$, $(Ax)_n := nx_n$. Wir behaupten, dass A abgeschlossen ist.

Sei dazu $(x(k))_{k\in\mathbb{N}}$ eine Folge in D und $v, w \in \ell^2_{\mathbb{N}}$ mit

$$\|x(k) - v\|^2 = \sum_{n=1}^{\infty} |x(k)_n - v_n|^2 \to 0,$$
$$\|Ax(k) - w\|^2 = \sum_{n=1}^{\infty} |nx(k)_n - w_n|^2 \to 0.$$

Das impliziert für alle $n \in \mathbb{N}$ die Limiten $|x(k)_n - v_n|^2 \to 0$ und $|nx(k)_n - w_n|^2 \to 0$ für $k \to \infty$, also $w_n = \lim_k nx(k)_n = nv_n$, und insbesondere $v \in D$. Also ist A abgeschlossen. Es ist auch klar, dass A nicht beschränkt ist.

b) Als eine kontinuierliche Variante betrachten wir $D := \mathscr{S}(\mathbb{R})$ in $\mathcal{H} := L^2(\mathbb{R})$ und den Impulsoperator $P : \mathscr{S}(\mathbb{R}) \to L^2(\mathbb{R})$, $(Pf)(x) = -if'(x)$.

Sei $(f_n) \subset \mathscr{S}(\mathbb{R})$ eine Folge mit $\|f_n - f\|_{L^2} \to 0$ und $\|Pf_n - g\|_{L^2} \to 0$ für $f, g \in L^2$. Wir verwenden die Fouriertransformation $f \mapsto \tilde{f}$ (ein unitärer Operator auf L^2) und erhalten mittels der bekannten Beziehung $\widetilde{Pf}(k) = k\tilde{f}(k)$

$$\int_{\mathbb{R}} |\tilde{f}_n(k) - \tilde{f}(k)|^2 dk \to 0,$$
$$\int_{\mathbb{R}} |k\tilde{f}_n(k) - \tilde{g}(k)|^2 dk \to 0.$$

Diese L^2-Konvergenzen implizieren punktweise Konvergenzen fast überall nach Übergang zu Teilfolgen (Analysis 3, Satz 2.35). Nach Übergang zu einer Teilfolge gilt also $\tilde{f}_n(k) \to \tilde{f}(k)$ fast überall und $k\tilde{f}_n(k) \to \tilde{g}(k)$ fast überall, dh $k\tilde{f}(k) = \tilde{g}(k)$ im Sinne von L^2.

In diesem Beispiel muss allerdings nicht $f \in \mathscr{S}(\mathbb{R}) = D$ gelten. Wir haben es also nicht ganz mit einem abgeschlossenen Operator zu tun, sondern nur mit einem abschließbaren Operator - dies sind Begrifflichkeiten, die wir erst in Funktionalanalysis 2 beim Studium von unbeschränkten Operatoren brauchen werden.

Lemma 3.33. *Seien X, Y Banachräume, $D \subset X$ ein Unterraum, und $A : D \to Y$ ein abgeschlossener Operator. Dann ist D ein Banachraum bzgl. der Norm*

$$\|x\|_A := \|x\| + \|Ax\|, \qquad x \in D,$$

und A ist als Abbildung $A : (D, \|\cdot\|_A) \to Y$ stetig.

Beweis. Sei $(x_n) \subset D$ eine $\|\cdot\|_A$-Cauchyfolge, dh $\|x_n - x_m\| + \|Ax_n - Ax_m\| \to 0$ für $n, m \to \infty$. Also sind $(x_n) \subset X$ und $(Ax_n)_n \subset Y$ Cauchyfolgen. Da X, Y Banachräume sind, konvergieren diese Folgen, $x_n \to x$ und $Ax_n \to y$. Aufgrund der Abgeschlossenheit von A folgt $x \in D$ und $Ax = y$. Das bedeutet $\|x_n - x\|_A \to 0$, dh $(D, \|\cdot\|_A)$ ein Banachraum.

Die Stetigkeit ist klar wegen $\|x\| \leq \|x\| + \|Ax\| = \|x\|_A$. □

Theorem 3.34 (Satz vom abgeschlossenen Graphen). *Seien X, Y Banachräume und $A : X \to Y$ linear. Dann ist A genau dann beschränkt, wenn A abgeschlossen ist.*

Beweis. Die Implikation stetig ⇒ abgeschlossen ist klar.

Für die andere Richtung erinnern wir uns, dass $\Gamma(A) \subset X \oplus Y$ ein Banachraum ist, und betrachten die beiden kanonischen Projektionen

$$\begin{aligned} \pi_1 : \Gamma(A) \to X, \qquad & (x, Ax) \mapsto x, \\ \pi_2 : \Gamma(A) \to X, \qquad & (x, Ax) \mapsto Ax. \end{aligned}$$

π_1 ist bijektiv, also ist π_1^{-1} stetig. Damit ist auch $A = \pi_2 \pi_1^{-1}$ stetig. □

Eine interessante Konsequenz des Satzes vom abgeschlossenen Graphen bezieht sich auf Hilbertraumoperatoren.

Korollar 3.35 (Satz von Hellinger-Töplitz). *Sei $\mathcal{H}$ ein Hilbertraum und $A : \mathcal{H} \to \mathcal{H}$ linear mit der Eigenschaft $\langle x, Ay \rangle = \langle Ax, y \rangle$ für alle $x, y \in \mathcal{H}$. Dann ist A beschränkt.*

Beweis. Nach dem Satz vom abgeschlossenen Graphen genügt es, zu zeigen, dass A abgeschlossen ist. Sei also $(x_n) \subset \mathcal{H}$ eine Folge und $x, y \in \mathcal{H}$ so, dass $x_n \to x$ und $Ax_n \to y$. Dann gilt für alle $z \in \mathcal{H}$

$$\langle z, y \rangle = \lim_{n\to\infty} \langle z, Ax_n \rangle = \lim_{n\to\infty} \langle Az, x_n \rangle = \langle Az, x \rangle = \langle z, Ax \rangle.$$

Also gilt $\langle z, y - Ax \rangle = 0$ für alle $z \in \mathcal{H}$ und deshalb $Ax = y$. Also ist A abgeschlossen. □

Natürlich erfüllen selbstadjungierte Operatoren $A = A^*$ die in diesem Korollar gemachte Symmetriebedingung $\langle x, Ay \rangle = \langle Ax, y \rangle$. Aber der Punkt bei dem Korollar ist, dass A a priori nicht als beschränkt angenommen wird, und wir für unbeschränkte Operatoren noch keinen Begriff eines Adjungierten haben.

In der Quantenmechanik werden physikalische Observable (Ort, Impuls, Energie, Drehimpuls, ...) durch Operatoren beschrieben, die $\langle x, Ay \rangle = \langle Ax, y \rangle$ erfüllen (für x, y in einem dichten Unterraum), zumeist aber unbeschränkt sind, was den unbeschränkten Werten für Ort, Impuls, etc entspricht. Der Satz von Hellinger-Töplitz besagt in diesem Zusammenhang, dass solche Operatoren nicht auf dem ganzen Hilbertraum definiert sein können. Wir werden deshalb später (in Funktionalanalysis 2) Operatoren studieren, die nur auf dichten Unterräumen von Hilberträumen definiert sind.

4 Spektraltheorie beschränkter Operatoren

In diesem Kapitel werden wir die Eigenwerttheorie der Linearen Algebra auf Banach- und insbesondere Hilbertraumoperatoren verallgemeinern. Dies beinhaltet auch sogenannte Funktionalkalküle, dh das Bilden von Funktionen von Operatoren.

4.1 Das Spektrum

Die aus der Linearen Algebra bekannte Definition eines *Eigenwerts/Eigenvektors* überträgt sich problemlos auf unendlichdimensionale Räume: Ist $A : X \to X$ eine lineare Abbildung auf einem beliebigen $\mathbb{K}$-Vektorraum X, so ist ein Eigenvektor ein Vektor $x \in X\backslash\{0\}$, so dass $Ax = \lambda \cdot x$ für eine Zahl $\lambda \in \mathbb{K}$, die dann Eigenwert genannt wird.

Im Endlichdimensionalen wissen Sie, dass jedes $A : X \to X$ Eigenwerte hat, wenn X komplex ist. Im reellen Fall hingegen muss es keine Eigenwerte geben. Im Unendlichdimensionalen kann es auch aber dann keine Eigenwerte geben, wenn wir über $\mathbb{C}$ arbeiten, wie das folgende Beispiel zeigt.

Beispiel 4.1. Wir betrachten auf dem komplexen Banachraum $X := C([0,1])$ die lineare Abbildung[a]

$$\begin{aligned} A : C([0,1]) &\to C([0,1]), \\ (Af)(x) &= x \cdot f(x) \end{aligned}$$

und bemerken, dass A *keine Eigenwerte* hat. Gilt nämlich $Af = \lambda \cdot f$ für ein $f \in C([0,1])$, $f \neq 0$, und $\lambda \in \mathbb{C}$, dann folgt

$$xf(x) = \lambda f(x) \qquad \forall x \in [0,1]$$

und damit $(x-\lambda)f(x) = 0$. Somit muss $f(x)$ für alle $x \neq \lambda$ verschwinden. Da f stetig ist, erhalten wir als einzige Lösung $f = 0$, was per Definition kein Eigenvektor ist.

[a] In der Quantenmechanik heißen dieser Operator und seine Varianten *Ortsoperator*.

Wir führen nun das Spektrum als Verallgemeinerung von Eigenwerten ein. Wir werden später zeigen, dass das Spektrum im komplexen Fall nie leer ist.

Die Idee ist die Folgende: Hat $A : X \to X$ im Endlichdimensionalen einen Eigenwert λ, so gilt $\ker(A - \lambda) \neq \{0\}$ (hier und im Folgenden heißt $A - \lambda$ immer $A - \lambda \mathrm{id}_X$), denn jeder Eigenvektor zu λ liegt in $\ker(A - \lambda)$. Also ist $(A - \lambda)$ in diesem Fall nicht invertierbar. Ist hingegen $(A - \lambda)$ nicht invertierbar, so ist $A - \lambda$ nicht bijektiv und aufgrund der Dimensionsformel (im Endlichdimensionalen!) nicht injektiv, dh λ ist ein Eigenwert.

Im Endlichdimensionalen können Eigenwerte also durch Invertierbarkeitseigenschaften ausdrücken, die wir im allgemeinen Fall zur Definition erheben werden.

Ist $A : X \to X$ ein beschränkter Operator auf einem Banachraum X, so nennen wir A invertierbar, wenn er bijektiv ist. Aus Korollar 3.24 wissen wir, dass die inverse Abbildung $(A - \lambda)^{-1}$ dann auch automatisch beschränkt ist.

Definition 4.2. Sei X ein Banachraum und $A \in \mathcal{B}(X)$. Die *Resolventenmenge* $\rho(A)$ von A ist

$$\rho(A) := \{\lambda \in \mathbb{C} : (A - \lambda) \text{ invertierbar}\}$$

und das *Spektrum* $\sigma(A)$ von A ist

$$\sigma(A) := \mathbb{C} \backslash \rho(A).$$

Die *Resolvente* (oder *Resolventenabbildung*) ist

$$R_A : \rho(A) \to \mathcal{B}(X), \qquad R_A(\lambda) := (A - \lambda)^{-1}.$$

Diese Definition ist eigentlich ganz algebraisch und kann deshalb genauso für Elemente von beliebigen unitalen Algebren getroffen werden.

Aufgrund der Definition ist klar:

- Alle Eigenwerte von A liegen in $\sigma(A)$.
- Für $\dim(X) < \infty$ ist $\sigma(A)$ die Menge aller Eigenwerte von A.

Im Allgemeinen ist das Spektrum größer als die Menge der Eigenwerte:

Beispiel 4.3. In diesem Beispiel betrachten wir noch einmal den eigenwertfreien Operator aus Beispiel 4.1 und zeigen, dass sein Spektrum keineswegs leer ist.

Sei also wieder $X = C([0, 1])$ und $A : X \to X$, $(Af)(x) = xf(x)$. Dann behaupten wir $\sigma(A) = [0, 1]$.

Zum Beweis sei $\lambda \in \mathbb{C} \backslash [0, 1]$. Dann ist

$$B_\lambda : C([0, 1]) \to C([0, 1]), \qquad (B_\lambda f)(x) = \frac{f(x)}{x - \lambda}$$

wohldefiniert und beschränkt wegen $\sup_{x \in [0,1]} |x - \lambda|^{-1} = d(\lambda, [0, 1])^{-1} < \infty$. Offenbar gilt $(A - \lambda)B_\lambda = B_\lambda(A - \lambda) = \text{id}$, dh $\mathbb{C} \backslash [0, 1] \subset \rho(A)$.

Sei nun $\lambda \in [0, 1]$. Dann hat für jedes $f \in C([0, 1])$ die Funktion $(A - \lambda)f : x \mapsto (x - \lambda)f(x)$ bei $x = \lambda$ eine Nullstelle. Also ist $A - \lambda$ nicht surjektiv und somit nicht invertierbar, d.h. $[0, 1] \subset \sigma(A)$.

Wir werden jetzt die grundlegenden Eigenschaften des Spektrums ableiten und zeigen, dass $\sigma(A)$ (für $A \in \mathcal{B}(X)$) immer eine abgeschlossene und beschränkte (kompakte) und im komplexen Fall nicht leere Teilmenge von $\mathbb{C}$ ist.

Zur Vorbereitung erinnern wir zunächst an Hausaufgabe H3.2. Dort wurde gezeigt: Hat ein Operator $A \in \mathcal{B}(X)$ Norm $\|A\| < 1$, so ist $(1 - A)$ invertierbar, und es gilt

$$(1 - A)^{-1} = \sum_{n=0}^{\infty} A^n,$$

(mit $A^0 := 1$), wobei die Reihe in Operatornorm konvergiert. Diese geometrische Reihe von Operatoren wird auch Neumann-Reihe genannt.

Im folgenden Beweis werden wir auch etwas Funktionentheorie benötigen. Da diese Vorlesung nicht vorausgesetzt wird, hier einige Kommentare.

- Eine auf einer offenen Menge $U \subset \mathbb{C}$ definierte Funktion $f : U \to \mathbb{C}$ heißt *holomorph*, wenn sie lokal durch eine konvergente Potenzreihe dargestellt wird[6]. Das heißt: es gibt für jeden Punkt $z_0 \in U$ einen Radius $r > 0$ mit $U_r(z_0) \subset U$ und f hat auf der Kreisscheibe $U_r(z_0)$ eine Potenzreihendarstellung hat, dh $f(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n$ mit Koeffizienten $a_n \in \mathbb{C}$ und die Reihe konvergiert (absolut). Insbesondere sind holomorphe Funktionen stetig.
- Man zeigt dann: Ist $f : U \to \mathbb{C}$ holomorph und $f(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n$ die Potenzreihendarstellung von f um $z_0 \in U$, so erfüllt der Konvergenzradius dieser Potenzreihe $r \geq d(z_0, \partial U)$ (Abstand von z_0 zum Rand von U).
- Der *Satz von Liouville* besagt: Jede auf ganz $\mathbb{C}$ holomorphe Funktion, die beschränkt ist (dh es gibt $C > 0$, so dass $|f(z)| \leq C$ für alle $z \in \mathbb{C}$), ist konstant.

  Anders formuliert: Ist $\sum_{n=0}^{\infty} a_n z^n$ eine für alle $z \in \mathbb{C}$ konvergente Potenzreihe, die auf ganz $\mathbb{C}$ beschränkt ist, so gilt $a_n = 0$ für alle $n > 0$.

Theorem 4.4. *Sei X ein $\mathbb{K}$-Banachraum und $A \in (X)$.*

a) $\sigma(A) \subset \mathbb{C}$ ist kompakt, dh abgeschlossen und beschränkt.

b) Ist $\mathbb{K} = \mathbb{C}$, so ist $\sigma(A)$ nicht leer.

Beweis. a) Wir zeigen zuerst, dass $\sigma(A)$ abgeschlossen ist, indem wir zeigen, dass die Resolventenmenge offen ist. Sei also $\lambda_0 \in \rho(A)$. Für $\lambda \in \mathbb{C}$ schreiben wir

$$A - \lambda = (A - \lambda_0) + (\lambda_0 - \lambda) = (A - \lambda_0)\Big(1 - (\lambda - \lambda_0)\,(A - \lambda_0)^{-1}\Big).$$

Wir wählen nun λ so nah an λ_0, dass $|\lambda_0 - \lambda| < \|(A - \lambda_0)^{-1}\|^{-1}$ gilt. Dann hat $B := (\lambda - \lambda_0)\,(A - \lambda_0)^{-1}$ Norm $\|B\| < 1$. Also konvergiert die Neumannreihe $\sum_{n=0}^{\infty} B^n = (1 - B)^{-1}$, insbesondere ist $1 - B$ invertierbar. Da auch $(A - \lambda_0)$ invertierbar ist, ist auch $A - \lambda = (A - \lambda_0)(1 - B)$ invertierbar, dh $\lambda \in \rho(A)$. Für die Resolvente erhalten wir die Gleichung

$$R_A(\lambda) = \sum_{n=0}^{\infty} (\lambda - \lambda_0)^n R_A(\lambda_0)^{n+1}, \qquad |\lambda_0 - \lambda| < \|(A - \lambda_0)^{-1}\|^{-1}.$$

Für den Beweis der Beschränktheit von $\sigma(A)$ wählen wir $\lambda \in \mathbb{C}$ mit $|\lambda| > \|A\|$ und behaupten $\lambda \in \rho(A)$. Um dies zu zeigen, benutzen wir noch einmal die Neumannreihe:

$$A - \lambda = -\lambda\left(1 - \frac{A}{\lambda}\right).$$

[6] Dies ist, nebenbei bemerkt, äquivalent zur komplexen Differenzierbarkeit von f.

Wegen $\|\frac{A}{\lambda}\| = \frac{\|A\|}{|\lambda|} < 1$ ist $\left(1 - \frac{A}{\lambda}\right)$ invertierbar. Da $\lambda \neq 0$, ist auch $A - \lambda$ invertierbar, dh $\lambda \in \rho(A)$.

b) Wir führen einen Widerspruchsbeweis, nehmen also $\sigma(A) = \emptyset$ an. Dann ist die Resolventenabbildung $R_A(\lambda) = (A - \lambda)^{-1}$ für alle $\lambda \in \mathbb{C}$ definiert. Im Beweis von a) haben wir gelernt, dass wir in einer kleinen Umgebung von λ_0 die Potenzreihendarstellung

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

haben. Sei $\varphi \in \mathcal{B}(X)'$ ein beliebiges stetiges Funktional auf dem Banachraum $\mathcal{B}(X)$. Dann ist die Funktion $f : \mathbb{C} \to \mathbb{C}$, $f(\lambda) := \varphi(R_A(\lambda))$ holomorph, da sie lokal die Potenzreihendarstellung $f(\lambda) = \sum_{n=0}^{\infty} (\lambda - \lambda_0)^n \varphi(R_A(\lambda_0)^{n+1})$ hat. Weiterhin ist f beschränkt. Um dies einzusehen, betrachten wir zunächst $|\lambda| > 2\|A\|$. Wegen $(A - \lambda)^{-1} = -\lambda^{-1}(1 - \frac{A}{\lambda})^{-1} = -\lambda^{-1} \sum_{n=0}^{\infty} (A/\lambda)^n$ gilt

$$\begin{aligned} |f(\lambda)| = |\varphi((A - \lambda)^{-1})| &\leq |\lambda|^{-1} \left| \varphi\left(\sum_{n=0}^{\infty} \frac{A^n}{\lambda^n} \right) \right| \\ &\leq |\lambda|^{-1} \|\varphi\| \sum_{n=0}^{\infty} \frac{\|A\|^n}{|\lambda|^n} \leq \frac{1}{2\|A\|} \|\varphi\| \sum_{n=0}^{\infty} \frac{1}{2^n} = \frac{\|\varphi\|}{2\|A\|}. \end{aligned}$$

Auf $G := \{\lambda \in \mathbb{C} : |\lambda| > 2\|A\|\}$ ist f also beschränkt. Auf der kompakten Kreisscheibe $\mathbb{C} \backslash G$ ist f als holomorphe (insbesondere stetige) Funktion aber ohnehin beschränkt. Also ist f holomorph und beschränkt, nach dem Satz von Liouville also konstant.

Dies führt nun zu einem Widerspruch. Wir betrachten die Potenzreihenentwicklung von f um 0, nämlich

$$f(\lambda) = \sum_{n=0}^{\infty} \lambda^n \varphi(R_A(0)^{n+1}).$$

Da diese Funktion konstant ist, verschwinden alle Terme mit $n > 0$. Für $n = 1$ besagt dies insbesondere $0 = \varphi(R_A(0)^2) = \varphi(A^{-2})$. Da $\varphi \in \mathcal{B}(X)'$ beliebig war, folgt mit Korollar 3.7 $A^{-2} = 0$. Aber das Inverse von A^2 kann natürlich nicht 0 sein. Also haben wir einen Widerspruch erhalten, dh $\sigma(A)$ ist nicht leer. □

Zur genaueren Beschreibung des Spektrums definieren wir noch den Spektralradius als die maximale radiale Ausdehnung des Spektrums.

> **Definition 4.5.** Der *Spektralradius* r_A von $A \in \mathcal{B}(X)$ ist
>
> $$r_A := \sup_{\lambda \in \sigma(A)} |\lambda| .$$

Im Laufe des Beweises von Satz 4.4 haben wir insbesondere gelernt, dass das Spektrum von A in der abgeschlossenen Kreisscheibe mit Mittelpunkt 0 und Radius $\|A\|$ enthalten ist, dh dass

$$r_A \leq \|A\|$$

gilt. Dies ist im Allgemeinen eine echte Ungleichung. Betrachten Sie zum Beispiel auf $X = \mathbb{C}^2$ den durch die Matrix

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

gegebenen Operator. Dann gilt $\sigma(A) = \{0\}$ (die Menge der Eigenwerte von A), also $r_A = 0$. Aber $\|A\| > 0$, da $A \neq 0$.

Wir geben noch ein weiteres (unendlichdimensionales) Beispiel.

Beispiel 4.6. Auf dem Banachraum $C([0,1])$ betrachten wir den linearen Operator $A : C([0,1]) \to C([0,1])$,

$$(Af)(x) := \int_0^x f(t)dt.$$

Mittels der Abschätzung $|(Af)(x)| \leq |x|\|f\|_\infty \leq \|f\|_\infty$ sehen wir, dass A beschränkt ist mit Norm $\|A\| \leq 1$. Betrachten wir die konstante Funktion $f(t) = 1$ (mit $\|f\|_\infty = 1$), so folgt $(Af)(x) = x$, also $\|Af\|_\infty = 1$. Daraus folgt $\|A\| = 1$. Wir behaupten

$$\sigma(A) = \{0\}.$$

Zuerst bemerken wir, dass $(Af)(0) = 0$ für alle $f \in C([0,1])$ gilt. Also ist A nicht surjektiv und damit $0 \in \sigma(A)$.

Um zu zeigen, dass jedes $\lambda \in \mathbb{C}\backslash\{0\}$ nicht im Spektrum von A liegt, zeigen wir zuerst, dass A keine Eigenwerte hat. Die Eigenwertgleichung $(Af)(x) = \int_0^x f(t)dt = \lambda f(x)$ impliziert für $\lambda \neq 0$, dass f stetig differenzierbar ist, und Differentiation liefert $f(x) = \lambda f'(x)$ bzw. $f(x) = 0$ im Falle $\lambda = 0$. Im letzteren Falle haben wir $f = 0$, also ist $\lambda = 0$ kein Eigenwert. Im Falle $\lambda \neq 0$ erhalten wir $f(x) = c\, e^{x/\lambda}$ mit einer Konstanten c. Aber $\lambda f(0) = (Af)(0) = 0$ impliziert dann $c = 0$, also ist auch $\lambda \neq 0$ kein Eigenwert.

Wir haben also gezeigt, dass $A - \lambda$ stets injektiv ist. Um zu zeigen, dass $\lambda \neq 0$ nicht im Spektrum von A liegt, zeigen wir nun auch noch, dass $A - \lambda$ surjektiv ist, dh dass $Af - \lambda f = g$ für beliebiges $g \in C([0,1])$ eine Lösung $f \in C([0,1])$ besitzt. Dazu betrachten wir die Gleichung

$$Af - \lambda f = g \;\Leftrightarrow\; \int_0^x f(t)dt - \lambda f(x) = g(x).$$

Wer sich mit Differentialgleichungen (Methode: Variation der Konstanten) auskennt, kann diese Gleichung lösen. Da wir dies hier nicht voraussetzen, geben wir die Lösung an:

$$f(x) = -\frac{1}{\lambda^2}\int_0^x e^{(x-t)/\lambda} g(t)dt - \frac{1}{\lambda} g(x).$$

Es lässt sich dann direkt nachprüfen, dass mit diesem f tatsächlich $(A-\lambda)f = g$ gilt. Also ist $A-\lambda$ surjektiv und damit $\lambda \neq 0$ nicht im Spektrum.

Auch hier gilt also eine echte Ungleichung $0 = r_A < \|A\| = 1$.

Wir werden bald eine exakte Formel für den Spektralradius herleiten. Vorher betrachten wir eine Zerlegung des Spektrums in unterschiedliche Teile.

Definition 4.7. Sei $A \in \mathcal{B}(X)$. Dann definieren wir

a) $\sigma_p(A) := \{\lambda \in \mathbb{C} : \; A-\lambda \text{ nicht injektiv}\}$ das *Punktspektrum* von A,

b) $\sigma_c(A) := \{\lambda \in \mathbb{C} : \; A-\lambda \text{ injektiv, nicht surjektiv, mit dichtem Bild}\}$ das *kontinuierliche Spektrum* von A,

c) $\sigma_r(A) := \{\lambda \in \mathbb{C} : \; A-\lambda \text{ injektiv, nicht dichtes Bild}\}$ das *Residualspektrum* von A.

Es ist klar, dass $\sigma_p(A)$, $\sigma_c(A)$, und $\sigma_r(A)$ disjunkte Teilmengen von $\sigma(A)$ sind, und dass

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

gilt. Das Punktspektrum besteht genau aus den Eigenwerten von A. Um die anderen Teile des Spektrums besser zu verstehen, führen wir verallgemeinerte Eigenwerte ein.

Definition 4.8. Ein *verallgemeinerter Eigenwert* von $A \in \mathcal{B}(X)$ ist eine Zahl $\lambda \in \mathbb{C}$, so dass eine Folge $(v_n)_{n\in\mathbb{N}}$ von Vektoren in $X\backslash\{0\}$ existiert, die

$$\lim_{n\to\infty} \frac{\|(A-\lambda)v_n\|}{\|v_n\|} = 0$$

erfüllt. Die Menge aller verallgemeinerten Eigenwerte wird auch das *approximative Punktspektrum* von A genannt und mit $\sigma_{ap}(A)$ bezeichnet.

Die Bezeichnung "verallgemeinerte Eigenwerte" ist dadurch motiviert, dass die Eigenwertgleichung $Av = \lambda v$ im Sinne obiger Definition immerhin "approximativ gilt". Die Beziehung zum Spektrum von A klären wir in den folgenden Ergebnissen.

Lemma 4.9. *Sei $A \in \mathcal{B}(X)$.*

a) $\lambda \notin \sigma_{ap}(A) \iff (A-\lambda)$ ist injektiv und $\operatorname{Ran}(A-\lambda)$ ist abgeschlossen.

b) Jeder verallgemeinerte Eigenwert von A liegt im Spektrum von A.

c) Jedes $\lambda \in \sigma_p(A) \cup \sigma_c(A)$ ist ein verallgemeinerter Eigenwert.

Zusammengefasst: $\sigma_p(A) \cup \sigma_c(A) \subset \sigma_{ap}(A) \subset \sigma(A)$.

Beweis. a) Wir bemerken zunächst, dass für eine komplexe Zahl λ die Aussage $\lambda \notin \sigma_{ap}(A)$ äquivalent ist zu: Es gibt ein $c > 0$, so dass

$$\|(A-\lambda)v\| \geq c\|v\| \qquad \forall v \in X. \tag{$\star$}$$

Nun zeigen wir die Äquivalenz von $(\star)$ und ($(A-\lambda)$ injektiv und $\operatorname{Ran}(A-\lambda)$ abgeschlossen): Gilt $(\star)$, so ist $(A-\lambda)$ sicherlich injektiv, da jeder Vektor $v \in \ker(A-\lambda)$ die Ungleichung $c\|v\| \leq \|(A-\lambda)v\| = 0$ erfüllt, also $v = 0$ ist.

Wir müssen zeigen, dass $(A-\lambda)$ abgeschlossenes Bild hat. Sei dazu $w \in \overline{\operatorname{Ran}(A-\lambda)}$ ein Vektor im Abschluss des Bildes und (v_n) eine Folge in X, die $(A-\lambda)v_n \to w$ erfüllt. Dann gilt wegen $(\star)$ auch

$$\|v_n - v_m\| \leq c^{-1}\|(A-\lambda)v_n - (A-\lambda)v_m\|.$$

Aber $(A-\lambda)v_n$ konvergiert, dh die rechte Seite geht gegen Null für $n, m \to \infty$. Also ist (v_n) eine Cauchyfolge, die wegen der Vollständigkeit von X einen Grenzwert $v \in X$ hat. Da $(A-\lambda)$ beschränkt ist, folgt

$$w = \lim_{n\to\infty}(A-\lambda)v_n = (A-\lambda)\lim_{n\to\infty} v_n = (A-\lambda)v.$$

Damit haben wir gezeigt, dass w bereits im Bild von $(A-\lambda)$ liegt. Also ist $\operatorname{Ran}(A-\lambda)$ abgeschlossen.

Für die andere Richtung nehmen wir an, dass $(A-\lambda)$ injektiv und $\operatorname{Ran}(A-\lambda)$ abgeschlossen ist. Dann hat nach Korollar 3.27 die Abbildung $X \to \operatorname{Ran}(A-\lambda)$, $v \mapsto (A-\lambda)v$ eine stetige Umkehrabbildung, was nach Satz 1.22 die Abschätzung $(\star)$ impliziert.

b) Angenommen, $\lambda \in \sigma_{ap}(A)$ liegt in der Resolventenmenge $\rho(A)$. Dann ergibt

$$1 = \frac{\|(A-\lambda)^{-1}(A-\lambda)v_n\|}{\|v_n\|} \leq \|(A-\lambda)^{-1}\|\frac{\|(A-\lambda)v_n\|}{\|v_n\|} \xrightarrow{n\to\infty} 0$$

einen Widerspruch. Also liegt jedes $\lambda \in \sigma_{ap}(A)$ insbesondere im Spektrum $\sigma(A)$ von A.

c) Wir zeigen: $\lambda \notin \sigma_{ap}(A) \Rightarrow \lambda \notin \sigma_p(A) \cup \sigma_c(A)$. Nach a) impliziert $\lambda \notin \sigma_{ap}(A)$, dass $(A-\lambda)$ injektiv ist, also $\lambda \notin \sigma_p(A)$. Angenommen, $\lambda \in \sigma_c(A)$. Dann wäre $\operatorname{Ran}(A-\lambda)$ dicht, nach a) aber auch abgeschlossen. Also wäre $(A-\lambda)$ bijektiv, was ein Widerspruch zu $\lambda \in \sigma(A)$ ist. $\square$

Eine Folge (v_n) wie in obiger Definition wird auch *Weylfolge* genannt. Wir geben ein Beispiel, dass die Methode illustriert.

Beispiel 4.10. Wir betrachten $\mathcal{H} = L^2(\mathbb{R}, dx)$ und den Multiplikationsoperator

$$(M_g f)(x) = g(x)f(x)$$

mit einer stetigen beschränkten Funktion g. Gegeben $x_0 \in \mathbb{R}$, so erfüllt die Funktion $f_{x_0}(x) := \delta_{x_0,x}$ die Gleichung $g \cdot f_{x_0} = g(x_0) \cdot f_{x_0}$, was wie die Eigenwertgleichung für M_g mit Eigenwert $\lambda = g(x_0)$ aussieht. Allerdings hat f_{x_0} Träger nur in $\operatorname{supp} f_{x_0} = \{x_0\}$, eine Lebesgue-Nullmenge. Also gilt $f_{x_0} = 0$ im Sinne von L^2.

Wir verwenden jetzt eine Weylfolge, um zu zeigen, dass $g(x_0)$ im Spektrum von M_g liegt. Die Idee ist, f_{x_0} auf gewisse Art und Weise zu approximieren. Dazu betrachten wir die Funktionen

$$v_n(x) := \chi_{[x_0-1/n, x_0+1/n]}$$

und, mit $\lambda = g(x_0)$,

$$\frac{\|(A - g(x_0))v_n\|^2}{\|v_n\|^2} = \frac{\int_{x_0-1/n}^{x_0+1/n} |g(x) - g(x_0)|^2 dx}{\int_{x_0-1/n}^{x_0+1/n} dx}.$$

Aufgrund der angenommenen Stetigkeit von g gibt es ein $\xi_n \in [x_0 - \frac{1}{n}, x_0 + \frac{1}{n}]$, so dass $\int_{x_0-1/n}^{x_0+1/n} |g(x) - g(x_0)|^2 dx = |g(\xi_n) - g(x_0)|^2 \int_{x_0-1/n}^{x_0+1/n} dx$ gilt. Für $n \to \infty$ gilt $|g(\xi_n) - g(x_0)| \to 0$, so dass wir $\frac{\|(A-g(x_0))v_n\|}{\|v_n\|} \to 0$ erhalten (hierbei haben wir die Stetigkeit von g verwendet).

Also liegt $g(x_0)$ im approximativen Punktspektrum und insbesondere im Spektrum von M_g. Da $\sigma(M_g)$ abgeschlossen ist, sehen wir

$$\overline{g(\mathbb{R})} \subset \sigma(M_g),$$

dh der Abschluss des Bildes von g liegt im Spektrum von M_g. Tatsächlich gilt hier Gleichheit: Für $\lambda \in \mathbb{C}\backslash\overline{g(\mathbb{R})}$ ist $x \mapsto \frac{1}{g(x)-\lambda}$ eine beschränkte Funktion auf $\mathbb{R}$, und Multiplikation mit dieser Funktion ist offenbar das Inverse zu $M_g - \lambda$. Also $\lambda \in \rho(M_g)$ in diesem Fall.

Diese Analyse zeigt noch nicht, in welchem Teil des Spektrums ein gegebener Bildwert $g(x_0)$ liegt (also Punktspektrum, kontinuierliches Spektrum, oder Residualspektrum). Dazu betrachten wir die Eigenwertgleichung

$$0 = (M_g - g(x_0))f \Leftrightarrow (g(x) - g(x_0))f(x) = 0.$$

Gibt es eine Menge $I \subset \mathbb{R}$ von Lebesguemaß $\mu(I) > 0$, auf der $g(x) = g(x_0)$ für alle $x \in I$ gilt, so ist $f = \chi_I$ ein Eigenvektor zu Eigenwert $g(x_0)$. In diesem Fall gilt also $g(x_0) \in \sigma_p(M_g)$. Gibt es eine solche Menge hingegen nicht, so impliziert die Eigenwertgleichung via $f(x) = 0$ für $g(x) \neq g(x_0)$, dass $f = 0$ in L^2 gilt. In diesem Fall ist $g(x_0)$ also kein Eigenwert.

Beachten Sie, dass die (normierte) Folge $\frac{v_n}{\|v_n\|} = \sqrt{\frac{n}{2}}\chi_{[x_0-1/n,x_0+1/n]}$ in L^2 nicht konvergiert. Dies ist typisch für eine Weylfolge: Würde $v_n/\|v_n\| \to v$ gegen einen Vektor $v \in \mathcal{H}$ konvergieren, so erhalten wir $\|(A - \lambda)v\| = 0$, dh λ ist ein Eigenwert.

Wir werden später eine zu $\overline{g(\mathbb{R})} = \sigma(M_g)$ ähnliche Formel für das Spektrum eines Multiplikationsoperators mit $g \in L^\infty(\Omega, \mu)$ auf einem Maßraum Ω mit Maß μ herleiten. Zunächst kommen wir aber zu der versprochenen exakten Formel für den Spektralradius.

Satz 4.11 (Spektralradiusformel). *Sei X ein Banachraum. Der Spektralradius eines Operators $A \in \mathcal{B}(X)$ ist durch*

$$r_A = \lim_{n\to\infty} \|A^n\|^{1/n} = \inf_{n\in\mathbb{N}} \|A^n\|^{1/n}$$

gegeben (Insbesondere existiert der Grenzwert).

Beweis. Wir zeigen zuerst, dass der angegebene Limes existiert und mit dem angegebenen

Infimum übereinstimmt Dazu setzen wir $x_n := \|A^n\|$ und $x := \inf\{x_n^{1/n} : n \in \mathbb{N}\}$, wir behaupten also $\lim_{n\to\infty} x_n^{1/n} = x$.

Wegen $\|A^{n+m}\| \leq \|A^n\|\|A^m\|$ gilt $x_{n+m} \leq x_n x_m$. Sei $\varepsilon > 0$. Dann finden wir $N \in \mathbb{N}$ mit $x_N^{1/N} < x + \varepsilon$ und setzen $y := \max\{x_1, \dots, x_N\}$. Beliebiges $n \in \mathbb{N}$ schreiben wir nun in der Form $n = kN + r$ mit $r \in \{1, \dots, N\}$. Mit diesem Trick finden wir

$$x_n^{1/n} = x_{kN+r}^{1/n} \leq (x_N^k x_r)^{1/n} \leq (x+\varepsilon)^{\frac{Nk}{n}} y^{\frac{1}{n}} = (x+\varepsilon)^{1-\frac{r}{n}} y^{1/n}.$$

Für $n \to \infty$ konvergiert die rechte Seite gegen $(x+\varepsilon)$, für n groß genug haben wir also $x \leq x_n^{1/n} \leq x + 2\varepsilon$. Das zeigt $\lim_{n\to\infty} \|A^n\|^{1/n} = \inf_{n\in\mathbb{N}} \|A^n\|^{1/n}$.

Für die erste behauptete Identität müssen wir $r_A = x$ zeigen, was wir in zwei Schritten ($x \geq r_A$ und $x \leq r_A$) tun. Sei zunächst $\lambda \in \mathbb{C}$ mit $|\lambda| > x$. Dann gilt

$$\lim_{n\to\infty} \|(\lambda^{-1}A)^n\|^{1/n} = \frac{\lim_{n\to\infty} \|A^n\|^{1/n}}{|\lambda|} = \frac{x}{|\lambda|} < 1.$$

Wir erinnern uns nun an das Wurzelkriterium aus Analysis 1: Gibt es eine Folge $t_n \geq 0$ mit $\lim_{n\to\infty} t_n^{1/n} < 1$, so konvergiert die Reihe $\sum_{n=0}^\infty t_n$. Übertragen auf Operatoren heißt dies:

$$\lim_{n\to\infty} \|\lambda^{-n}A^n\|^{1/n} < 1 \;\Rightarrow\; -\lambda \sum_{n=0}^\infty \lambda^{-n}A^n = R_A(\lambda)$$

im Sinne von Operatornormkonvergenz (Details: Übung). Insbesondere gilt dann $\lambda \in \rho(A)$.

Da genau diese Situation vorliegt, sehen wir, dass jedes λ mit $|\lambda| > x$ in der Resolventenmenge liegt. Das heißt $r_A \leq x$.

Für den Beweis der umgekehrten Ungleichung $x \leq r(A)$ betrachten wir zwei offene Gebiete $\Omega, \tilde{\Omega} \subset \rho(A)$,

$$\tilde{\Omega} := \{\lambda \in \mathbb{C} : |\lambda| > x\} \subset \Omega := \{\lambda \in \mathbb{C} : |\lambda| > r_A\}.$$

Wie oben erläutert, konvergiert die Reihe $-\lambda^{-1}\sum_{n=0}^\infty \lambda^{-n}A^n = R_A(\lambda)$ für λ in dem potentiell kleineren Gebiet $\tilde{\Omega}$. Um Ergebnisse der Funktionentheorie nutzbar zu machen, wählen wir nun ein (beliebiges) stetiges lineares Funktional $\varphi \in \mathcal{B}(\mathcal{H})'$ und betrachten die zahlenwertige Funktion

$$f : \rho(A) \to \mathbb{C}, \qquad f(\lambda) := \varphi(R_A(\lambda)),$$

die für $\lambda \in \tilde{\Omega}$ durch

$$f(\lambda) = -\lambda^{-1} \sum_{n=0}^\infty \lambda^{-n} \varphi(A^n)$$

gegeben ist. Dies ist eine holomorphe Funktion (sie hat lokal eine Potenzreihendarstellung, siehe Funktionentheorie). In der Funktionentheorie beweist man, dass in diesem Fall die Reihe automatisch auf dem größten Kreisring konvergiert, in dem f holomorph ist. Dies ist nach Definition des Spektralradius genau das größere Gebiet Ω.

Für $\lambda \in \Omega$ konvergiert also die Reihe $\sum_{n=0}^{\infty} \lambda^{-n}\varphi(A^n)$: Insbesondere muss demnach $\lambda^{-n}\varphi(A^n) \to 0$ für $n \to \infty$ konvergieren. Da φ beliebig war, heißt das, dass die Operatorfolge $\lambda^{-n}A^n$ schwach gegen Null konvergiert. Nun können wir das Prinzip der gleichmäßigen Beschränktheit in der Form von Lemma 3.19 anwenden und sehen, dass die Folge $(\lambda^{-n}A^n)_n$ *in Operatornorm* beschränkt ist. Es gibt also $c > 0$ mit $\|\lambda^{-n}A^n\| \leq c$ für alle $n \in \mathbb{N}$.

Aus dieser Umgleichung erhalten wir leicht

$$x \leftarrow \|A^n\|^{1/n} \leq c^{1/n}|\lambda| \to |\lambda|,$$

wobei die Pfeile Konvergenz für $n \to \infty$ bedeuten. Also gilt $x \leq |\lambda|$ für jedes $\lambda \in \Omega$. Im Grenzfall $|\lambda| \to r_A$ (vgl. Definition von Ω) ergibt das die gewünschte Ungleichung $x \leq r_A$. □

4.2 Normale Hilbertraumoperatoren

Nachdem wir bisher Spektren von (beschränkten) Operatoren auf allgemeinen Banachräumen betrachtet haben, wenden wir uns nun dem spezielleren Hilbertraumfall zu, dh betrachten $A \in \mathcal{B}(\mathcal{H})$ mit einem Hilbertraum $\mathcal{H}$. Da das Spektrum im Falle eines reellen Hilbertraums leer sein kann, betrachten wir hier nur komplexes $\mathcal{H}$.

Wir haben damit also die Adjunktion $A \mapsto A^*$ (Def. 2.21) zur Verfügung und können über selbstadjungierte, unitäre, normale Operatoren und orthogonale Projektionen sprechen (siehe Kapitel 2). Insbesondere normale Operatoren – also die $A \in \mathcal{B}(\mathcal{H})$, die $AA^* = A^*A$ erfüllten – werden von Interesse sein. In diesem kurzen Abschnitt sammeln wir etwas Information zu normalen Operatoren.

Lemma 4.12. *Sei $A \in \mathcal{B}(\mathcal{H})$ normal.*

a) Für jedes $\xi \in \mathcal{H}$ gilt $\|A^*\xi\| = \|A\xi\|$.

b) $\ker(A^*) = \ker(A)$.

c) Ist ξ ein Eigenvektor von A mit Eigenwert λ, so ist ξ auch ein Eigenvektor von A^* mit Eigenwert $\overline{\lambda}$.

d) Eigenvektoren von A zu verschiedenen Eigenwerten stehen aufeinander senkrecht.

Beweis. a) $\|A^*\xi\|^2 = \langle A^*\xi, A^*\xi\rangle = \langle \xi, AA^*\rangle = \langle \xi, A^*A\xi\rangle = \langle A\xi, A\xi\rangle = \|A\xi\|^2$.

b) folgt sofort aus a) (und ist ein Spezialfall von c)).

c) Mit A ist auch $(A-\lambda)$ normal, da $(A-\lambda)^* = (A^* - \overline{\lambda})$ mit $(A-\lambda)$ kommutiert. Falls $A\xi = \lambda\xi$, so gilt $\xi \in \ker(A-\lambda) = \ker(A^* - \overline{\lambda})$, dh $A^*\xi = \overline{\lambda}\xi$.

d) Seien $\xi_1, \xi_2 \in \mathcal{H}$ mit $A\xi_1 = \lambda_1\xi_1$, $A\xi_2 = \lambda_2\xi_2$, $\lambda_1 \neq \lambda_2$. Dann gilt

$$\lambda_2\langle \xi_1, \xi_2\rangle = \langle \xi_1, \lambda_2\xi_2\rangle = \langle \xi_1, A\xi_2\rangle = \langle A^*\xi_1, \xi_2\rangle \overset{c)}{=} \langle \overline{\lambda_1}\xi_1, \xi_2\rangle = \lambda_1\langle \xi_1, \xi_2\rangle,$$

also $0 = (\lambda_1 - \lambda_2)\langle \xi_1, \xi_2\rangle$ und wegen $\lambda_1 \neq \lambda_2$ folgt $\langle \xi_1, \xi_2\rangle = 0$. □

Allgemein bestehen folgende Zusammenhänge zwischen dem Spektrum eines (nicht notwendigerweise normalen) Operatores A und dem Spektrum seines Adjungierten A^*.

Satz 4.13. *Sei $A \in \mathcal{B}(\mathcal{H})$. Dann gilt:*

a) $\sigma(A^*) = \{\overline{\lambda} : \lambda \in \sigma(A)\}$.

b) Falls $\lambda \in \sigma_p(A)$, *so gilt* $\overline{\lambda} \in \sigma_p(A^*) \cup \sigma_r(A^*)$.

c) Falls $\lambda \in \sigma_r(A)$, *so gilt* $\overline{\lambda} \in \sigma_p(A^*)$.

Beweis. a) Hausaufgabe H9.1 b).

b), c) Wir benutzen die Beziehung $\ker(B) = \operatorname{Ran}(B^*)^\perp$ für $B \in \mathcal{B}(\mathcal{H})$. Angewendet auf $B = A - \lambda$ ergibt sich $\ker(A-\lambda) = \operatorname{Ran}(A^* - \bar{\lambda})^\perp$. Falls $\lambda \in \sigma_p(A)$, so gilt $\ker(A-\lambda) \neq \{0\}$ und deshalb ist $\operatorname{Ran}(A^* - \bar{\lambda})$ nicht dicht. Also $\bar{\lambda} \notin \sigma_c(A^*)$. Da $\bar{\lambda} \in \sigma(A^*)$ gemäß a), muss $\bar{\lambda}$ entweder im Punkt- oder Residualspektrum von A^* liegen.

Für c) setzen wir $B^* = A - \lambda$ und betrachten $\lambda \in \sigma_r(A)$, also $\operatorname{Ran}(A - \lambda)$ nicht dicht und damit $\{0\} \neq \operatorname{Ran}(A - \lambda)^\perp = \ker(A^* - \bar{\lambda})$. Also $\bar{\lambda} \in \sigma_p(A^*)$. □

Satz 4.14. *Sei $A \in \mathcal{B}(\mathcal{H})$ normal. Dann besteht das Spektrum von A genau aus den verallgemeinerten Eigenwerten von A und das Residualspektrum von A ist leer, d.h.* $\sigma(A) = \sigma_{ap}(A) = \sigma_p(A) \cup \sigma_c(A)$.

Beweis. Wir zeigen zuerst $\sigma_r(A) = \emptyset$: Für $\lambda \in \sigma_r(A)$ gilt $\bar{\lambda} \in \sigma_p(A^*)$ nach dem vorigen Satz, und mit Lemma 4.12 c) auch $\lambda \in \sigma_p(A)$. Aber dies ist ein Widerspruch, da das Punkt- und Residualspektrum disjunkt sind. Also gibt es $\lambda \in \sigma_r(A)$ nicht.

Nun folgt mit Lemma 4.9

$$\sigma_{ap}(A) \subset \sigma(A) \subset \sigma_p(A) \cup \sigma_c(A) \subset \sigma_{ap}(A)$$

auch $\sigma_{ap}(A) = \sigma(A)$. □

Für die wichtige Klasse der normalen Operatoren reduziert sich der Begriff des Spektralwerts also auf den Begriff des verallgemeinerten Eigenwerts, und der Begriff des Spektrums auf das Punktspektrum und das kontinuierliche Spektrum.

Auch die Formel für den Spektralradius vereinfacht sich für normale Operatoren stark.

Satz 4.15. *Sei $A \in \mathcal{B}(\mathcal{H})$ ein normaler Operator. Dann gilt $\|A^2\| = \|A\|^2$, und der Spektralradius von A ist $r_A = \|A\|$.*

Beweis. Wir betrachten zuerst den Spezialfall eines selbstadjungierten Operators $A = A^*$. Dann gilt wegen der C^*-Eigenschaft $\|A^2\| = \|A^*A\| = \|A\|^2$ die behauptete Normgleichung. Ist nun A nur normal, argumentieren wir wie folgt:

$$\begin{aligned}
\|A^2\|^2 &= \|(A^2)^*A^2\| \\
&= \|(A^*A)^2\| && \text{(weil } A \text{ normal ist)} \\
&= \|A^*A\|^2 && \text{(weil } A^*A \text{ selbstadjungiert ist)} \\
&= \|A\|^4 && \text{(wegen der } C^*\text{-Eigenschaft).}
\end{aligned}$$

Dies ergibt $\|A^2\| = \|A\|^2$ auch für normale Operatoren. Iteriert man diese Gleichung, so sieht man $\|A^{2^n}\| = \|A\|^{2^n}$ für alle $n \in \mathbb{N}$.

Aus der allgemeinen Grenzwertformel für den Spektralradius (Satz 4.11) erhalten wir nun

$$r_A = \lim_{n\to\infty} \|A^n\|^{1/n} = \lim_{n\to\infty} \|A^{2^n}\|^{2^{-n}} = \|A\|.$$

□

Wichtige Spezialfälle von normalen Operatoren sind unitäre Operatoren (A invertierbar mit $A^* = A^{-1}$) und vor allem selbstadjungierte Operatoren ($A = A^*$). In Hausaufgabe H9.1c) haben Sie sich bereits davon überzeugt, dass das Spektrum eines unitären Operators stets im Einheitskreis $\{\lambda \in \mathbb{C} : |\lambda| = 1\}$ enthalten ist. Wir betrachten nun das Spektrum eines selbstadjungierten Operators.

Satz 4.16. *Das Spektrum eines selbstadjungierten Operators $A = A^* \in \mathcal{B}(\mathcal{H})$ ist reell, $\sigma(A) \subset \mathbb{R}$.*

Beweis. Sei $\lambda = a + ib$ die Zerlegung eines Spektralwertes $\lambda \in \sigma(A)$ in Real- und Imaginärteil, und $x \in \mathbb{R}$ beliebig. Dann gilt $a + i(b + x) \in \sigma(A + ix)$, und

$$\begin{aligned}
a^2 + b^2 + 2bx + x^2 &= |a + i(b + x)|^2 \\
&\leq \|A + ix\|^2 && \text{(Spektralradius)} \\
&= \|(A + ix)^*(A + ix)\| && \text{(C^*-Eigenschaft)} \\
&= \|(A - ix)(A + ix)\| && \text{(Selbstadjungiertheit)} \\
&= \|A^2 + x^2\| \\
&\leq \|A\|^2 + x^2.
\end{aligned}$$

Das ergibt $a^2 + b^2 + 2bx \leq \|A\|^2$ für alle $x \in \mathbb{R}$, was nur für $b = 0$ richtig ist. □

Wir betrachten nun eingehend die spezielle aber wichtige Klasse von Multiplikationsoperatoren.

Beispiel 4.17. Sei (Ω, Σ, μ) ein Maßraum mit einem σ-endlichen Maß μ (d.h. Ω ist eine abzählbare Vereinigung von messbaren Mengen Δ_n, $n \in \mathbb{N}$, mit $\mu(\Delta_n) < \infty$, wie zB das Lebesguemaß auf $\mathbb{R}^d$), $\mathcal{H} := L^2(\Omega, d\mu)$, und $g \in L^\infty(\Omega, d\mu)$ eine messbare wesentlich beschränkte Funktion (siehe Analysis 3). Wir betrachten den Multiplikationsoperator

$$M_g : L^2(\Omega, d\mu) \to L^2(\Omega, d\mu), \qquad M_g f := g \cdot f.$$

- **(Operatornorm)** Da $|g(x)| \leq \|g\|_{L^\infty}$ für μ-fast alle $x \in \Omega$ gilt, haben wir

  $$\|gf\|_{L^2}^2 = \int_\Omega |g(x)f(x)|^2 d\mu(x) \leq \|g\|_{L^\infty}^2 \int_\Omega |f(x)|^2 d\mu(x) = \|g\|_{L^\infty}^2 \|f\|_{L^2}^2.$$

  Also ist M_g beschränkt mit $\|M_g\| \leq \|g\|_{L^\infty}$. Wir behaupten sogar

  $$\|M_g\| = \|g\|_{L^\infty}$$

und müssen noch die Ungleichung "$\geq$", also $|g(x)| \leq \|M_g\|$ für μ-fast alle $x \in \Omega$ zeigen. Dazu wählen wir eine Konstante $0 \leq a < \|g\|_{L^\infty}$ und betrachten die messbare Menge $E_a := \{|g| > a\}$ (siehe Analysis 3), die nach Konstruktion $\mu(E_a) > 0$ erfüllt. Da μ σ-endlich ist, enthält E_a eine messbare Teilmenge $\tilde{E}_a$ mit $0 < \tilde{\mu}(E_a) < \infty$ (Schneide E_a mit einer der Mengen Δ_n). Wir betrachten die charakteristische Funktion $\chi_{\tilde{E}_a} \in L^2(\Omega, d\mu)$ und erhalten

$$a^2\|\chi_{\tilde{E}_a}\|_{L^2}^2 = \int_{\tilde{E}_a} a^2\, d\mu(x) \leq \int_{\tilde{E}_a} |g(x)|^2\, d\mu(x) = \|g\chi_{\tilde{E}_a}\|_{L^2}^2 \leq \|M_g\|^2\|\chi_{\tilde{E}_a}\|_{L^2}^2,$$

also $a \leq \|M_g\|$. Da $0 \leq a < \|g\|_{L^\infty}$ beliebig war, folgt im Limes $a \to \|g\|_{L^\infty}$ die Behauptung.

- **(Adjungiertes und Normalität)** Wir behaupten

  $$M_g{}^* = M_{\bar{g}}.$$

  Beweis: Für $f, h \in L^2(\Omega, d\mu)$ gilt

  $$\langle f, M_g h\rangle = \int_\Omega \overline{f(x)}g(x)h(x)d\mu(x) = \int_\Omega \overline{\overline{g(x)}f(x)}h(x)d\mu(x) = \langle M_{\bar{g}}f, h\rangle,$$

  und damit folgt die Behauptung.

  Da alle Multiplikationsoperatoren untereinander kommutieren, folgt, dass M_g normal ist. M_g ist genau dann selbstadjungiert, wenn g fast überall reell ist, genau dann unitär, wenn g fast überall Betrag 1 hat, und eine orthogonale Projektion, wenn $g(x) \in \{0,1\}$ für fast alle $x \in \Omega$ erfüllt.

- **(Spektrum)** Das *wesentliche Bild* $\operatorname{ess.im}(g)$ (essential range) der messbaren Funktion g ist definiert als

  $$\operatorname{ess.im}(g) = \{\lambda \in \mathbb{C} \,:\, \mu(\{x \,:\, |g(x) - \lambda| < \varepsilon\}) > 0\, \forall\varepsilon > 0\}.$$

  Wir behaupten, dass $\operatorname{ess.im}(g)$ eine abgeschlossene Teilmenge des Abschlusses $\overline{\operatorname{im}(g)}$ des Bildes von g ist.

  Sei $\lambda \in \operatorname{ess.im}(g)$. Dann ist nach Voraussetzung die Menge $\{x : |g(x)-\lambda| < \varepsilon\}$ für kein $\varepsilon > 0$ leer (denn sonst hätte sie Maß 0). In jeder ε-Umgebung von λ finden sich also Punkte $g(x)$ im Bild von g, dh λ liegt im Abschluss vom Bild von g.

  Zur Abgeschlossenheit: Sei $(\lambda_n)_{n\in\mathbb{N}}$ eine Folge in $\operatorname{ess.im}(g)$, die gegen $\lambda \in \mathbb{C}$ konvergiert, und sei $\varepsilon > 0$ beliebig. Dann gibt es $N_\varepsilon \in \mathbb{N}$, so dass $|\lambda_{N_\varepsilon} - \lambda| < \frac{\varepsilon}{2}$. Sei nun $x \in \Omega$ so, dass $|g(x) - \lambda_{N_\varepsilon}| < \frac{\varepsilon}{2}$ gilt. Dann gilt auch $|g(x) - \lambda| < \varepsilon$ per Dreiecksungleichung. Also $\{x : |g(x) - \lambda_{N_\varepsilon}| < \frac{\varepsilon}{2}\} \subset \{x : |g(x) - \lambda| < \varepsilon\}$. Wegen $\lambda_{N_\varepsilon} \in \operatorname{ess.im}(g)$ hat die kleinere dieser Mengen positives Maß, also (Monotonie des Maßes) hat auch die größere Menge positives Maß. Das bedeutet aber $\lambda \in \operatorname{ess.im}(g)$. Also ist $\operatorname{ess.im}(g)$ abgeschlossen.

Wir behaupten

$$\sigma(M_g) = \operatorname{ess.im}(g).$$

Durch Komplementbildung der Definition des wesentlichen Bildes erhält man eine für den Beweis hilfreiche Charakterisierung von $\operatorname{ess.im}(g)$:

$$\mathbb{C}\setminus \operatorname{ess.im}(g) = \left\{ \lambda \in \mathbb{C} \;:\; \exists c > 0 : \frac{1}{|g(x) - \lambda|} \leq c \;\text{ für fast alle } x \in \Omega \right\}.$$

Wir zeigen $\rho(M_g) = \mathbb{C}\setminus \operatorname{ess.im}(g)$. Angenommen, $\lambda \notin \operatorname{ess.im}(g)$. Dann ist die Funktion $(g - \lambda)^{-1} : x \mapsto (g(x) - \lambda)^{-1}$ für fast alle x wohldefiniert und wesentlich beschränkt, also $(g - \lambda)^{-1} \in L^\infty(\Omega, d\mu)$. Der zugehörige Multiplikationsoperator $M_{(g-\lambda)^{-1}}$ ist also beschränkt. Aber offenbar gilt $M_{g-\lambda}^{-1} = M_{(g-\lambda)^{-1}}$. Also $\lambda \in \rho(M_g)$.

Nehmen wir nun $\lambda \in \rho(M_g)$ an. Dann existiert $M_{g-\lambda}^{-1}$ als beschränkter Operator, und wir haben

$$(g(x) - \lambda)(M_{g-\lambda}^{-1} f)(x) = f(x), \qquad f \in L^2(\Omega, d\mu).$$

Da diese Gleichung für alle $f \in L^2(\Omega, d\mu)$ gilt, sehen wir, dass $x \mapsto g(x) - \lambda$ nur auf einer Nullmenge verschwinden kann. Wir können also durch $g(x) - \lambda$ teilen und sehen, dass $M_{g-\lambda}^{-1} = M_{(g-\lambda)^{-1}}$ gilt. Also ist $(g - \lambda)^{-1}$ wesentlich beschränkt mit $\|(g - \lambda)^{-1}\|_{L^\infty} = \|M_{(g-\lambda)}^{-1}\| < \infty$. Wir haben also $\lambda \notin \operatorname{ess.im}(g)$.

Falls g stetig ist, stimmt $\operatorname{ess.im}(g)$ mit $\overline{\operatorname{im}(g)}$ überein (Übung).

Wir bemerken noch, dass das Supremum aller Werte im wesentlichen Bild das wesentliche Supremum ist

$$\|g\|_{L^\infty} = \sup\{|\lambda| \;:\; \lambda \in \operatorname{ess.im}(g)\}.$$

Dies bestätigt die vereinfachte Spektralradiusformel $r_A = \|A\|$ für den normalen Operator $A = M_g$.

Es hat einen guten Grund, dass wir uns die speziellen Multiplikationsoperatoren so genau ansehen: Später werden wir zeigen, dass bis auf eine unitäre Transformation jeder normale Operator ein Multiplikationsoperator ist!

Die Multiplikationsoperatoren sind also ein Modell für beliebige normale Operatoren; auch im endlichdimensionalen Fall. Dort ist ein normaler Operator durch eine normale Matrix $A \in M_n(\mathbb{C})$ gegeben. Eine normale Matrix ist diagonalisierbar, dh es gibt eine Orthonormalbasis $(e_k)_{k=1,\ldots,d}$ unseres (für jetzt endlichdimensionalen) Hilbertraums $\mathcal{H}$, so dass $Ae_k = \lambda_k \cdot e_k$. Um dies als Multiplikationsoperator zu formulieren, wählen wir $\Omega = \{1, \ldots, d\}$, μ das Zählmaß auf Ω, und $g(k) := \lambda_k$, $k = 1, \ldots, d$. Wir betrachten den unitären Operator

$$U : \mathcal{H} \to L^2(\Omega, d\mu), \qquad Ue_k := \delta_k,$$

wobei $\delta_k(j) = \delta_{k,j}$. Dann gilt $UAU^{-1} = M_g$ (Übung), dh die normale Matrix A wird auf dem Raum $L^2(\Omega, d\mu)$ durch den Multiplikationsoperator M_g dargestellt.

Eine solche Darstellung auch im Unendlichdimensionalen zu finden, ist ein schwieriges Problem, das wir mit Hilfe einer detaillierten Spektralanalyse lösen werden.

4.3 Der polynomielle und stetige Funktionalkalkül

Wir wenden uns nun dem Studium von sogenannten Funktionalkalkülen zu. Darunter ist im Wesentlichen die Frage zu verstehen, wie "Funktionen von Operatoren" gebildet/definiert werden können, dh gegeben einen Operator A (mit gewissen Eigenschaften, zB normal) und eine Funktion (mit gewissen Eigenschaften) f, was sollten wir unter "$f(A)$" verstehen? Was ist zum Beispiel die Wurzel aus einer Matrix, oder der Logarithmus eines Integraloperators? Je nachdem, welche Klassen von Operatoren und welche Klassen von Funktionen betrachtet werden, gibt es eine Reihe von unterschiedlichen Funktionalkalkülen.

Wir beginnen in diesem Abschnitt mit der elementarsten Klasse von Funktionen, nämlich Polynomen. Gegeben einen Hilbertraumoperator $A \in \mathcal{B}(\mathcal{H})$ und eine Polynomfunktion $\mathbb{C} \ni z \mapsto f(z) = \sum_{k=0}^{n} c_k\, z^k$ in einer Variable (mit beliebigen Koeffizienten $c_k \in \mathbb{C}$), so möchten wir

$$f(A) := \sum_{k=0}^{n} c_k\, A^k$$

definieren, wobei $A^0 := 1$ und $A^k = A \circ A \circ \ldots \circ A$ die k-fache Komposition von A mit sich selbst bezeichnet.

Diese Definition ist rein algebraisch und scheint keinerlei Schwierigkeiten zu machen. Wir beobachten: Ist v ein Eigenvektor von A zu Eigenwert λ, so ist v auch ein Eigenvektor von $f(A)$ zum Eigenwert $f(\lambda)$. Diese Beobachtung legt nahe, dass die Einschränkung von f auf das Spektrum $\sigma(A)$ eine wichtige Rolle spielen sollte.

Allerdings hängt $f(A)$ im Allgemeinen nicht nur von $f|_{\sigma(A)}$ ab. Betrachten Sie zum Beispiel

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

als einen (nicht normalen) Operator auf $\mathbb{C}^2$, so gilt $\sigma(A) = \{0\}$. Die beiden Polynome $f_1(z) = 0$ und $f_2(z) = z - z^2$ stimmen auf $\sigma(A)$ überein, aber

$$f_2(A) = A - A^2 = A \neq 0 = f_1(A).$$

Diese Beobachtung zeigt, dass wir einen gut mit Spektraltheorie zusammenspielenden polynomiellen Funktionalkalkül nur für normale Operatoren erreichen werden. Um zu zeigen, dass für normales A keine Probleme auftreten, betrachten wir in der nächsten Proposition das Zusammenspiel von Polynomen mit Spektrum und Norm. Dazu definieren wir für beliebige Polynome (oder gleich stetige Funktionen $f : \sigma(A) \to \mathbb{C}$) die Norm

$$\|f\|_{C(\sigma(A))} := \max\{|f(\lambda)| \,:\, \lambda \in \sigma(A)\}.$$

Das Maximum existiert, da f eine stetige Funktion auf der kompakten Menge $\sigma(A)$ ist.

Satz 4.18. *Sei $A \in \mathcal{B}(\mathcal{H})$ und f eine Polynomfunktion.*

a) Das spektrale Abbildungsprinzip *gilt:*

$$\sigma(f(A)) = f(\sigma(A)) = \{f(\lambda) : \lambda \in \sigma(A)\}.$$

b) $\|f(A)\| \geq \|f\|_{C(\sigma(A))}$.

c) Wenn A normal ist, gilt

$$\|f(A)\| = \|f\|_{C(\sigma(A))}.$$

Beweis. a) Sei f ein Polynom. Falls f konstant ist, ist nichts zu zeigen. Falls f nicht konstant ist, betrachten wir für zunächst beliebiges $\lambda \in \mathbb{C}$ das Polynom $f(X) - \lambda$ und zerlegen es in Linearfaktoren, $f(z) - \lambda = c(z - w_1) \cdots (z - w_n)$, mit $c \neq 0$. Dann haben wir

$$f(A) - \lambda = c(A - w_1) \cdots (A - w_n).$$

Die linke Seite $f(A) - \lambda$ ist genau dann invertierbar, wenn $\lambda \in \rho(f(A))$. Die rechte Seite ist genau dann invertierbar, wenn alle Faktoren $(A - w_k)$ invertierbar sind[7], d.h. genau dann, wenn $w_1, \ldots, w_n \in \rho(A)$.

Unter Beachtung von $f(w_1) = \ldots = f(w_k) = \lambda$ haben wir dann die Äquivalenzen

$$\begin{aligned}
\lambda \in \sigma(f(A)) &\iff f(A) - \lambda \text{ nicht invertierbar} \\
&\iff \exists k \in \{1, \ldots, n\} : (A - w_k) \text{ nicht invertierbar} \\
&\iff \exists k \in \{1, \ldots, n\} : w_k \in \sigma(A) \\
&\iff \lambda = f(w_k) \in f(\sigma(A)).
\end{aligned}$$

b) Wir verwenden die Ungleichung $r_{f(A)} \leq \|f(A)\|$ zwischen Spektralradius und Norm von $f(A)$. Nach Teil a) gilt

$$\begin{aligned}
r_{f(A)} &= \max\{|\lambda| : \lambda \in \sigma(f(A))\} \\
&= \max\{|\lambda| : \lambda \in f(\sigma(A))\} \\
&= \max\{|f(\mu)| : \mu \in \sigma(A)\} \\
&= \|f\|_{C(\sigma(A))},
\end{aligned}$$

was $\|f(A)\| \geq r_{f(A)} = \|f\|_{C(\sigma(A))}$ zeigt.

c) Wir bemerken, dass für normales A auch $f(A)$ normal ist, also $r_{f(A)} = \|f(A)\|$ gilt (Prop. 4.15). Damit folgt c) genau analog zu b). □

Teil c) besagt insbesondere, dass $f(A)$ nur von der Einschränkung von f auf das Spektrum von A abhängt, falls A normal ist: Denn falls $f(\lambda) = 0$ für alle $\lambda \in \sigma(A)$, so gilt $\|f(A)\| = \|f\|_{C(\sigma(A))} = 0$, also $f(A) = 0$.

[7] Sind alle Faktoren invertierbar, so gilt $((A - w_1) \cdots (A - w_n))^{-1} = (A - w_n)^{-1} \cdots (A - w_1)^{-1}$. Ist ein Faktor $(A - w_k)$ nicht invertierbar, so ist $(A - w_k)$ entweder nicht injektiv oder nicht surjektiv. Da alle Faktoren kommutieren, folgt in beiden Fällen, dass $c(A - w_1) \cdots (A - w_n)$ nicht invertierbar ist (falls $(A - w_k)$ nicht injektiv ist, sortiere diesen Faktor nach rechts, falls $(A - w_k)$ nicht surjektiv ist, sortiere ihn nach links).

Es lohnt sich, an dieser Stelle kurz innezuhalten und die gerade entwickelte Struktur zu betrachten und in einen allgemeineren Kontext zu stellen. Wir betrachten eine Menge $\mathcal{F}$ von Funktionen $f : \sigma(A) \to \mathbb{C}$ (im aktuellen Fall sind dies Einschränkungen von Polynomen auf das Spektrum von A) und definieren zu jeder Funktion $f \in \mathcal{F}$ einen Operator $f(A)$, der der Idee "Auswerten von f in A" entspricht. Dies gibt uns also eine Abbildung $\phi_A : \mathcal{F} \to \mathcal{B}(\mathcal{H})$, $\phi_A(f) := f(A)$.

Um die wesentlichen Eigenschaften dieser Abbildung zu betonen, machen wir die folgende Definition.

Definition 4.19. Sei $A \in \mathcal{B}(\mathcal{H})$ und $\mathcal{F}$ eine unitale *-Algebra von Funktionen $f : \sigma(A) \to \mathbb{C}$, dh eine Menge mit den Eigenschaften

a) $f, g \in \mathcal{F} \Rightarrow f + g \in \mathcal{F}$ und $f \cdot g \in \mathcal{F}$,

b) $f \in \mathcal{F} \Rightarrow \bar{f} \in \mathcal{F}$,

c) Die Einsfunktion $f(\lambda) = 1$ liegt in $\mathcal{F}$,

die weiterhin auch die Identitätsfunktion $\mathrm{id} : \sigma(A) \to \mathbb{C}$, $\lambda \mapsto \lambda$, enthält.

Ein *Funktionalkalkül* für A und $\mathcal{F}$ ist ein Homomorphismus von unitalen *-Algebren

$$\phi_A^{\mathcal{F}} : \mathcal{F} \to \mathcal{B}(\mathcal{H}),$$

das heißt

a) $\phi_A^{\mathcal{F}}$ ist linear,

b) $\phi_A^{\mathcal{F}}(fg) = \phi_A^{\mathcal{F}}(f)\phi_A^{\mathcal{F}}(g)$,

c) $\phi_A^{\mathcal{F}}(\bar{f}) = \phi_A^{\mathcal{F}}(f)^*$.

d) $\phi_A^{\mathcal{F}}(1) = 1$.

Weiterhin wird $\phi_A^{\mathcal{F}}(\mathrm{id}) = A$ gefordert.

Beispiele für unitale *-Algebren von Funktionen $f : \sigma(A) \to \mathbb{C}$ auf dem Spektrum eines Operators sind $C(\sigma(A))$ (stetige Funktionen) oder messbare Funktionen (bzgl der Borel-σ-Algebra von $\sigma(A)$). Unsere bisher betrachtete Menge $\mathcal{P}(\sigma(A))$ von Polynomfunktionen ist abgeschlossen unter Linearkombinationen und Produkten und enthält 1 und id. Allerdings ist $\mathcal{P}(\sigma(A))$ im Allgemeinen *nicht* unter komplexer Konjugation abgeschlossen, weil $z \mapsto \overline{z}$ *kein* Polynom in z ist[8].

Diese Schwierigkeit verschwindet allerdings, wenn wir *selbstadjungiertes* A betrachten, da dann $\sigma(A) \subset \mathbb{R}$ gilt und die Polynomfunktionen einer *reellen* Variable unter komplexer Konjugation abgeschlossen sind. Wir beschränken uns deshalb zunächst auf selbstadjungiertes A. In diesem Fall bemerken wir:

[8] Sie können das zB so sehen: Für ein Polynom $f : \mathbb{C} \to \mathbb{C}$ existiert der Grenzwert $\lim_{z\to 0} \frac{f(z)-f(0)}{z}$ (Ableitung nach z bei $z = 0$) und hängt nicht davon ab, auf welchem Weg z gegen Null konvergiert. Insbesondere gilt $\lim_{x\to 0} \frac{f(x)-f(0)}{x} = \lim_{y\to 0} \frac{f(iy)-f(0)}{iy}$, wobei x, y jeweils reell sind. Andererseits gilt für die komplexe Konjuation $\frac{\bar{z}}{z} = 1$ für $z \in \mathbb{R}$ und $\frac{\bar{z}}{z} = -1$ für $z \in i\mathbb{R}$. Also ist $z \mapsto \bar{z}$ kein Polynom.

Korollar 4.20. *Sei $A = A^* \in \mathcal{B}(\mathcal{H})$ selbstadjungiert und $\mathcal{F} = \mathcal{P}(\sigma(A))$ die Menge der Polynomfunktionen auf $\sigma(A)$. Dann ist*

$$\phi_A^{\mathcal{F}} : \mathcal{F} \to \mathcal{B}(\mathcal{H}), \qquad f \mapsto f(A)$$

(mit $f(A)$ wie vorher definiert) der eindeutige Funktionalkalkül. Wird $\mathcal{P}(\sigma(A))$ mit der Supremumsnorm $\|\cdot\|_{C(\sigma(A))}$ ausgestattet, so ist $\phi_A^{\mathcal{F}}$ isometrisch und insbesondere stetig. Weiterhin erfüllt $\phi_A^{\mathcal{F}}$ das spektrale Abbildungsprinzip.

Beweis. Der Nachweis der Eigenschaften a)–d) eines Funktionalkalküls ist durch direkte Rechnung zu führen, zB zeigt für $f(x) = \sum_{k=0}^n c_k x^k$

$$\phi_A^{\mathcal{F}}(\bar{f}) = \bar{f}(A) = \sum_{k=0}^n \overline{c_k} A^k = \sum_{k=0}^n \overline{c_k} (A^*)^k = \left(\sum_{k=0}^n c_k \, A^k \right)^* = \phi_A^{\mathcal{F}}(A)^*$$

Eigenschaft c), und $\phi_A^{\mathcal{F}}(\mathrm{id}) = \mathrm{id}(A) = A$ ist Teil der Definition des Kalküls. Da jedes Polynom aus 1 und id durch Linearkombinationen und Produktbildung hervorgeht, legt dies $\phi_A^{\mathcal{F}}$ bereits eindeutig fest.

Das spektrale Abbildungsprinzip und die Isometrieeigenschaft von $\phi_A^{\mathcal{F}}$ wurden in Satz 4.18 gezeigt. □

Wir möchten diesen Funktionalkalkül jetzt auf eine größere Algebra von Funktionen ausdehnen, nämlich die stetigen Funktionen $C(\sigma(A)) \supset \mathcal{P}(\sigma(A))$ (vorerst für $A = A^*$). Die Idee ist, stetige Funktionen durch Polynome zu approxmieren. Dazu verwenden wir den Satz von Stone-Weierstraß (siehe Topologie-Vorlesung oder z.B. Heuser, Lehrbuch der Analysis, Teil 2, Satz 116.1) in der folgenden Form.

Satz 4.21 (Satz von Stone-Weierstraß). *Sei X eine kompakte Teilmenge eines normierten Raumes E und $\mathcal{A}$ eine unitale *-Unteralgebra der Algebra $C(X)$ aller stetigen Funktionen $f : X \to \mathbb{C}$ (d.h. insbesondere $1 \in \mathcal{A}$ und $(f \in \mathcal{A} \Rightarrow \bar{f} \in \mathcal{A})$ die punktetrennend ist in dem Sinne, das für beliebige $x, y \in X$, $x \neq y$, stets $f \in \mathcal{A}$ existiert mit $f(x) \neq f(y)$. Dann ist $\mathcal{A} \subset C(X)$ dicht in der Supremumsnorm $\|f\|_X = \sup_{x \in X} |f(x)|$.*

Wir bemerken, dass $X := \sigma(A) \subset E := \mathbb{R}$ und $\mathcal{A} := \mathcal{P}(\sigma(A))$ die Voraussetzungen des Satzes erfüllen: $\sigma(A)$ ist kompakt (Satz 4.4) und $\mathcal{P}(\sigma(A))$ ist eine unitale *-Unteralgebra von $C(\sigma(A))$ (s.o.), die die Punkte trennt: Betrachte $x, y \in \sigma(A)$, $x \neq y$. Dann erfüllt das Identitätspolynom $p = \mathrm{id}$ natürlich $p(x) \neq p(y)$.

Theorem 4.22 (Stetiger Funktionalkalkül für selbstadjungierte Operatoren).
Sei $A \in \mathcal{B}(\mathcal{H})$ ein selbstadjungierter beschränkter Operator. Dann gibt es einen eindeutigen Funktionalkalkül $\phi_A^{\mathit{stetig}} : C(\sigma(A)) \to \mathcal{B}(\mathcal{H})$, der stetig ist, also

$$\|\phi_A^{\mathit{stetig}}(f)\| \leq c\|f\|_\infty$$

für ein $c > 0$ und alle $f \in C(\sigma(A))$ erfüllt. Die Abbildung ϕ_A^{stetig} hat außerdem folgende Eigenschaften. Für alle $f \in C(\sigma(A))$ gilt:

- $\|\phi_A^{\mathit{stetig}}(f)\| = \|f\|_\infty$.

- *$\sigma(\phi_A^{stetig}(f)) = f(\sigma(A))$ (spektrales Abbildungsprinzip)*
- *Sei $B \in \mathcal{B}(\mathcal{H})$ ein mit A kommutierender Operator, $[A, B] = 0$. Dann kommutiert B auch mit $\phi_A^{stetig}(f)$ für alle $f \in C(\sigma(A))$.*
- *$\phi_A^{stetig}(f)$ ist normal.*

Beweis. Der polynomielle Funktionalkalkül

$$\phi_A^{\text{pol}} : \mathcal{P}(\sigma(A)) \to \mathcal{B}(\mathcal{H})$$

ist nach Proposition 4.18 c) eine lineare isometrische (insbesondere stetige) Abbildung, wenn $\mathcal{P}(\sigma(A))$ mit der Supremumsnorm $\|\cdot\|_{C(\sigma(A))}$ und $\mathcal{B}(\mathcal{H})$ mit der Operatornorm betrachtet wird. Nach dem Satz von Stone-Weierstraß ist $\mathcal{P}(\sigma(A)) \subset C(\sigma(A))$ ein dichter Unterraum. Also hat ϕ_A^{pol} eine eindeutige stetige lineare Fortsetzung auf $C(\sigma(A))$ (=die Vervollständigung von $\mathcal{P}(\sigma(A))$), siehe Satz 1.33 b). Wir nennen diese Fortsetzung

$$\phi_A^{\text{stetig}} : C(\sigma(A)) \to \mathcal{B}(\mathcal{H}).$$

Als stetige Fortsetzung einer Isometrie ist auch ϕ_A^{stetig} eine Isometrie. Dies wird unser stetiger Funktionalkalkül sein.

Konkret bedeutet das: Für eine stetige Funktion $f : \sigma(A) \to \mathbb{C}$ wählen wir eine Folge von Polynomfunktionen $p_n : \sigma(A) \to \mathbb{C}$, so dass $\sup_{\lambda\in\sigma(A)} |f(\lambda) - p_n(\lambda)| \to 0$ (diese Folge gibt es nach Stone-Weierstraß) und setzen dann $f(A) := \lim_n p_n(A)$. Dieser Limes existiert in Opertornorm und hängt nicht von der Wahl der approximierenden Polynomfolge $(p_n)_{n\in\mathbb{N}}$ ab.

Wir haben bereits gezeigt, dass ϕ_A^{stetig} isometrisch (und insbesondere stetig) ist. Jetzt betrachten wir die anderen Eigenschaften von ϕ_A^{stetig}. Um zu zeigen, dass ϕ_A^{stetig} ein Funktionalkalkül ist, betrachten wir die Eigenschaften a)–d) aus Definition 4.19. Davon sind a), d) und $\phi_A^{\text{stetig}}(\text{id}) = A$ klar, denn diese Eigenschaften liegen schon auf $\mathcal{P}(\sigma(A))$ fest. Für b) betrachten wir $f, g \in C(\sigma(A))$ mit approximierenden Polynomen (p_n) bzw (q_m). Dann gilt $p_n(\lambda)q_n(\lambda) \to f(\lambda)g(\lambda)$ gleichmäßig in λ (siehe Analyis 1), und damit

$$\phi_A^{\text{stetig}}(fg) = \lim_n \phi_A^{\text{pol}}(p_n q_n) = \lim_n \phi_A^{\text{pol}}(p_n) \cdot \lim_n \phi_A^{\text{pol}}(q_n) = \phi_A^{\text{stetig}}(f)\phi_A^{\text{stetig}}(g).$$

Für Eigenschaft c) eines Funktionalkalküls bemerken wir, dass auch $\overline{p_n} \to \overline{f}$ gleichmäßig gilt (komplexe Konjugation ändert Betrag nicht), und damit aufgrund der Stetigkeit der Adjunktion $B \mapsto B^*$

$$\phi_A^{\text{stetig}}(f)^* = \left(\lim_n \phi_A^{\text{pol}}(p_n)\right)^* = \lim_n \phi_A^{\text{pol}}(p_n)^* = \lim_n \phi_A^{\text{pol}}(\overline{p_n}) = \phi_A^{\text{stetig}}(\overline{f}).$$

Damit haben wir gezeigt, dass ϕ_A^{stetig} ein Funktionalkalkül ist. Er ist der eindeutige stetige Funktionalkalkül $C(\sigma(A)) \to \mathcal{B}(\mathcal{H})$, da ein Funktionalkalkül per Definition auf $\mathcal{P}(\sigma(A))$ eindeutig festgelegt ist, und die Stetigkeit auch die Eindeutigkeit der Ausdehnung auf $C(\sigma(A))$ festlegt.

Dies zeigt die Eindeutigkeit von ϕ_A^{stetig}. Als nächstes betrachten wir das spektrale Abbildungsprinzip, das wir für Polynomfunktionen bereits in Satz 4.18 a) gezeigt haben.

Sei dazu $\mu \notin f(\sigma(A))$. Dann ist $g := (f-\mu)^{-1} \in C(\sigma(A))$, und $g\cdot(f-\mu) = (f-\mu)\cdot g = 1$. Also gilt

$$g(A)(f(A)-\mu) = (g\cdot(f-\mu))(A) = 1(A) = 1,$$

und analog $(f(A)-\mu)g(A) = 1$. Also folgt $g(A) = (f(A)-\mu)^{-1}$ und insbesondere $\mu \notin \sigma(f(A))$. Dies zeigt die Inklusion $\sigma(f(A)) \subset f(\sigma(A))$. Für die umgekehrte Inklusion sei $\mu = f(\lambda)$ mit $\lambda \in \sigma(A)$. Wir wählen ein Polynom p_n mit $\|f - p_n\|_{C(\sigma(A))} \leq \frac{1}{n}$ und $|f(\lambda) - p_n(\lambda)| \leq \frac{1}{n}$. Da wir das spektrale Abbildungsprinzip für Polynome bereits kennen, gilt $p_n(\lambda) \in p_n(\sigma(A)) = \sigma(p_n(A)) = \sigma_{ap}(p_n(A))$ (hier haben wir Satz 4.14 benutzt). Wir finden also normierte Vektoren (v_n), so dass $\|(p_n(A) - p_n(\lambda))v_n\| \leq \frac{1}{n}$. Damit erhalten wir

$$\|(f(A)-\mu)v_n\| \leq \frac{2}{n} + \|(p_n(A)-p_n(\lambda))v_n\| \leq \frac{3}{n},$$

also $\mu \in \sigma_{ap}(f(A)) = \sigma(f(A))$.

Die letzten beiden Aussagen sind leicht zu zeigen. Erfüllt $B \in \mathcal{B}(\mathcal{H})$ die Gleichung $AB = BA$, so gilt natürlich auch $p(A)B = Bp(A)$ für jedes Polynom p. Per Approximation folgt dann auch $f(A)B = Bf(A)$ für alle $f \in C(\sigma(A))$.

Da die Funktionen $f(A) = \phi_A^{\text{stetig}}(f)$ alle Operatornorm-Grenzwerte von Polynomen $p(A)$ sind, und die $p(A)$ alle normal sind, kommutieren $f(A)$ und $g(A)$ für beliebige $f, g \in C(\sigma(A))$. Insbesondere kommutiert $f(A)$ mit $f(A)^* = \overline{f}(A)$, dh $f(A)$ ist normal. ☐

Zeigen Sie, dass im Kontext von diesem Theorem auch $f(\sigma_p(A)) \subset \sigma_p(f(A))$ gilt. Wieso gilt im Allgemeinen nicht Gleichheit?

Der stetige Funktionalkalkül hat viele Anwendungen. Wir werden im Folgenden meist die Schreibweise $f(A)$ anstelle von $\phi_A^{\text{stetig}}(f)$ verwenden. Beachten Sie, dass Ausdrücke wie e^A, $\sin(A)$, $|A|$, $\sqrt{|A|}$, etc jetzt eine wohldefinierte Bedeutung haben (für $A = A^* \in \mathcal{B}(\mathcal{H})$). Wir betrachten einige Beispiele.

Beispiel 4.23. Sei $A = A^* \in \mathcal{B}(\mathcal{H})$. Dann ist $U = e^{iA}$ unitär.

Beweis. Wir betrachten die Funktion $f(\lambda) = e^{i\lambda}$, die auf ganz $\mathbb{C}$ und damit insbesondere auf $\sigma(A) \subset [-\|A\|, \|A\|]$ stetig ist. Für reelles λ gilt offenbar $\overline{f}(\lambda) = e^{-i\lambda} = (f(\lambda))^{-1}$, also $f\overline{f} = 1 = \overline{f}f$. Damit folgt

$$UU^* = f(A)f(A)^* = (f\overline{f})(A) = 1(A) = 1,$$

und $U^*U = 1$ in völlig analoger Art und Weise. Also ist U unitär. ☐

Als nächstes Beispiel betrachten wir eine Funktion $f : \sigma(A) \to \mathbb{C}$, die durch eine konvergente Potenzreihe gegeben ist. In diesem Fall konvergiert auch die Operatorreihe (bei geeignetem Konvergenzradius, genau wie bei unserer Diskussion der Neumann-Reihe). Wir zeigen, dass die Operatorreihe in diesem Fall mit dem stetigen Kalkül übereinstimmt.

Beispiel 4.24. Sei $A = A^* \in \mathcal{B}(\mathcal{H})$ ein selbstadjungierter Operator und $f(z) = \sum_{n=0}^{\infty} c_n\, z^n$ eine Potenzreihe mit Konvergenzradius $r > r_A$ (dh insbesondere konvergiert die Reihe für alle $z \in \sigma(A)$ absolut). Dann konvergiert auch $\sum_{n=0}^{\infty} c_n\, A^n$ in $\mathcal{B}(\mathcal{H})$, und es gilt

$$f(A) = \sum_{n=0}^{\infty} c_n\, A^n.$$

Zum *Beweis* bemerken wir zunächst, dass wir aufgrund von $r > r_A = \|A\|$

$$\left\| \sum_{n=N}^{M} c_n A^n \right\| \to 0, \qquad N, M \to \infty$$

haben. Also ist $\left(\sum_{n=0}^{N} c_n A^n\right)_{N\in\mathbb{N}}$ ist eine Cauchyfolge, deren Grenzwert in $\mathcal{B}(\mathcal{H})$ wir mit $\sum_{n=0}^{\infty} c_n A^n$ bezeichnen.

Um die Übereinstimmung dieses Operators mit $f(A)$ zu zeigen, betrachten wir die Polynome $p_N(z) := \sum_{n=0}^{N} c_n z^n$, die auf $\sigma(A)$ gleichmäßig gegen f konvergieren (siehe Analysis 1). Per Definition des stetigen Funktionalkalküls gilt $f(A) = \lim_N p_N(A) = \sum_{n=0}^{\infty} c_n A^n$.

Eine interessante Anwendung des stetigen Funktionalkalküls, die über Potenzreihen hinausgeht, werden wir im nächsten Abschnitt bei der Betrachtung der Betragsfunktion $f(z) = |z|$ kennenlernen. In dem Zusammenhang werden wir auch folgendes Resultat über Komposition von stetigen Funktionen benötigen:

Lemma 4.25. *Sei $A = A^* \in \mathcal{B}(\mathcal{H})$ und $g \in C(\sigma(A))$ reell, sowie $f \in C(\sigma(g(A)))$. Dann gilt*

$$f(g(A)) = (f \circ g)(A).$$

Der Beweis dieses Lemmas erfolgt in den Übungen.

Zum Abschluss dieses Abschnittes wenden wir uns jetzt der Frage zu, wie der stetige Funktionalkalkül von *selbstadjungierten* Operatoren auf allgemeine *normale* Operatoren A ausgedehnt werden kann. Im Unterschied zu selbstadjungierten Operatoren haben normale Operatoren im Allgemeinen komplexes Spektrum (d.h. $\sigma(A) \not\subset \mathbb{R}$). Wenn wir eine kompakte Teilmenge $X \subset \mathbb{C}$ betrachten, erfüllt die Polynomalgebra $\mathcal{A} = \mathcal{P}(X)$ aber nicht die Voraussetzungen des Satzes von Stone-Weierstraß, da sie nicht unter komplexer Konjuation abgeschlossen ist.

Wir müssen unsere Approximationsidee also modifizieren und verwenden statt $\mathcal{P}(\sigma(A))$ die Algebra $\mathcal{P}_{\mathbb{R}^2}(\sigma(A))$, definiert als die Menge aller Polynome in Real- und Imaginärteil von $z \in \sigma(A)$, mit komplexen Koeffizienten. D.h. Elemente von $\mathcal{P}_{\mathbb{R}^2}(\sigma(A))$ sind von der Form

$$p(z) = \sum_{n=0}^{N} \sum_{m=0}^{M} c_{n,m} \mathrm{Re}(z)^n \mathrm{Im}\,(z)^m\,, \qquad z \in \sigma(A).$$

Die Funktionenalgebra $\mathcal{A} = \mathcal{P}_{\mathbb{R}^2}(\sigma(A))$ und die Menge $X = \sigma(A)$ erfüllen die Voraussetzungen des Satzes von Stone-Weierstraß. Also lässt sich jede stetige Funktion $f \in C(\sigma(A))$ gleichmäßig durch Polynome in $\mathcal{P}_{\mathbb{R}^2}(\sigma(A))$ approximieren.

Die Idee, einen normalen Operator A in einem solchen Polynom auszuwerten, basiert auf einer Operatorversion von Real- und Imaginärteil. Zunächst bemerken wir, dass *jeder* Operator $A \in \mathcal{B}(\mathcal{H})$ in selbstadjungierte "Real- und Imaginärteile" zerlegt werden kann, nämlich

$$A = A_+ + iA_-, \qquad A_+ = \frac{1}{2}(A + A^*), \quad A_- = \frac{1}{2i}(A - A^*), \qquad {A_\pm}^* = A_\pm.$$

Für *normales* A gilt $AA^* = A^*A$, so dass A_+ und A_- kommutieren. Wir definieren einen polynomiellen Funktionalkalkül für normale Operatoren also gemäß

$$\phi_A^{\text{pol,normal}}(p) = \sum_{n=0}^{N}\sum_{m=0}^{M} c_{n,m} A_+^n A_-^m \qquad \text{für } p(z) = \sum_{n=0}^{N}\sum_{m=0}^{M} c_{n,m} \mathrm{Re}(z)^n \mathrm{Im}\,(z)^m.$$

Aufgrund der Normalität von A kommutieren A_+ und A_-. Das hat zur Folge, dass $\phi_A^{\text{pol,normal}}(f)$ mit $\phi_A^{\text{pol}}(f)$ übereinstimmt, falls f ein Polynom ist; denn die Faktoren A_+^n und A_-^m können genau wie $(\mathrm{Re}\, z)^n$ und $(\mathrm{Im}\, z)^m$ beliebig vertauscht werden.

Um nun in Analogie zum selbstadjungierten Fall vorzugehen, benötigen wir wieder ein spektrales Abbildungsprinzip:

Satz 4.26. *Sei $A \in \mathcal{B}(\mathcal{H})$ normal, $f : \sigma(A) \to \mathbb{C}$ ein Polynom in zwei Variablen, und $\phi_A^{\text{pol,normal}}(f) = f(A)$ wie oben definiert. Dann gilt das spektrale Abbildungsprinzip,*

$$\sigma(f(A)) = f(\sigma(A)).$$

Im Gegensatz zum Fall von Polynomen in einer Variablen ist der Beweis dieses Satzes deutlich komplizierter. Er kann entweder mit Hilfe der Theorie kommutativer C^*-Algebren gegeben werden (siehe z.B. [Rud91]) oder mit Hilfe des Kalküls für selbstadjungierte Operatoren (siehe z.B. [Hal13, Thm. 10.23]). Wir werden den Beweis von dieser Version des spektralen Abbildungsprinzips später aus einem Spektralsatz erhalten und vorerst ohne Beweis akzeptieren.

Dann ist die weitere Entwicklung des stetigen Kalküls für normale Operatoren einfach. Da wir bereits wissen, dass für normale Operatoren Spektralradius und Norm übereinstimmen, erhalten wir wieder

$$\|f(A)\| = \|f\|_{C(\sigma(A))}$$

für normales A und $f(z) = p(\mathrm{Re}\, z, \mathrm{Im}\, z)$ ein Polynom in $\mathrm{Re}\, z, \mathrm{Im}\, z$. Damit erhalten wir eine eindeutige Isometrie

$$\phi_A^{\text{pol,normal}} : \mathcal{P}_{\mathbb{R}^2}(\sigma(A)) \to \mathcal{B}(\mathcal{H}).$$

Es ist klar, dass $\phi_A^{\text{pol,normal}}$ ein Algebrenhomomorphismus ist. Für das Verhalten unter Adjunktion gehen wir wie folgt vor. Sei $f(z) = p(\mathrm{Re}\, z, \mathrm{Im}\, z) = \sum_{n=0}^{N}\sum_{m=0}^{M} c_{n,m}(\mathrm{Re}\, z)^n(\mathrm{Im}\, z)^m$, dann gilt $\bar{f}(z) = \sum_{n=0}^{N} \sum_{m=0}^{M} \overline{c_{n,m}} (\operatorname{Re} z)^n (\operatorname{Im} z)^m$, und

$$\begin{aligned}
\phi_A^{\text{pol,normal}}(f)^* &= \left(\sum_{n=0}^{N} \sum_{m=0}^{M} c_{n,m} (A_+)^n (A_-)^m \right)^* \\
&= \sum_{n=0}^{N} \sum_{m=0}^{M} \overline{c_{n,m}} (A_-)^m (A_+)^n \\
&= \sum_{n=0}^{N} \sum_{m=0}^{M} \overline{c_{n,m}} (A_+)^n (A_-)^m \\
&= \phi_A^{\text{pol,normal}}(\bar{f}).
\end{aligned}$$

Mit Anwendung des Satzes von Stone-Weierstraß erhalten wir dann eine eindeutige stetige Fortsetzung zu einem (isometrischen, insbesondere also stetigen) Funktionalkalkül

$$\phi_A^{\text{stetig,normal}} : C(\sigma(A)) \to \mathcal{B}(\mathcal{H}).$$

Dies ist unser bisher allgemeinster Funktionalkalkül. Er enthält den polynomiellen Kalkül für normale Operatoren und den stetigen (und insbesondere polynomiellen) Kalkül für selbstadjungierte Operatoren als Spezialfälle.

Wir fassen zusammen:

Theorem 4.27. *Sei $A \in \mathcal{B}(\mathcal{H})$ normal. Dann existiert ein eindeutiger stetiger Funktionalkalkül*

$$\phi_A^{\textit{stetig,normal}} : C(\sigma(A)) \to \mathcal{B}(\mathcal{H}).$$

Dieser Kalkül ist isometrisch und erfüllt das spektrale Abbildungsprinzip. Die Operatoren $\phi_A^{\textit{stetig,normal}}(f) = f(A)$ sind für alle $f \in C(\sigma(A))$ normal.

Da der Beweis sehr ähnlich zum selbstadjungierten Fall ist, verzichten wir auf eine detaillierte Diskussion.

4.4 Positive Operatoren und Polarzerlegung

Als Anwendung des stetigen Kalküls besprechen wir die *Polarzerlegung* von Operatoren, dies wird eine Zerlegung von Operatoren analog zur Zerlegung $z = |z| e^{i \arg(z)}$ einer komplexen Zahl z in Betrag $|z|$ und Phase $e^{i \arg(z)}$ sein. Als Vorbereitung dafür diskutieren wir *positive* Operatoren.

Satz 4.28. *Sei $A \in \mathcal{B}(\mathcal{H})$. Dann sind die folgenden drei Aussagen äquivalent:*

a) $A = A^*$ *und* $\sigma(A) \subset [0, \infty)$.

b) *Es gibt* $B \in \mathcal{B}(\mathcal{H})$ *mit* $A = B^* B$.

c) $\langle v, Av \rangle \geq 0$ *für alle* $v \in \mathcal{H}$.

Beweis. $a) \Rightarrow b)$ Wegen $\sigma(A) \subset [0, \infty)$ ist $\sqrt{\cdot} : \sigma(A) \to \mathbb{R}$ eine wohldefinierte stetige Funktion. Mit Hilfe des stetigen Kalküls definieren wir $B := \sqrt{A}$. Da $\sqrt{\cdot}$ reell ist, gilt

$B^* = (\sqrt{A})^* = \sqrt{A} = B$. Weiterhin gilt $(\sqrt{\lambda})^2 = \lambda$ für alle $\lambda \geq 0$. Da der stetige Funktionalkalkül Kompositionen respektiert (dh $(f \circ g)(A) = f(g(A))$ für $f \in C(\sigma(g(A)))$, siehe Lemma 4.25), gilt $B^2 = (\sqrt{A})^2 = A$, also $A = B^*B$.

$b) \Rightarrow c)$ folgt sofort aus $\langle v, B^*Bv\rangle = \|Bv\|^2 \geq 0$.

$c) \Rightarrow a)$: Wir betrachten die Sesquilinarform

$$s : \mathcal{H} \times \mathcal{H} \to \mathbb{C}, \qquad s(v, w) := \langle v, Aw\rangle,$$

die nach Voraussetzung c) positiv semidefinit ist. Also ist sie hermitesch, das heißt

$$\langle v, Aw\rangle = s(v, w) = \overline{s(w, v)} = \overline{\langle w, Av\rangle} = \langle Av, w\rangle = \langle v, A^*w\rangle$$

für alle $v, w \in \mathcal{H}$. Also gilt $A = A^*$, was $\sigma(A) \subset \mathbb{R}$ impliziert. Sei $\lambda \in \mathbb{R}$, $\lambda < 0$. Wir müssen $\lambda \in \rho(A)$ zeigen. Dazu schätzen wir für beliebiges $v \in \mathcal{H}$ ab:

$$\langle v, (A - \lambda)v\rangle = \langle v, Av\rangle - \lambda\|v\|^2 \geq |\lambda|\|v\|^2.$$

Andererseits gilt $0 \leq \langle v, (A - \lambda)v\rangle \leq \|(A - \lambda)v\|\|v\|$, so dass wir $|\lambda|\|v\| \leq \|(A - \lambda)v\|$ erhalten. Da $|\lambda| > 0$, sehen wir, dass λ nicht im approximativen Punktspektrum liegen kann. Aber für selbstadjungiertes A ist das approximative Punktspektrum das ganze Spektrum. Also $\lambda \in \rho(A)$. ☐

Definition 4.29. Ein Operator $A \in \mathcal{B}(\mathcal{H})$ heißt *positiv*, falls er die Bedingungen in der vorherigen Proposition erfüllt. Wir schreiben dann auch $A \geq 0$.

Bemerkung: Die positiven Operatoren definieren eine partielle Ordnung auf der Menge der selbstadjungierten Operatoren durch $A \geq B :\Leftrightarrow A - B \geq 0$.

Definition und Satz 4.30. Sei $A \in \mathcal{B}(\mathcal{H})$. Dann ist der Betrag von A definiert als

$$|A| := \sqrt{A^*A}.$$

$|A|$ ist der eindeutige positive Operator $X \geq 0$, der $X^2 = A^*A$ erfüllt.

Beweis. Beachten Sie, dass A hier nicht als normal vorausgesetzt ist. Trotzdem ist A^*A selbstadjungiert und sogar positiv, so dass $\sqrt{\cdot}$ wohldefiniert und stetig auf $\sigma(A^*A) \subset [0, \infty)$ ist. Damit ist $|A|$ definiert im Sinne des stetigen Funktionalkalküls von A^*A.

$|A|$ ist normal per Funktionalkalkül und erfüllt $\sigma(|A|) = \{\sqrt{\lambda^*\lambda} : \lambda \in \sigma(A)\} \subset [0, \infty)$. Damit ist $|A|$ selbstadjungiert (vgl Hausaufgaben) und angesichts unserer Charakterisierungen von Positivität positiv, $|A| \geq 0$. Wie bereits oben argumentiert, gilt auch $|A|^2 = A^*A$.

Sei nun $X \geq 0$ ein positiver Operator mit $X^2 = A^*A \geq 0$. Dann gilt wegen Lemma 4.25 $|A| = \sqrt{A^*A} = \sqrt{X^2} = X$. Dies zeigt die Eindeutigkeit. ☐

Nachdem wir den Betrag eines Operators verstanden haben, wenden wir uns der Phase zu. Die Phase $\varphi : z \mapsto e^{i \arg(z)} = z/|z|$ ist unstetig bei $z = 0$, so dass wir $\varphi(A)$ nicht mit Hilfe des stetigen Kalküls definieren können[9]. Wir gehen deshalb etwas anders vor und erinnern zunächst an den Begriff der partiellen Isometrie (Hausaufgabe H5.3).

[9] Falls A invertierbar ist, können wir den Funktionalkalkül verwenden.

Definition 4.31. Eine partielle Isometrie ist ein Operator $V \in \mathcal{B}(\mathcal{H})$ mit der Eigenschaft, dass V^*V eine orthogonale Projektion ist.

In der Hausaufgabe haben Sie bereits gezeigt, dass mit V auch V^* eine partielle Isometrie ist, dh auch VV^* ist eine orthogonale Projektion, und dass $V^*VV^* = V^*$ gilt.

Es ist allerdings noch nicht klar, inwiefern der Name "partielle Isometrie" gerechtfertigt ist, der ja eine Nähe zu dem Begriff einer Isometrie U (dh. $\|Ux\| = \|x\|$ für alle $x \in \mathcal{H}$) andeutet. Um dies zu verstehen, betrachten wir das folgende Lemma.

Lemma 4.32. *Sei $V \in \mathcal{B}(\mathcal{H})$ eine partielle Isometrie.*

a) $\operatorname{Ran} V$ *ist abgeschlossen.*

b) V^*V *ist die orthogonale Projektion auf* $(\ker V)^\perp = \operatorname{Ran}(V^*)$*, und* VV^* *ist die orthogonale Projektion auf* $\operatorname{Ran}(V) = (\ker V^*)^\perp$.

c) Die eingeschränkte Abbildung $V : (\ker V)^\perp \to \operatorname{Ran} V$ *ist eine Isometrie mit Inversem* $V^* : \operatorname{Ran} V \to (\ker V)^\perp$.

Ein $V \in \mathcal{B}(\mathcal{H})$ ist genau dann eine partielle Isometrie, wenn $V|_{(\ker V)^\perp}$ eine Isometrie ist.

Man nennt V^*V die *initiale* und VV^* die *finale* Projektion von V, da V als Isometrie von $\operatorname{Ran} V^* = \operatorname{Ran}(V^*V)$ nach $\operatorname{Ran} V = \operatorname{Ran}(VV^*)$ angesehen werden kann.

Beweis. a) Wir haben $\operatorname{Ran}(VV^*) \subset \operatorname{Ran}(V)$ und wegen $V = VV^*V$ auch $\operatorname{Ran}(V) = \operatorname{Ran}(VV^*V) \subset \operatorname{Ran}(VV^*)$. Also $\operatorname{Ran}(V) = \operatorname{Ran}(VV^*)$, und letzterer Raum ist als Bild einer orthogonalen Projektion abgeschlossen.

b) folgt sofort aus a), der allgemeinen Beziehung $\operatorname{Ran} V = (\ker V^*)^\perp$ (da $\operatorname{Ran} V$ abgeschlossen), und der Tatsache, dass mit V auch V^* eine partielle Isometrie ist.

c) Sei $x \in (\ker V)^\perp = \operatorname{Ran} V^*$, also $x = V^*y$. Dann gilt

$$\|VV^*y\|^2 = \langle V^*y, V^*VV^*y\rangle = \langle V^*y, V^*y\rangle = \|V^*y\|^2.$$

Also ist $V : (\ker V)^\perp \to \operatorname{Ran} V$ isometrisch. Die Aussage über das Inverse folgt aus der bereits bekannten Gleichung $V^*VV^*x = V^*x$, $x \in \mathcal{H}$.

Bzgl. der letzten Behauptung haben wir bereits gezeigt, dass jede partielle Isometrie V die Eigenschaft hat, dass $V|_{(\ker V)^\perp}$ eine Isometrie ist. Nehmen wir nun $V \in \mathcal{B}(\mathcal{H})$ mit $V|_{(\ker V)^\perp}$ Isometrie. Dann müssen wir zeigen, dass $P := V^*V$ eine orthogonale Projektion ist. $P^* = P$ ist klar. Um auch $P^2 = P$ einzusehen, zerlegen wir den Hilbertraum als $\mathcal{H} = \ker V \oplus (\ker V)^\perp$. Für $x \in (\ker V)^\perp$ gilt dann nach Voraussetzung

$$\langle x, Px\rangle = \langle Vx, Vx\rangle = \|Vx\|^2 = \langle x, x\rangle,$$

also $Px = x + y$ mit $y \perp x$. Aber die obige Rechnung impliziert $\|P\| \leq 1$, was auf $y = 0$ führt. Also gilt $Px = x$ und insbesondere $P^2x = Px$ für $x \in (\ker V)^\perp$. Aber für $x \in \ker V$ gilt sowieso $Px = V^*Vx = 0$ und somit $P^2x = P0 = 0 = Px$. Also gilt $P^2 = P$. □

Eine partielle Isometrie muss also weder injektiv noch surjektiv sein, bildet aber den abgeschlossenen Unterraum $(\ker V)^\perp$ isometrisch und bijektiv auf den abgeschlossenen Unterraum $\operatorname{Ran} V$ ab.

Beispiel 4.33. Sei $\mathcal{H} = L^2(\mathbb{R}_+)$ und $(Vf)(x) := f(x+1)$, $x \geq 0$, der Linksshift. Dann gilt $\ker V = L^2([0,1]) = \{f \in L^2(\mathbb{R}_+) : \operatorname{supp} f \subset [0,1]\}$, dh V ist nicht injektiv und sicher keine Isometrie. Aber V ist eine partielle Isometrie: Eingeschränkt auf $(\ker V)^\perp = L^2([1,\infty)) = \{f \in L^2(\mathbb{R}_+) : \operatorname{supp}(f) \subset [1,\infty)\}$ wird V isometrisch, denn

$$\|Vf\|^2 = \int_0^\infty |f(x+1)|^2 dx = \int_1^\infty |f(x)|^2 dx \overset{\operatorname{supp} f \subseteq [1,\infty)}{=} \int_0^\infty |f(x)|^2 dx = \|f\|^2.$$

In diesem Fall ist die Initialprojektion die orthogonale Projektion auf $L^2([1,\infty) = \mathcal{H}_{\text{in}}$, und die finale Projektion die orthogonale Projektion auf $L^2(\mathbb{R}_+) = \mathcal{H}_{\text{fin}}$, dh $VV^* = 1$ (weil V surjektiv ist).

Wir zeigen jetzt, wie partielle Isometrien bei der Polarzerlegung auftreten.

Satz 4.34. (Polarzerlegung) *Sei $A \in \mathcal{B}(\mathcal{H})$. Dann gibt es eine partielle Isometrie $V \in \mathcal{B}(\mathcal{H})$, so dass*

$$A = V|A|$$

gilt. Die partielle Isometrie V ist eindeutig durch $\ker V = \ker A$ *festgelegt. Weiterhin gilt* $\operatorname{Ran} V = \overline{\operatorname{Ran} A}$.

Beweis. Per Definition des Betrags $|A|$ von A gilt für beliebiges $v \in \mathcal{H}$

$$\||A|v\|^2 = \langle |A|v, |A|v\rangle = \langle v, A^*Av\rangle = \langle Av, Av\rangle = \|Av\|^2.$$

Wir können also eine lineare Abbildung $V : \operatorname{Ran}|A| \to \operatorname{Ran} A$ definieren via

$$V|A|v = Av, \qquad v \in \mathcal{H},$$

dies ist wohldefiniert und isometrisch nach obiger Bemerkung. Also gibt es eine isometrische Ausdehnung $\overline{\operatorname{Ran}|A|} \to \overline{\operatorname{Ran} A}$, die wir weiterhin mit V bezeichnen.

Auf dem orthogonalen Komplement $\overline{\operatorname{Ran}|A|}^\perp = \ker|A|^* = \ker|A|$ definieren wir V als 0, d.h. $\ker V = \ker|A| = \ker A$. Damit ist V eine partielle Isometrie, und es gilt die Polarzerlegung $A = V|A|$.

Sei V' eine weitere partielle Isometrie, die $A = V'|A|$ und $\ker V' = \ker A$ erfüllt. Dann gilt $V'|A| = V|A|$, also $V'x = Vx$ für alle $x \in \overline{\operatorname{Ran}|A|} = (\ker|A|)^\perp = (\ker A)^\perp$. Wegen $\ker(V') = \ker(A) = \ker(V)$ gilt auch $V'x = 0 = Vx$ für $x \in \ker A$, also $V' = V$. □

Ist A invertierbar, so ist die partielle Isometrie V aus seiner Polarzerlegung sogar unitär: Denn dann gilt $\ker V = \ker A = \{0\}$ und $\operatorname{Ran} V = \overline{\operatorname{Ran} A} = \mathcal{H}$.

Beispiel 4.35. Sei $\mathcal{H} = \mathbb{C}^n$ und $A \in GL(n)$ eine invertierbare $(n \times n)$-Matrix. Dann ist A^*A eine positiv definite Matrix, die wir diagonalisieren können. Es gibt also eine unitäre Matrix U, so dass

$$UA^*AU^* = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

wobei die $\lambda_k > 0$ die Eigenwerte von A^*A sind. Wir behaupten

$$U|A|U^* = \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix} =: D.$$

Tatsächlich gilt $D \geq 0$ und damit $U^*DU \geq 0$, sowie $(U^*DU)^2 = U^*D^2U = A^*A$ (s.o.). Gemäßt Satz 4.30 folgt $U^*DU = |A|$ wie behauptet.

Jetzt sehen wir leicht, wie die Polarzerlegung von A aussieht: Wir suchen V (in diesem Fall unitär) mit $V|A| = A$, also

$$V = A|A|^{-1} = AU^*D^{-1}U.$$

Beispiele von Polarzerlegungen auf unendlichdimensionalen Räumen werden in den Übungen besprochen.

Hier einige weitere Bemerkungen zu Beträgen und positiven Operatoren:

a) Für $A \geq 0$ ist auch $BAB^* \geq 0$ für alle $B \in \mathcal{B}(\mathcal{H})$ (leichte Übung).

b) Im Allgemeinen sind die folgenden Gleichungen allesamt **falsch**: i) $|AB| = |A||B|$, ii) $|A^*| = |A|$, iii) $|A + B| \leq |A| + |B|$.

Es ist eine interessante Frage, welche Funktionen f operatormonoton sind in dem Sinne, dass sie $A \leq B \Rightarrow f(A) \leq f(B)$ erfüllen (mit f stetige Funktion auf den Spektren von A und B). Offenbar müssen operatormonotone Funktionen monoton sein, denn wir können A, B als Vielfache der Identität wählen. Aber längst nicht jede monotone Funktion ist auch operatormonoton! Zum Beispiel ist $x \mapsto x^2$ nicht operatormonoton, aber $x \mapsto -x^{-1}$ ist es. Eine volle Charakterisierung ist als Löwner's Theorem bekannt.

4.5 Spektralmaße und der messbare Kalkül

Ziel dieses Abschnittes ist es, den stetigen Funktionalkalkül noch einmal entscheidend zu erweitern, also die Algebra der betrachteten Funktionen $C(\sigma(A))$ zu vergößern. Als Motivation, wieso eine solche Erweiterung wünschenswert ist, beweisen wir den Spektralsatz der Linearen Algebra.

Satz 4.36 (Spektralsatz für normale Operatoren auf endlichdimensionalen Hilberträumen). *Sei $\mathcal{H}$ ein endlichdimensionaler Hilbertraum und $A \in \mathcal{B}(\mathcal{H})$ normal. Dann gibt es eine eindeutige Familie von Operatoren $\{P_\lambda : \lambda \in \sigma(A)\}$ mit den Eigenschaften*

a) Jedes P_λ ist eine orthogonale Projektion,

b) $P_\lambda P_\mu = 0$ für $\lambda \neq \mu$, *c)* $\sum_{\lambda\in\sigma(A)} P_\lambda = 1$,

d) $A = \sum_{\lambda\in\sigma(A)} \lambda P_\lambda$.

Beweis. Zu jedem Eigenwert $\lambda \in \sigma(A)$ betrachten wir die charakteristische Funktion

$$\chi_{\{\lambda\}} : \sigma(A) \to \mathbb{C}, \qquad \mu \mapsto \delta_{\lambda,\mu}.$$

Da $\sigma(A)$ aus endlich vielen Punkten besteht, ist dies eine stetige Funktion, ja sogar Einschränkung eines Polynoms auf $\sigma(A)$, nämlich

$$\chi_{\{\lambda\}}(\mu) = \prod_{\substack{\nu\in\sigma(A)\\ \nu\neq\lambda}} \frac{\mu-\nu}{\lambda-\nu} = \delta_{\lambda,\mu}.$$

Wir können also den polynomiellen Funktionalkalkül verwenden, um die Operatoren $P_\lambda := \chi_{\{\lambda\}}(A)$ zu studieren: Wegen $\chi_{\{\lambda\}}^2 = \chi_{\{\lambda\}} = \overline{\chi_{\{\lambda\}}}$ gilt $P_\lambda^2 = P_\lambda = P_\lambda^*$, dh P_λ ist eine orthogonale Projektion. Analog folgen die behaupteten Eigenschaften b),c),d) aus $\chi_{\{\lambda\}}\chi_{\{\mu\}} = 0$ für $\lambda \neq \mu$ sowie

$$1 = \sum_{\lambda\in\sigma(A)} \chi_{\{\lambda\}}, \qquad \mathrm{id}_{\sigma(A)} = \sum_{\lambda\in\sigma(A)} \lambda\cdot\chi_{\{\lambda\}}.$$

Eindeutigkeit: Wir behaupten, dass die Bilder der Projektionen P_λ die Eigenräume sind, dh Ran $P_\lambda = \ker(A-\lambda)$, was die Eindeutigkeit dann zeigt. Dazu bemerken wir, dass ein Vektor $\psi \in \mathcal{H}$ genau dann in $\ker(A-\lambda)$ (für ein $\lambda \in \sigma(A)$) liegt, wenn

$$\begin{aligned}
(A-\lambda)\psi &= 0 \\
\Leftrightarrow P_\mu(A-\lambda)\psi &= 0 \qquad \forall\mu\in\sigma(A) \\
\Leftrightarrow (\mu-\lambda)P_\mu\psi &= 0 \qquad \forall\mu\in\sigma(A) \\
\Leftrightarrow P_\mu\psi &= 0 \qquad \forall\mu\in\sigma(A)\backslash\{\lambda\} \\
\Leftrightarrow \psi &= P_\lambda\psi,
\end{aligned}$$

also $\ker(A-\lambda) = \mathrm{Ran}(P_\lambda)$. □

Die Darstellung

$$A = \sum_{\lambda\in\sigma(A)} \lambda P_\lambda$$

wird *Spektraldarstellung* von A genannt und ist in Theorie und Anwendungen von zentraler Bedeutung. Sie ist die basisunabhängige Formulierung der Aussage, dass sich normale Endomorphismen auf komplexen Hilberträumen diagonalisieren lassen. Auf Grundlage der Spektraldarstellung sieht man sofort

$$f(A) = \sum_{\lambda\in\sigma(A)} f(\lambda)P_\lambda$$

für jedes Polynom f, was das spektrale Abbildungsprinzip besonders deutlich hervortreten lässt.

Wir hätten deshalb auch gerne Spektraldarstellungen im unendlichdimensionalen Fall (einige Modifikationen gegenüber der endlichdimensionalen Situation werden wir dabei allerdings in Kauf nehmen müssen). Die wesentliche Idee des Beweises des Spektralsatzes war, die *Spektralprojektionen* P_λ als charakeristische Funktionen $P_\lambda = \chi_{\{\lambda\}}(A)$ zu erkennen. Für die Verallgemeinerung auf unendlichdimensionale Hilberträume werden wir auch einen Funktionalkalkül brauchen, der charakeristische Funktionen auf $\sigma(A)$ beinhaltet. Diese sind im Allgemeinen nicht stetig, so dass es einer Erweiterung des stetigen Kalküls bedarf.

Um den Überblick über die kommenden Entwicklungen zu behalten, erläutern wir kurz die grobe Idee der Erweiterung des Funktionalkalküls: Wir betrachten selbstadjungiertes $A \in \mathcal{B}(\mathcal{H})$, zwei beliebige Vektoren $v, w \in \mathcal{H}$, und die Abbildung

$$C(\sigma(A)) \ni f \longmapsto \langle v, \phi_A^{\text{stetig}}(f)w\rangle \in \mathbb{C}.$$

Dies ist ein stetiges lineares Funktional, also ein Element von $C(\sigma(A))'$. Wir werden gleich einen Satz besprechen, der besagt, dass der Dualraum von $C(\sigma(A))$ isometrisch isomorph zu einem Banachraum von (komplexen) Maßen ist. Das bedeutet, dass es ein eindeutiges (komplexes) Borelmaß $\mu_{v,w}^A$ auf $C(\sigma(A))$ gibt, so dass

$$\langle v, f(A)w\rangle = \int\limits_{\sigma(A))} f(\lambda)\, d\mu_{v,w}^A(\lambda)$$

für alle $f \in C(\sigma(A))$ gilt. Die rechte Seite macht aber nicht nur für stetige, sondern auch für beschränkte messbare Funktionen f Sinn, für die wir dann die linke Seite (mit messbarem f) als die rechte Seite *definieren* werden. Dies wird einen Operator $f(A)$ festlegen (siehe Lemma 2.19) und den messbaren Kalkül definieren.

Soweit der Überblick, nun zu den Details. Wir werden einige maßtheoretische Hintergrundresultate aus Zeitgründen nicht beweisen.

Definition 4.37. Sei Ω eine Menge und Σ eine σ-Algebra auf Ω. Ein *signiertes (bzw komplexes) Maß* ist eine σ-additive Abbildung $\mu : \Sigma \to \mathbb{R}$ (bzw $\mu : \Sigma \to \mathbb{C}$), dh für jede Folge paarweise disjunkter Mengen $A_n \in \Sigma$, $n \in \mathbb{N}$, gilt

$$\mu\left(\bigsqcup_{n\in\mathbb{N}} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n).$$

Die Menge aller komplexen Maße auf Ω wird mit $M(\Omega, \Sigma)$ bezeichnet.

Im Unterschied zur Definition eines *Maßes* muss ein signiertes bzw komplexes Maß nicht nur nicht-negative Werte annehmen. Andererseits nimmt ein signiertes/komplexes Maß nur endliche Werte an, nicht $\pm\infty$. Deshalb gilt, wie für Maße, $\mu(\emptyset) = 0$. Weiterhin gilt, da Σ unter abzählbaren Vereinigungen abgeschlossen ist, die Konvergenz von $\sum_{n=1}^{\infty} \mu(A_n)$ als Bestandteil der Definition.

Beispiel 4.38. Beispiele von komplexen Maßen sind:

a) $\mu = c_1\,\delta_{x_1} + \ldots + c_n\,\delta_{x_n}$ mit $c_1, \ldots, c_n \in \mathbb{C}$ und $x_1, \ldots, x_n \in \Omega$ (Hier bezeichnet δ_x das Dirac-Maß bei x)

b) Für $\Omega = \mathbb{R}$ mit der Borel-σ-Algebra und $f \in L^1(\mathbb{R})$ ist $\mu(A) := \int_A f(x)dx$ (Lebesgue-Integral) ein komplexes Maß.

Da reelle (komplexe) Linearkombinationen von signierten (komplexen) Maßen wieder signierte (komplexe) Maße sind, bilden signierte (komplexe) Maße einen Vektorraum. Jedem solchen "Maß" μ lässt sich ein echtes, dh positives, Maß $|\mu|$ zuordnen via

$$|\mu|(A) := \sup_Z \sum_{E \in Z} |\mu(E)|,$$

wobei das Supremum über alle Zerlegungen von A in endlich viele paarweise disjunkte messbare Mengen läuft. $|\mu|$ ist ein endliches Maß (ohne Beweis). In den oben angeführten Beispielen erhält man $|\mu| = |c_1|\,\delta_{x_1} + \ldots + |c_n|\,\delta_{x_n}$ bzw $|\mu|(A) = \int_A |f(x)|dx$.

Wir definieren nun die *Variationsnorm*

$$\|\mu\| := |\mu|(\Omega).$$

Satz 4.39. *Die Variationsnorm ist eine Norm, und $M(\Omega, \Sigma)$ mit der Variationsnorm ist ein Banachraum.*

Für einen Beweis siehe zB Rudin: Real and Complex Analysis, McGraw-Hill, 1986.

Für die Verbindung zu Räumen von stetigen Funktionen betrachten wir einen metrischen Raum Ω, die Borel-σ-Algebra Σ_B auf Ω, und setzen $M(\Omega) := M(\Omega, \Sigma_B)$. Real- und Imaginärteil eines komplexes Maßes $\mu \in M(\Omega)$ sind signierte Maße, und ein signiertes Maß ν kann auf eindeutige Art und Weise als Differenz $\nu = \nu_+ - \nu_-$ mit zwei positiven Maßen geschrieben werden. Es gilt dann $|\nu| = \nu_+ + \nu_-$, und eine messbare Funktion $f : \Omega \to \mathbb{R}$ heißt ν-integrierbar, wenn sie $|\nu|$-integrierbar ist. Man setzt dann $\int f d\nu := \int f d\nu_+ - \int f d\nu_-$ und entsprechend für ein komplexes Maß.

Satz 4.40 (Rieszscher Darstellungssatz). *Sei Ω ein kompakter metrischer Raum. Dann ist*

$$T : M(\Omega) \to C(\Omega)', \qquad (T\mu)(f) := \int_\Omega f\, d\mu$$

ein isometrischer Isomorphismus von Banachräumen. Ist $\varphi \in C(\Omega)'$ ein positives *Funktional in dem Sinne, dass $\varphi(f) \geq 0$ für alle $f \geq 0$, so gibt es ein eindeutiges endliches positives Borelmaß μ auf Ω, so dass $\varphi(f) = \int_\Omega f\, d\mu$ gilt.*

Dieser Satz existiert in vielen Varianten, die sich in den betrachteten Funktionenräumen ($C(\Omega)$, $C_0(\Omega)$, $C_c(\Omega)$, $C_b(\Omega)$, ..) und Anforderungen an Ω (kompakt, lokal kompakt, topologischer Hausdorff-Raum, ...) unterscheiden. Oft werden diese Sätze auch als Satz von Riesz-Markov oder Riesz-Markov-Kakutani bezeichnet.

Auch diesen Satz werden wir nicht beweisen (der Beweis ist lang und schwierig, siehe wieder Rudin), aber zumindest etwas erläutern. Dazu bemerken wir, dass für positives $\mu \in M(\Omega)$ (also ein endliches Maß im Sinne der Analysis 3) und $f \in C(\Omega)$

$$\left| \int_\Omega f\, d\mu \right| \leq \|f\|_\infty \mu(\Omega) = \|f\|_\infty \|\mu\|$$

gilt. In diesem Fall sehen wir also $T\mu \in C(\Omega)'$ und $\|T\mu\| \leq \|\mu\|$. Diese Abschätzungen übertragen sich auf allgemeine Elemente $\mu \in M(\Omega)$; dies zeigt dann, dass T stetig ist mit $\|T\| \leq 1$. Für den Beweis der Bijektivität und Isometrie verweisen wir auf das Buch von Rudin.

Als letztes Element aus der Maßtheorie benötigen wir eine Charakterisierung eines Raumes von beschränkten messbaren Funktionen. Für eine kompakte Menge $\Omega \subset \mathbb{C}$ (wir denken an $\Omega = \sigma(A)$ das Spektrum eines beschränkten Operators) betrachten wir die Menge $\mathscr{B}(\Omega)$ aller beschränkten messbaren (bzgl der Borel-σ-Algebra) Funktionen $f : \Omega \to \mathbb{C}$. Es ist klar, dass $\mathscr{B}(\Omega)$ ein Banachraum mit Supremumsnorm $\|\cdot\|_\infty$ ist. Weiterhin ist $\mathscr{B}(\Omega)$ der kleinste Funktionenraum $\Omega \to \mathbb{C}$ ist, der die stetigen Funktionen enthält und unter punktweisen Limiten von gleichmäßig beschränkten Folgen abgeschlossen ist:

Lemma 4.41. *Sei $\Omega \subset \mathbb{C}$ kompakt und $\mathcal{U} \subset \mathscr{B}(\Omega)$ eine Teilmenge mit den zwei Eigenschaften:*

a) $C(\Omega) \subset \mathcal{U}$

b) Falls $f_n \in \mathcal{U}$, $\sup_n \|f_n\|_\infty < \infty$ und $\lim_{n\to\infty} f_n(\lambda) =: f(\lambda)$ für alle $\lambda \in \Omega$ existiert, dann $f \in \mathcal{U}$.

Dann $\mathcal{U} = \mathscr{B}(\Omega)$.

Ein Beweis findet sich z.B. in [Wer95, Lemma VII.1.5].

Nach den maßtheoretischen Wiederholungen können wir den messbaren Kalkül für selbstadjungierte Operatoren konstruieren.

Theorem 4.42. *Sei $A = A^* \in \mathcal{B}(\mathcal{H})$. Dann existiert ein eindeutiger stetiger Funktionalkalkül*

$$\phi_A^{mb.} : \mathscr{B}(\sigma(A)) \to \mathcal{B}(\mathcal{H}),$$

so dass gilt:

(∗) Falls eine Folge $f_n \in \mathscr{B}(\sigma(A))$, $\sup_n \|f_n\|_\infty < \infty$, für alle $\lambda \in \sigma(A)$ punktweise konvergiert, d.h. $f_n(\lambda) \to f(\lambda)$, dann gilt für alle $v \in \mathcal{H}$

$$\lim_{n\to\infty} \|\phi_A^{mb.}(f_n)v - \phi_A^{mb.}(f)v\| = 0.$$

Wie bei den bisher besprochenen Funktionalkalkülen gilt auch hier, dass $f(A) := \phi_A^{mb.}(f)$ für alle $f \in \mathscr{B}(\sigma(A))$ normal ist, dass für reellwertige f der Operator $f(A)$ selbstadjungiert ist, und dass $f(A)$ und $g(A)$ für alle $f, g \in \mathscr{B}(\sigma(A))$ kommutieren – dies folgt ja einfach aus den definierenden Eigenschaften eines Funktionalkalküls.

Weiterhin ist klar, dass ϕ_A^{mb} eine Fortsetzung von ϕ_A^{stetig} ist, da $\phi_A^{mb}|_{C(\sigma(A))}$ als stetiger Funktionalkalkül eindeutig ist, also mit ϕ_A^{stetig} übereinstimmt (Satz 4.22).

Beweis. **Existenz:** Die Idee ist, den Rieszschen Darstellungssatz zu verwenden. Dazu betrachten wir für $f \in C(\sigma(A))$ und $v \in \mathcal{H}$ die Abbildung

$$C(\sigma(A)) \ni f \longmapsto \langle v, f(A)v \rangle \in \mathbb{C}.$$

Dies ist offenbar ein lineares Funktional. Da der stetige Kalkül isometrisch ist, gilt weiterhin

$$|\langle v, f(A)v\rangle| \leq \|v\|^2\|f\|_\infty.$$

Also ist das Funktional stetig, mit Norm höchstens $\|v\|^2$, und nach dem Rieszschen Darstellungssatz existiert ein eindeutiges endliches Maß μ_v^A auf der Borel-σ-Algebra von $\sigma(A)$, so dass für alle $f \in C(\sigma(A))$

$$\langle v, f(A)v\rangle = \int_{\sigma(A)} f(\lambda)\, d\mu_v^A(\lambda). \tag{$\star$}$$

Das Maß μ_v^A ist endlich, mit $\mu_v^A(\sigma(A)) \leq \|v\|^2$. Die wesentliche Beobachtung ist nun, dass die rechte Seite von $(\star)$ auch für $f \in \mathscr{B}(\sigma(A))$ sinnvoll ist. In der Tat gilt

$$\left|\int_{\sigma(A)} f\, d\mu_v^A\right| \leq \|f\|_\infty \mu_v^A(\sigma(A)) \leq \|f\|_\infty\|v\|^2,$$

Wir definieren nun einen Operator $f(A)$, indem wir seine Matrixelemente $\langle v, f(A)w\rangle$ definieren. Wir setzen $\langle v, f(A)v\rangle := \int_{\sigma(A)} f\, d\mu_v^A$, und für $v \neq w$ per Polarisation

$$\begin{aligned}\langle v, f(A)w\rangle := \frac{1}{4}\Big(&\langle v+w, f(A)(v+w)\rangle - \langle v-w, f(A)(v-w)\rangle \\ &+ i\langle v-iw, f(A)(v-iw)\rangle - i\langle v+iw, f(A)(v+iw)\rangle\Big).\end{aligned}$$

Da diese Sesquilinarform durch $|\langle v, f(A)w\rangle| \leq \|f\|_\infty\|v\|\|w\|$ beschränkt ist, definiert dies einen beschränkten Operator mit $\|f(A)\| \leq \|f\|_\infty$ (siehe Lemma 2.19). Dies ist die Definition des messbaren Kalküls $\phi_A^{mb.}(f) = f(A)$. Beachten Sie, dass $f(A)$ offensichtlich linear von f abhängt und $\|f(A)\| \leq \|f\|_\infty$ erfüllt, d.h. $\phi_A^{mb.}$ ist stetig. Die Matrixelemente $\langle v, f(A)w\rangle = \int_{\sigma(A)} f(\lambda) d\mu_{v,w}^A(\lambda)$ können auch direkt durch komplexe Maße $\mu_{v,w}^A$ ausgedrückt werden.

Es bleibt zu zeigen, dass $\phi_A^{mb.}$ ein Funktionalkalkül ist und die Eigenschaft $(*)$ hat. Dazu bemerken wir zuerst, dass $\phi_A^{mb.}(1) = 1$ und $\phi_A^{mb.}(\mathrm{id}) = A$ automatisch folgt, da $\phi_A^{mb.}$ auf $C(\sigma(A))$ ja mit dem stetigen Kalkül übereinstimmt.

Um zu demonstrieren, dass $\phi_A^{mb.}(fg) = \phi_A^{mb.}(f)\phi_A^{mb.}(g)$ gilt, zeigen wir zuerst eine abgeschwächte Form von $(*)$: Wir behaupten, dass für eine Folge $f_n \in \mathscr{B}(\sigma(A))$, $\sup_n \|f_n\|_\infty < \infty$, so dass für alle $\lambda \in \sigma(A)$ punktweise konvergiert, für alle $v, w \in \mathcal{H}$ gilt:

$$\lim_{n\to\infty} \langle v, \phi_A^{mb.}(f_n)w\rangle = \langle v, \phi_A^{mb.} w\rangle.$$

(Konvergenz $\phi_A^{mb.}(f_n) \to \phi_A^{mb.}(f)$ in der schwachen Operatortopologie.) Diese Aussage bedeutet also

$$\lim_{n\to\infty} \int f_n\, d\mu_{v,w}^A = \int f\, d\mu_{v,w}^A$$

und folgt sofort mit Polarisation und majorisierter Konvergenz.

Nun möchten wir zeigen, dass $\phi_A^{mb.}$ ein Algebren-Homomorphismus ist. Um dies zu zeigen, wählen wir zunächst $g \in C(\sigma(A))$ und betrachten die Menge

$$\mathcal{U} := \{f \in \mathscr{B}(\sigma(A)) \,:\, \phi_A^{mb.}(gf) = \phi_A^{mb.}(g)\phi_A^{mb.}(f)\}.$$

Natürlich gilt $C(\sigma(A)) \subset \mathcal{U}$. Für eine gleichmäßig beschränkte punktweise konvergente Folge $\mathcal{U} \ni f_n \to f$ gilt weiterhin ($v, w \in \mathcal{H}$ beliebig):

$$\langle v, \phi_A^{mb.}(gf_n)w\rangle = \langle v, \phi_A^{mb.}(g)\phi_A^{mb.}(f_n)w\rangle \to \langle v, \phi_A^{mb.}(g)\phi_A^{mb.}(f)w\rangle.$$

Andererseits ist aber auch gf_n eine gleichmäßig beschränkte Folge, die punktweise gegen gf konvergiert. Also gilt auch $\langle v, \phi_A^{mb.}(gf_n)w\rangle \to \langle v, \phi_A^{mb.}(gf)w\rangle$. Da v, w beliebig waren, folgt $\phi_A^{mb.}(gf) = \phi_A^{mb.}(g)\phi_A^{mb.}(f)$, d.h. $f \in \mathcal{U}$ und mit Lemma 4.41 $\mathcal{U} = \mathscr{B}(\sigma(A))$.

In einem zweiten Schritt wollen wir nun die anfangs gewählte stetige Referenzfunktion g eliminieren. Sei nun also $f \in \mathscr{B}(\sigma(A))$, und definiere

$$\mathcal{V} := \{g \in \mathscr{B}(\sigma(A)) \,:\, \phi_A^{mb.}(gf) = \phi_A^{mb.}(g)\phi_A^{mb.}(f)\}.$$

Nach dem obigen Beweisschritt gilt $C(\sigma(A)) \subset \mathcal{V}$. Sei wieder g_n eine gleichmäßig beschränkte Folge in $\mathcal{V}$, die punktweise gegen g konvergiert. Dann folgt wie oben $g \in \mathcal{V}$, und mit Lemma 4.41 $\mathcal{V} = \mathscr{B}(\sigma(A))$, was die Homomorphismus-Eigenschaft von $\phi_A^{mb.}$ beweist.

Der Beweis von $\phi_A^{mb.}(f)^* = \phi_A^{mb.}(\bar{f})$ verläuft sehr ähnlich (und einfacher) und wird deshalb zur Übung überlassen.

Als letzten Schritt müssen wir zeigen, dass statt der gezeigten schwachen Konvergenz (das heißt $\lim_{n\to\infty}\langle v, \phi_A^{mb.}(f_n)w\rangle = \langle v, \phi_A^{mb.} w\rangle$.) sogar die starke Konvergenz ($*$) gilt. Dazu betrachten wir eine gleichmäßig beschränkte Folge $f_n \in \mathscr{B}(\sigma(A))$, die punktweise gegen f konvergiert, und berechnen zunächst auf beliebigem $v \in \mathcal{H}$

$$\begin{aligned}\|f_n(A)v\|^2 &= \langle f_n(A)v, f_n(A)v\rangle = \langle v, f_n(A)^* f_n(A)v\rangle \\ &= \langle v, |f_n|^2(A)v\rangle \to \langle v, |f|^2(A)v\rangle = \|f(A)v\|^2.\end{aligned}$$

Wir haben also gezeigt, dass die Folge $(f_n(A)v)_n$ schwach gegen $f(A)v$ konvergiert, und die Normen von $\|f_n(A)v\|$ gegen die Norm $\|f(A)v\|$ konvergiert. Damit folgt die Behauptung (gezeigt in den Übungen).

Eindeutigkeit: Zuerst bemerken wir, dass jeder stetige Funktionalkalkül $\phi : \mathscr{B}(\sigma(A)) \to \mathcal{B}(\mathcal{H})$ auf der Unteralgebra $C(\sigma(A))$ eindeutig festgelegt ist: Auf Polynomen ist ϕ durch $\phi(1) = 1$ und $\phi(\mathrm{id}) = A$ und die Homorphismus-Eigenschaft fixiert, und wegen dem Satz von Stone-Weierstraß ist jeder *stetige* Kalkül auf $C(\sigma(A))$ notwendigerweise gleich dem stetigen Kalkül ϕ_A^{stetig}.

Seien nun ϕ, ϕ' zwei stetige Funktionalkalküle mit der angegebenen Zusatzeigenschaft ($*$). Wir müssen $\phi = \phi'$ auf ganz $\mathscr{B}(\sigma(A))$ zeigen. Dazu definieren wir

$$\mathcal{U} := \{f \in \mathscr{B}(\sigma(A)) \,:\, \phi(f) = \phi'(f)\};$$

nach den obigen Bemerkungen gilt also $C(\sigma(A)) \subset \mathcal{U}$. Sei nun f_n eine beschränkte Folge in $\mathcal{U}$, so dass $f_n \to f$ punktweise. Dann gilt für alle $v \in \mathcal{H}$

$$\phi(f_n)v \to \phi(f)v, \qquad \phi'(f_n)v \to \phi'(f)v$$

in der Normtopologie von $\mathcal{H}$ nach ($*$). Da aber $f_n \in \mathcal{U}$, gilt auch $\phi(f_n) = \phi'(f_n)$. Es folgt $\phi(f) = \phi'(f)$, d.h. $f \in \mathcal{U}$. Nach Lemma 4.41 folgt $\mathcal{U} = \mathscr{B}(\sigma(A))$, also $\phi = \phi'$. $\square$

Bemerkungen:

a) In diesem Satz wird kein spektrales Abbildungsprinzipg für den messbaren Kalkül behauptet, und tatsächlich gilt im Allgemeinen $\sigma(f(A)) \neq f(\sigma(A))$. In Hausaufgabe H12.2 sollen Sie dies per Beispiel zeigen und außerdem herausfinden, welche Variante des spektralen Abbildungsprinzips gilt.

   Unabhängig davon betrachten wir hier das Punktspektrum: Für $\lambda \in \sigma_p(A)$ mit zugehörigem Eigenvektor $v \in \mathcal{H}$ gilt für alle stetigen Funktionen $f \in C(\sigma(A))$ die Gleichung $f(A)v = f(\lambda) \cdot v$ (offensichtlich für Polynome, und für allgemeines f durch Approximation). In diesem Fall ist das durch den Satz von Riesz-Markov gegebene Maß μ_v also ein Dirac-Maß

$$f(\lambda)\|v\|^2 = \langle v, f(A)v\rangle = \int_{\sigma(A)} f(x)d\mu_v(x) \;\Rightarrow\; \mu_v = \|v\|^2\delta_\lambda,$$

   und wir sehen $f(A)v = f(\lambda)v$ für alle $f \in \mathscr{B}(\sigma(A))$.

b) Der messbare Kalkül erfüllt $\|\phi_A^{mb}(f)\| \leq \|f\|_\infty$, ist aber im Allgemeinen nicht isometrisch. Sei als Beispiel $\mathcal{H} = L^2([0,1])$ und $(Av)(x) = x{\cdot}v(x)$. Dann gilt $(f(A)v)(x) = f(x)v(x)$ (siehe Aufgabe H12.3). Für die Funktion $f = \chi_{\{0\}}$ gilt dann $\|f\|_\infty = 1$ aber $\|f(A)\| = 0$.

c) Für $A, B \in \mathcal{B}(\mathcal{H})$ mit $A = A^*$ und $AB = BA$ gilt auch $f(A)B = Bf(A)$ für alle $f \in \mathscr{B}(\sigma(A))$. Da wir dies für stetiges f bereits wissen, gilt für $v, w \in \mathcal{H}$

$$\langle B^*v, f(A)w\rangle = \int_{\sigma(A)} f(\lambda)d\mu_{B^*v,w}(\lambda) = \langle v, f(A)Bw\rangle = \int_{\sigma(A)} f(\lambda)d\mu_{v,Bw}(\lambda)$$

   und damit $\mu_{B^*v,w} = \mu_{v,Bw}$, was die Behauptung impliziert.

Die wichtigsten Funktionen f, die wir durch die Ausdehnung des stetigen auf den messbaren Kalkül hinzugewonnen haben, sind charakteristische Funktionen χ_S von Borelmengen $S \subset \sigma(A)$. Wir bemerken, dass $\chi_S(A)$ orthogonale Projektionen sind (da $\chi_S^2 = \chi_S = \overline{\chi_S}$). Da wir im Folgenden viel über (stets orthogonale) Projektionen sprechen werden, führen wir die Bezeichnung $\mathcal{P}(\mathcal{H})$ für die Menge aller Projektionen in $\mathcal{B}(\mathcal{H})$ ein.

Motiviert vom Spektralsatz in endlicher Dimension ($A = \sum_{\lambda\in\sigma(A)} \lambda P_\lambda^A$), wobei P_λ^A die Projektion auf den Eigenraum $\ker(A - \lambda)$ ist, möchten wir jetzt die Summe $\sum_{\lambda\in\sigma(A)}$ zu einem geeigneten Integral erweitern, dass auch für kontinuierliches Spektrum sinnvoll ist. Dazu führen wir zuerst projektorwertige Maße ein.

Definition 4.43. Ein *projektorwertiges Maß* E auf der Borel-σ-Algebra Σ von $\mathbb{R}$ ist eine Abbildung

$$E : \Sigma \to \mathcal{P}(\mathcal{H}),$$

so dass gilt:

a) $E_\emptyset = 0$, $E_\mathbb{R} = 1$,

b) Sind $S_n \in \Sigma$ abzählbar viele paarweise disjunkte Borelmengen, so gilt für alle

$v \in \mathcal{H}$

$$\sum_n E_{S_n} v = E_{\bigcup_n S_n} v.$$

Im Falle von unendlich vielen S_n ist die linke Seite als Konvergenz in der Normtopologie von $\mathcal{H}$ zu verstehen (d.h. starke Konvergenz von $\sum_{n=1}^N E_{S_n}$ für $N \to \infty$).

Der *Träger von* E, geschrieben $\operatorname{supp} E$, ist die kleinste abgeschlossene Menge $K \subset \mathbb{R}$ mit $E_K = 1$.

Ein Beispiel eines projektorwertigen Maßes wird in Aufgabe P12.2 besprochen. Hier noch ein weiteres Beispiel:

Beispiel 4.44. Sei $\mathcal{H}$ ein separabler Hilbertraum, $(v_n)_n$ eine Orthonormalbasis, und $(\lambda_n)_{n\in\mathbb{N}}$ eine Folge von paarweise verschiedenen reellen Zahlen. Sei P_n die orthogonale Projektion auf den eindimensionalen Unterraum $\mathbb{C}v_n$. Dann ist

$$E_S \xi := \sum_{\substack{n\in\mathbb{N}\\ \lambda_n \in S}} P_n \xi = \sum_{\substack{n\in\mathbb{N}\\ \lambda_n \in S}} \langle v_n, \xi \rangle v_n$$

ein projektorwertiges Maß.

Diese Reihe konvergiert (siehe Abschnitt 2.3). Da die P_n paarweise orthogonal sind, sind die E_S orthogonale Projektionen. Offenbar gilt $E_\emptyset = 0$ und $E_\mathbb{R} = 1$ wegen der Basisentwicklungsformel $\xi = \sum_n \langle v_n, \xi \rangle v_n$, $\xi \in \mathcal{H}$. Die σ-Additivität b) folgt auch einfach: Sind S_m, $m \in \mathbb{N}$, messbare paarweise disunkte Teilmengen von $\mathbb{R}$, und $S := \bigsqcup_m S_m$, so gilt für $\xi \in \mathcal{H}$

$$E_S \xi = \sum_{\substack{n\in\mathbb{N}\\ \lambda_n \in S}} \langle v_n, \xi \rangle v_n = \sum_m \sum_{\substack{n\in\mathbb{N}\\ \lambda_n \in S_m}} \langle v_n, \xi \rangle v_n = \sum_m E_{S_m} \xi.$$

Dieses Maß hat kompakten Träger genau dann, wenn die Folge $(\lambda_n)_{n\in\mathbb{N}}$ beschränkt ist.

Das folgende Lemma zeigt, dass sich die Projektionen analog zu charakeristischen Funktionen verhalten.

Lemma 4.45. *Sei E ein projektorwertiges Maß und $S, T \in \Sigma$ Borelmengen.*

a) Falls $S \subset T$, so gilt $E_S \leq E_T$ (siehe auch Aufgabe P12.1)

b) $E_S E_T = E_{S\cap T} = E_T E_S$.

c) Für $v \in \mathcal{H}$ ist $\mu_v(S) := \|E_S v\|^2$ ein endliches Borelmaß mit $\mu_v(\mathbb{R}) = \|v\|^2$.

d) Für $S \cap \operatorname{supp} E = \emptyset$ gilt $E_S = 0$.

Beweis. a) Es gilt $E_T = E_{S\sqcup(T\setminus S)} = E_S + E_{T\setminus S} \geq E_S$, da $E_{T\setminus S} \geq 0$ (Projektionen sind positiv).

b) Wie in Aufgabe P12.1 gezeigt wird, gilt für Projektionen P, Q die Äquivalenz $P \leq Q \Leftrightarrow QP = P = PQ$. Damit erhalten wir für disjunkte S, T aus $E_{S\sqcup T} = E_S + E_T$ sofort

$E_S = E_S + E_S E_T$, also $E_S E_T = 0$. Für allgemeine Mengen S, T haben wir

$$\begin{aligned} E_{S\cup T} &= E_{S\setminus T} + E_{S\cap T} + E_{T\setminus S} \\ \Rightarrow E_S &= E_{S\setminus T} + E_{S\cap T} \quad (\text{Mult. mit } E_S) \\ \Rightarrow E_S E_T &= E_{S\cap T} \quad (\text{Mult. mit } E_T) \end{aligned}$$

c) μ_v ist eine Abbildung $\Sigma \to [0,\infty)$, die für paarweise disjunkte Borelmengen S_n

$$\|\mu_v(\bigsqcup_n S_n)\|^2 = \sum_{n,m} \langle E_{S_n} v, E_{S_m} v\rangle = \sum_n \langle E_{S_n} v, v\rangle = \sum_{n,m} \mu_v(S_n)$$

erfüllt, also σ-additiv ist. Offenbar gilt $\mu_v(\emptyset) = 0$ und $\mu_v(\mathbb{R}) = \|v\|^2$.

d) Wir haben $E_{\text{supp}\, E \sqcup S} = E_{\text{supp}\, E} + E_S = 1 + E_S$. Da dies eine orthogonale Projektion ist, folgt $E_S = 0$. □

Projektorwertige Maße lassen sich unabhängig von Spektraltheorie definieren, sind aber eng mit dem messbaren Kalkül verbunden. Wir zeigen nun, dass jeder selbstadjungierte Operator A ein projektorwertiges Maß definiert.

Lemma 4.46. *Sei $A = A^* \in \mathcal{B}(\mathcal{H})$. Dann ist das* **Spektralmaß von** A

$$E^A : \Sigma \to \mathcal{P}(\mathcal{H}), \qquad S \mapsto \chi_{S\cap\sigma(A)}(A)$$

ein projektorwertiges Maß mit kompaktem Träger supp $E^A \subset \sigma(A)$.

Beweis. Für $S \in \Sigma$ mit $S \cap \sigma(A) = \emptyset$ gilt $\chi_{S\cap\sigma(A)} = 0$, also $E^A_S = 0$ und insbesondere $E^A_\emptyset = 0$. Für $S \in \Sigma$ mit $S \supset \sigma(A)$ gilt $\chi_{S\cap\sigma(A)} = \chi_{\sigma(A)}$, also $E^A_S = 1$ und insbesondere $E^A_\mathbb{R} = 1$.

Für den Beweis der σ-Additivität betrachten wir abzählbar viele paarweise disjunkte $S_n \in \Sigma$, und $v \in \mathcal{H}$. Die Partialsummen $\varphi_N := \sum_{n=1}^N \chi_{S_n}$ konvergieren punktweise gegen $\varphi := \sum_{n=1}^\infty \chi_{S_n} = \chi_{\bigcup_{n=1}^\infty S_n}$ für $N \to \infty$, und erfüllen $\sup_N \|\varphi_N\|_\infty \leq 1 < \infty$. Nach der starken Stetigkeit des messbaren Kalküls gilt also ($v \in \mathcal{H}$ beliebig)

$$E^A_{\bigcup_{n=1}^\infty S_n} v = \lim_{N\to\infty} \sum_{n=1}^N \chi_{S_n\cap\sigma(A)}(A)v = \sum_{n=1}^\infty \chi_{S_n\cap\sigma(A)}(A)v = \sum_{n=1}^\infty E^A_{S_n} v.$$

E^A hat kompakten Träger, da $\sigma(A) \subset \mathbb{R}$ kompakt ist und $E^A_{\sigma(A)} = 1$. □

Wir möchten nun ein Integral der Form $\int f\, dE \in \mathcal{B}(\mathcal{H})$ definieren, wobei $f \in \mathscr{B}(\mathbb{R})$ eine beschränkte messbare zahlenwertige Funktion und E ein projektorwertiges Maß sein soll. Wir gehen ähnlich wie bei der Konstruktion des Lebesgue-Integrals vor und definieren $\int f\, dE$ zunächst für einfache Funktionen.

Falls $f = \chi_S$ mit $S \in \Sigma$, setzen wir

$$\int \chi_S\, dE := E_S.$$

Ist f etwas allgemeiner eine Treppenfunktion, $f = \sum_{k=1}^{n} \alpha_k \, \chi_{S_k}$, so setzen wir

$$\int \sum_{k=1}^{n} \alpha_k \, \chi_{S_k} \, dE := \sum_{k=1}^{n} \alpha_k \, E_{S_k}.$$

Um einzusehen, dass das wohldefiniert ist, muss man zeigen, dass diese Definition nicht von der Wahl der Zerlegung einer Treppenfunktion in charakteristische Funktionen abhängt (siehe Analysis 3).

Für allgemeines $f \in \mathscr{B}(\mathbb{R})$ approximieren wir durch Treppenfunktionen. Dazu erinnern wir uns (Analysis 3), dass es zu $f \in \mathscr{B}(\mathbb{R})$ stets eine Folge φ_n von Treppenfunktionen gibt, die gleichmäßig gegen f konvergiert.

Lemma 4.47. *Sei E ein projektorwertiges Maß.*

a) Für jede Treppenfunktion φ gilt

$$\left\| \int \varphi \, dE \right\| \leq \|\varphi\|_\infty.$$

b) Sei $f \in \mathscr{B}(\mathbb{R})$ und φ_n eine Folge von Treppenfunktionen mit $\|f - \varphi_n\|_\infty \to 0$. Dann existiert der Limes

$$\int f \, dE := \lim_{n\to\infty} \int \varphi_n \, dE$$

und hängt nicht von der Wahl der approximierenden Folge ab.

Beweis. a) Wir können annehmen, dass $\varphi = \sum_{k=1}^{n} \alpha_k \chi_{S_k}$ mit paarweise disjunkten S_k gilt (falls nicht, zerlege die S_k feiner). In der folgenden Rechnung benutzen wir $E_S E_T = E_{S\cap T}$ und $E_\emptyset = 0$ zusammen mit dem Satz von Pythagoras. Für beliebiges $v \in \mathcal{H}$ erhalten wir so

$$\begin{aligned}\left\| \left(\int \varphi \, dE \right) v \right\|^2 &= \left\| \sum_{k=1}^{n} \alpha_k E_{S_k} v \right\|^2 = \sum_{k,l=1}^{n} \overline{\alpha_l} \alpha_k \langle E_{S_l} v, E_{S_k} v \rangle = \sum_{k=1}^{n} |\alpha_k|^2 \|E_{S_k} v\|^2 \\ &\leq \sup_k |\alpha_k|^2 \sum_l \|E_{S_l} v\|^2 = \|\varphi\|_\infty^2 \|E_{\bigcup_l S_l} v\|^2 \leq \|\varphi\|_\infty^2 \|v\|^2.\end{aligned}$$

Im letzten Schritt haben wir verwendet, dass eine Projektion Norm höchstens 1 hat. Die Behauptung folgt nun.

b) Da die Folge (φ_n) konvergiert, ist sie Cauchy bzgl Supremumsnorm. Also zeigt a) via $\| \int \varphi_n dE - \int \varphi_m dE \| \leq \|\varphi_n - \varphi_m\|_\infty \to 0$, dass $(\int \varphi_n dE)_n$ eine Cauchyfolge in $\mathcal{B}(\mathcal{H})$ ist, die wegen der Vollständigkeit von $\mathcal{B}(\mathcal{H})$ einen Grenzwert hat.

Ist $(\tilde{\varphi}_n)$ eine weitere Folge von Treppenfunktionen, die gleichmäßig gegen f konvergiert, so haben wir $\| \int \varphi_n dE - \int \tilde{\varphi}_n dE \| \leq \|\varphi_n - \tilde{\varphi}_n\| \to 0$, dass der Limes nicht von der Wahl der approximerenden Folge abhängt. $\square$

Mit Hilfe dieses Lemmas können wir nun also für $f \in \mathscr{B}(\mathbb{R})$ das Integral $\int f\,dE$ definieren als

$$\int f\,dE := \lim_{n\to\infty} \int \varphi_n\,dE,$$

wobei φ_n Treppenfunktionen sind, die gleichmäßig gegen f konvergieren, und der obige Limes in der Normtopologie von $\mathcal{B}(\mathcal{H})$ gilt. Diese Konstruktion verlief genau analog zur Konstruktion des Integrals bzgl eines Maßes. Betrachten wir das zahlenwertige Maß $\mu_v(S) := \|E_S v\|^2 = \langle v, E_S v\rangle$ (mit $v \in \mathcal{H}$), so haben wir für Treppenfunktionen $\varphi = \sum_k \alpha_k \chi_{S_k}$

$$\langle v, \int \varphi\,dE\,v\rangle = \sum_k \alpha_k \langle v, E_{S_k} v\rangle = \sum_k \alpha_k \mu_v(S_k) = \int \varphi(\lambda) d\mu_v(\lambda).$$

Durch gleichmäßige Approximation einer allgemeinen Funktion $f \in \mathscr{B}(\mathbb{R})$ durch Treppenfunktionen erhalten wir deshalb per majorisierter Konvergenz

$$\langle v, \int f\,dE\,v\rangle = \int f(\lambda) d\mu_v(\lambda).$$

Weitere gebräuchliche Notationen sind $\int f(\lambda)\,dE(\lambda) = \int f\,dE$ sowie $\langle v, \int f\,dE\,v\rangle = \int f\,d\langle v, Ev\rangle$ und $\int_S f\,dE := \int f\chi_S\,dE$ für messbares S.

Da $E_S = 0$ für $S \cap \operatorname{supp} E = \emptyset$, gilt weiterhin

$$\int f\,dE = \int f|_{\operatorname{supp} E}\,dE.$$

Für reelle Treppenfunktion φ_n ist $\int \varphi_n dE$ per Definition selbstadjungiert. Eine reelle Funktion $f \in \mathscr{B}(\mathbb{R})$ können wir durch reelle Treppenfunktionen approximieren. Also ist für reelles f der Operator $\int f\,dE$ selbstadjungiert. Insbesondere erhalten wir zu jedem projektorwertigen Maß mit kompaktem Träger einen selbstadjungierten Operator

$$B := \int \lambda\,dE(\lambda),$$

da die Identitätsfunktion $\lambda \mapsto \lambda$ auf kompaktem $\operatorname{supp} E$ beschränkt ist.

Theorem 4.48. **(Spektralsatz)** *Sei $A = A^* \in \mathcal{B}(\mathcal{H})$. Dann existiert ein eindeutiges projektorwertiges Maß $E = E^A$ mit kompaktem Träger, das Spektralmaß von A, so dass*

$$A = \int \lambda\,dE^A(\lambda)$$

gilt. Der messbare Funktionalkalkül von A ist durch

$$\phi_A^{mb}(f) = \int f\,dE^A$$

gegeben.

Beweis. Wir wissen, dass E^A ein projektorwertiges Maß mit kompaktem Träger ist. Also ist $B := \int \lambda dE^A(\lambda)$ ein wohldefinierter selbstadjungierter Operator. Allgemeiner ist klar, dass $\psi : \mathscr{B}(\sigma(B)) \to \mathcal{B}(\mathcal{H})$, $\psi(f) := \int f\, dE^A$, eine wohldefinierte lineare stetige Abbildung mit $\|\psi\| \leq 1$ ist. Weiterhin gilt $\psi(fg) = \psi(f)\psi(g)$ und $\psi(f)^* = \psi(\overline{f})$: Die erste Gleichung folgt für Treppenfunktionen $f = \sum_k a_k \chi_{S_k}$, $g = \sum_l b_l \chi_{T_l}$ aus $fg = \sum_{k,l} a_k b_l \chi_{S_k \cap T_l}$ und $E^A_{S_k \cap T_l} = E^A_{S_k} E^A_{T_l}$ und für allgemeines f, g per Approximation, und die zweite Gleichung folgt für eine Treppenfunktion direkt und für allgemeines f wieder per Approximation.

Es genügt, die schwache Version der Stetigkeitseigenschaft $(*)$ aus Theorem 4.42 zu zeigen, also $\langle v, \psi(f_n)w\rangle \to \langle v, \psi(f)w\rangle$ für beschränkte punktweise konvergente Folgen $\mathscr{B} \ni f_n \to f$. Für $w = v$ ist dies eine Konsequenz des Satzes über majorisierte Konvergenz:

$$\langle v, \psi(f_n)v\rangle = \int f_n\, d\langle v, Ev\rangle \to \int f\, d\langle v, Ev\rangle = \langle v, \psi(f)v\rangle.$$

Für $v \neq w$ haben wir ein komplexes Maß $\langle v, Ew\rangle$. Da sich dies in eine Linearkombination von positiven Maßen zerlegen lässt, folgt die Behauptung auch in diesem Fall.

Wir behaupten außerdem $\psi(1) = 1$. Dies folgt aus $E^A_{\sigma(B)} = 1$ und lässt sich durch ein etwas technisches Argument zeigen (keine Details hier, siehe Werner S. 357/358). Dann folgt $\psi(\mathrm{id}) = \int \lambda\, dE^A(\lambda) = B$. Aufgrund der Eindeutigkeit des messbaren Kalküls zeigt dies $\psi = \phi_B^{mb}$.

Damit folgt für messbares $S \subset \mathbb{R}$

$$E^B(S) = \chi_S(B) = \phi_B^{mb}(\chi_S) = \int \chi_S\, dE^A = E^A(S),$$

also $E^B = E^A$. Dies zeigt auch die Eindeutigkeitsaussage.

Wir wollen jetzt noch $A = B$ zeigen. Sei dazu $\varepsilon > 0$ und $\varphi : \sigma(A) \to \mathbb{R}$ eine Treppenfunktion mit $\|\varphi - \mathrm{id}\|_\infty < \varepsilon$. Wegen $E^A(S) = E^B(S)$ gilt $\varphi(A) = \int \varphi(\lambda) dE^A(\lambda)$ und

$$\|A - B\| \leq \|A - \varphi(A)\| + \left\| \int \varphi(\lambda) dE^A(\lambda) - \int \lambda\, dE^A(\lambda) \right\|.$$

Der erste Term ist $\leq \varepsilon$ nach Satz 4.42, und der zweite Term ist gleich $\|\phi_B^{mb}(\varphi) - \phi_B^{mb}(\mathrm{id})\| \leq \|\varphi - \mathrm{id}\|_\infty = \varepsilon$. Damit folgt $A = B$. □

Die Formel

$$A = \int \lambda\, dE(\lambda)$$

wird *Spektraldarstellung von A* genannt. Dies ist die gesuchte Verallgemeinerung von der endlichdimensionalen Diagonalisierungsformel $A = \sum_{\lambda \in \sigma(A)} \lambda\, P_\lambda$.

Beispiel 4.49. Sei $\dim \mathcal{H} < \infty$ und $A = A^*$. Dann ist das Spektralmaß von A

$$E^A_S = \sum_{\lambda \in S} P^A_\lambda,$$

wobei P^A_λ die orthogonale Projektion auf den Eigenraum $\ker(A - \lambda)$ ist.
Beweis: Wir sehen, dass $\tilde{E}_S := \sum_{\lambda \in S} P^A_\lambda$ ein projektorwertiges Maß ist: Die Werte sind

orthogonale Projektionen, weil die P_λ^A paarweise orthogonal sind, $\tilde{E}_\emptyset^A = 0$ und $\tilde{E}_\mathbb{R}^A = 1$ sind klar, und σ-Additivität über der endlichen Menge $\sigma(A)$ ist trivial.

Auf der endlichen Menge $\sigma(A)$ ist jede Funktion $f = \sum_{\lambda\in\sigma(A)} f(\lambda)\chi_{\{\lambda\}}$ eine Treppenfunktion. Das durch $\tilde{E}$ definierte Integral ist also

$$\int f(\lambda)\, d\tilde{E}(\lambda) = \sum_{\lambda\in\sigma(A)} f(\lambda)\tilde{E}(\{\lambda\}) = \sum_{\lambda\in\sigma(A)} f(\lambda)\tilde{P}_\lambda^A.$$

Wegen $\sum_{\lambda\in\sigma(A)} \lambda P_\lambda^A = A$ (Satz 4.36) folgt nun $\tilde{E} = E^A$ aus Satz 4.48. □

Beispiel 4.50. Sei $\mathcal{H} = L^2([0,1])$ und A der selbstadjungierte Multiplikationsoperator $(Af)(x) = x \cdot f(x)$. Dann ist das Spektralmaß von A

$$E^A(S) = M_{\chi_S}, \qquad M_{\chi_S} f := \chi_S \cdot f.$$

Dies wird in Hausaufgabe H12.3 bewiesen.

Diese beiden grundlegenden Beispiele zeigen, wie Sie über die Spektralmaße E^A nachdenken sollten. Im ersten, "diskreten", Beispiel projiziert E_S^A auf die direkte Summe von allen Eigenräumen von A zugehörig zu Eigenwerten $\lambda \in S$. Im zweiten, "kontinuierlichen", Beispiel sollten Sie sich $\operatorname{Ran} E_S^A = \{f \in L^2([0,1]) : \ \operatorname{supp} f \subset S\}$ als einen zu Spektralwerten $\lambda \in S \subset \sigma(A)$ gehörigen Raum von verallgemeinerten Eigenvektoren.

Als eine erste Illustration zeigen wir das folgende Lemma.

Lemma 4.51. *Sei $A = A^* \in \mathcal{B}(\mathcal{H})$ und $\lambda \in \mathbb{R}$.*

a) $\lambda \in \rho(A) \Leftrightarrow$ *es existiert eine offene Umgebung* $\lambda \in \mathcal{O} \subset \mathbb{R}$ *mit* $E_\mathcal{O}^A = 0$.

b) $\lambda \in \sigma_p(A) \Leftrightarrow E_{\{\lambda\}}^A \neq 0$.

c) $\operatorname{supp} E^A = \sigma(A)$.

Beweis. a) $\Rightarrow$ Falls $\lambda \in \rho(A)$, so können wir die offene Umgebung $\mathcal{O} = \rho(A)$ wählen, die $E_{\rho(A)}^A = \chi_{\rho(A)\cap\sigma(A)}(A) = 0$ erfüllt.

$\Leftarrow$ Sei $\lambda \in \mathcal{O} \subset \mathbb{R}$ offen mit $E_\mathcal{O}^A = 0$. Betrachte die messbare beschränkte Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) := \chi_{\mathbb{R}\setminus\mathcal{O}}(x)/(\lambda - x)$. Es gilt per Funktionalkalkül mit der Funktion $g(x) := (x - \lambda)$

$$f(A) \cdot (A - \lambda) = (fg)(A) = \chi_{\mathbb{R}\setminus\mathcal{O}}(A) = E^A(\mathbb{R}\setminus\mathcal{O}) = 1,$$

und analog $(A - \lambda)f(A) = 1$. Also ist $(A - \lambda)$ invertierbar, dh $\lambda \in \rho(A)$.

b) Zunächst sehen wir für $\lambda \in \sigma(A)$

$$(A - \lambda)E_{\{\lambda\}}^A = \int (x - \lambda)\chi_{\{\lambda\}}(x)\, dE^A(\lambda) = 0.$$

Falls $v \in \operatorname{Ran} E_{\{\lambda\}}^A$, so gilt also $Av = AE_{\{\lambda\}}^A v = \lambda E_{\{\lambda\}}^A v = \lambda v$, dh $v \in \ker(A - \lambda)$. Dies zeigt die Implikation $E_{\{\lambda\}}^A \neq 0 \Rightarrow \lambda \in \sigma_p(A)$.

Ist umgekehrt $\lambda \in \sigma_p(A)$ und $Av = \lambda v$, $v \neq 0$, so wissen wir $f(A)v = f(\lambda)v$ für alle $f \in \mathscr{B}(\sigma(A))$. Also gilt für $f = \chi_{\{\lambda\}}$

$$v = \chi_{\{\lambda\}}(\lambda)v = \chi_{\{\lambda\}}(A)v = E^A_{\{\lambda\}}v,$$

d.h. $v \in \operatorname{Ran} E^A_{\{\lambda\}}$ und damit $E^A_{\{\lambda\}} \neq 0$.

c) Sei $K \subset \mathbb{R}$ kompakt und $E^A_K = 1$. Dann gilt $E^A_{\mathbb{R}\setminus K} = 0$. Also gilt nach Teil a) $\lambda \in \rho(A)$ für alle $\lambda \notin K$, dh $\sigma(A) \subset K$. Zusammen mit $\operatorname{supp} E^A \subset \sigma(A)$ impliziert dies die Behauptung. □

Als weitere Anwendung betrachten wir Zerlegungen des Hilbertraums $\mathcal{H}$ mit Hilfe von Spektralprojektionen in spektrale Unterräume. Dazu betrachten wir Borelmengen $S \subset \mathbb{R}$, wobei wir vor allem an Teilmengen des Spektrums, dh $S \subset \sigma(A)$, denken, und setzen $\mathcal{H}_S := \operatorname{Ran} E^A_S$. Im endlichdimensionalen Fall ist z.B. $\sigma(A) = \{\lambda_1, \dots, \lambda_n\}$ und $S = \{\lambda_1, \dots, \lambda_k\}$ mit $k \leq n$. Dann ist $\mathcal{H}_S$ die direkte Summe der Eigenräume mit Eigenwerten $\lambda_1, \dots, \lambda_k$. Im Allgemeinen erhält man:

> Zeigen Sie, dass die spektralen Unterräume $\mathcal{H}_S$ eines selbstadjungierten Operators $A \in \mathcal{B}(\mathcal{H})$ die folgenden Eigenschaften haben:
>
> a) $\mathcal{H}_S$ ist invariant unter A, dh $A\mathcal{H}_S \subset \mathcal{H}_S$.
>
> b) Für $S \cap T = \emptyset$ gilt $\mathcal{H}_S \subset (\mathcal{H}_T)^\perp$.
>
> c) $\sigma(A|_{\mathcal{H}_S}) \subset \overline{S}$.

Der Spektralsatz hat verschiedene Formulierungen. Wir haben bisher die Formulierung durch projektorwertige Maße und den im wesentlichen äquivalenten messbaren Kalkül kennengelernt. Nun gehen wir auch noch kurz auf die sogenannte Multiplikationsoperator-Version ein.

> **Definition 4.52.** Ein Vektor $v \in \mathcal{H}$ heißt *zyklisch* für einen selbstadjungierten Operator $A \in \mathcal{B}(\mathcal{H})$, falls der Unterraum $\operatorname{span}\{A^n v \,:\, n \in \mathbb{N}_0\}$ ein *dichter* Unterraum von $\mathcal{H}$ ist.

> **Beispiel 4.53.**
>
> a) Sei $A = M_{\text{id}}$ der Multiplikationsoperator mit der Identitätsfunktion $\text{id} \in C_b(\mathbb{R})$ auf $L^2([0,1])$. Dann ist die Funktion $f(x) = 1$ ein zyklischer Vektor für A. Denn $\operatorname{span}\{A^n v \,:\, n \in \mathbb{N}_0\}$ enthält alle Polynome auf $[0,1]$, die nach Stone-Weierstraß in $C([0,1])$ dicht liegen, was wiederum in $L^2([0,1])$ dicht liegt.
>
> b) Sei $\mathcal{H}$ beliebig mit $\dim \mathcal{H} > 1$. Dann hat $A = 1$ keine zyklischen Vektoren (klar).

Weitere Beispiele werden in den Übungen besprochen.

Wir betrachten nun den speziellen Fall, dass A einen zyklischen Vektor hat und zeigen, dass A in diesem Fall "ein Multiplikationsoperator ist".

Satz 4.54. *Sei $A = A^* \in \mathcal{B}(\mathcal{H})$ und sei v ein zyklischer Vektor für A. Sei $\mu_v := \langle v, E^A v\rangle$. Dann gibt es einen unitären Operator $U : \mathcal{H} \to L^2(\mathbb{R}, d\mu_v)$, so dass*

$$(UAU^{-1}f)(x) = xf(x)$$

μ_v-fast überall gilt.

Beweis. Sei $g \in C(\sigma(A))$. Dann gilt

$$\int_{\sigma(A)} |g(\lambda)|^2 \, d\mu_v(\lambda) = \langle v, \int_{\sigma(A)} |g|^2 \, dE^A \, v\rangle = \langle v, g(A)^* g(A) v\rangle = \|g(A)v\|^2.$$

Also ist die Abbildung $\tilde{V} : C(\sigma(A)) \to \mathcal{H}$, $\tilde{V}g := g(A)v$, wohldefiniert und isometrisch, wenn wir $C(\sigma(A))$ als (dichten) Unterraum von $L^2(\mathbb{R}, d\mu_v)$ auffassen. Es gibt also eine isometrische Ausdehnung $V : L^2(\mathbb{R}, d\mu_v) \to \mathcal{H}$. Wegen der Zyklizität von v für A ist das Bild von V dicht, denn es enthält alle Vektoren $A^n v$. Aber ein isometrischer linearer Operator mit dichtem Bild ist unitär[10]

Damit erhalten wir, $g \in L^2(\mathbb{R}, d\mu_v)$

$$AVg = Ag(A)v = V(\mathrm{id} \cdot g) \;\Rightarrow\; V^{-1}AVg = \mathrm{id} \cdot g.$$

□

Dieser Satz besagt, dass jeder selbstadjungierte Operator mit zyklischem Vektor ein Multiplikationsoperator ist. Man kann dieses Ergebnis auch auf beliebige selbstadjungierte Operatoren ausdehnen, da $\mathcal{H}$ stets in eine direkte Summe von Unterräumen $\mathcal{H}_k$ zerlegt werden kann, die jeweils unter A invariant sind und zyklische Vektoren für $A|_{\mathcal{H}_k}$ enthalten (siehe z.B. [RS72, Theorem VII.3]).

Zum Abschluss unserer Diskussion des messbaren Kalküls skizzieren wir die Ausdehnung auf *normale* Operatoren (ohne Details).

Sei $A \in \mathcal{B}(\mathcal{H})$ normal, und

$$A = A_+ + iA_-, \qquad A_+ = \frac{1}{2}(A + A^*), \quad A_- = \frac{1}{2i}(A - A^*)$$

die Zerlegung in selbstadjungierte “Real- und Imaginärteile” $A_\pm = (A_\pm)^*$. Da A normal ist, kommutieren A_+ und A_-. Mit Übung 5, Blatt 8, folgt, dass die Spektralmaße E_S^{A+} und E_T^{A-} für beliebige Borelmengen $S, T \subset \mathbb{R}$ kommutieren.

Wir definieren nun eine Abbildung auf Produktmengen $S \times T \subset \mathbb{R}^2$ durch

$$E_{S\times T} := E_S^{A+} E_T^{A-}.$$

[10] Erinnerung: Ist $U : \mathcal{H} \to \mathcal{H}$ isometrisch mit dichtem Bild und $x \in \mathcal{H}$ mit $x_n = Uy_n \to x$, dann ist wegen der Isometrie von U auch (y_n) eine Cauchyfolge, konvergiert also gegen ein $y \in \mathcal{H}$. Das zeigt $x = Uy$, also ist U surjektiv und damit unitär.

Man sieht leicht, dass dies orthogonale Projektionen sind. Mit etwas mehr Arbeit kann man zeigen, dass E sich zu einem projektorwertigen Maß auf der Borel-σ-Algebra von $\mathbb{R}^2$ fortsetzen lässt. Unter der Identifizierung $\mathbb{R}^2 \cong \mathbb{C}$ erhält man dann ein projektorwertiges Maß auf $\mathbb{C}$ mit kompaktem Träger, das wiederum $A = \int_{\sigma(A)} \lambda \, dE(\lambda)$ erfüllt. Dies führt im Endeffekt auf den Spektralsatz für normale Operatoren:

Theorem 4.55. *Sei $A \in \mathcal{B}(\mathcal{H})$ normal. Dann gibt es ein eindeutiges projektorwertiges Maß E^A auf der Borel-σ-Algebra von $\mathbb{C}$ mit kompaktem Träger, das*

$$A = \int_{\sigma(A)} \lambda \, dE^A(\lambda)$$

erfüllt. Die Abbildung

$$\mathscr{B}(\sigma(A)) \ni f \mapsto \int_{\sigma(A)} f \, dE^A$$

ist der eindeutige messbare Funktionalkalkül für A.

Als Ausblick skizzieren wir, wie Spektraltheorie in der Quantenmechanik Anwendung findet. Historisch gesehen waren diese Anwendungen eine der Hauptmotivationen für die Entwicklung der Spektraltheorie.

In der Quantenmechanik werden observable Größen (kurz: Observable) wie zB Energie, Impuls, Ort, Drehimpuls durch selbstadjungierte Operatoren auf einem Hilbertraum $\mathcal{H}$ beschrieben. Diese Operatoren sind oft nicht beschränkt und damit noch etwas außerhalb der Reichweite unserer Methoden. Nichtdestotrotz lassen sich die Hauptideen auch mit beschränkten Operatoren beschreiben.

Zustände, also Beschreibungen des Zustandes, in dem ein physikalisches System vorliegt, werden (im einfachsten Fall) durch Einheitsvektoren $\psi \in \mathcal{H}$ beschrieben.

In der klassischen Physik hat eine Observable wie zB "x-Koordinate des Ortes" in einem gegebenen Zustand zu jedem Zeitpunkt einen festen Wert. In der Quantenphysik ist dies überraschenderweise nicht der Fall! Mathematisch formuliert ordnet die Quantenmechanik einem Paar aus einer Observablen und einem Zustand nicht eine Zahl (=Messwert), sondern eine Wahrscheinlichkeitsverteilung zu. Es werden prinzipiell nur statistische Aussagen über die Verteilung der Messergebnisse bei häufiger Wiederholung der Messung gemacht.

Diese Situation kann gut mit Spektraltheorie beschrieben werden. Ist $A = A^* \in \mathcal{B}(\mathcal{H})$ unsere Observable und $\psi \in \mathcal{H}$, $\|\psi\| = 1$, unser Zustand, so ist das Spektrum $\sigma(A)$ die Menge der möglichen Messergebnisse von A, und das gesuchte Wahrscheinlichkeitsmaß ist $\mu_\psi^A = \langle v, E^A \psi \rangle$.

Das heißt insbesondere, dass der *Erwartungswert* (Mittelwert von Messergebnissen) von A im Zustand ψ durch

$$\int_{\sigma(A)} \lambda \, d\mu_\psi^A = \langle \psi, A\psi \rangle$$

gegeben ist.

Eine andere interessante Größe ist die Standardabweichung der Messergebnisse vom Mittelwert, die in der Physik *Unschärfe* genannt wird, nämlich

$$\Delta_\psi(A) := \sqrt{\langle \psi, (A - \langle \psi, A\psi \rangle)^2 \psi \rangle} = \sqrt{\int_{\sigma(A)} (\lambda - \langle \psi, A\psi \rangle)^2 d\mu_\psi^A(\lambda)}.$$

Man sieht leicht

$$\Delta_\psi(A) = \|(A - \langle \psi, A\psi \rangle)\psi\| \geq 0,$$

d.h. Zustände mit verschwindender Unschärfe für A sind genau durch die Eigenvektoren von A gegeben (Punktspektrum).

Eine berühmte Beziehung zwischen den Unschärfen zweier selbstadjungierter Operatoren A, B ist die *Unschärferelation*

$$\Delta_\psi(A)\Delta_\psi(B) \geq \frac{1}{2}|\langle \psi, [A, B]\psi \rangle|,$$

die sich elementar beweisen lässt.

Auch in der Beschreibung der *Dynamik* (Zeitentwicklung) spielt die Spektraltheorie eine große Rolle. In der Quantenmechanik wird die Zeitentwicklung durch die Festlegung eines selbstadjungierten Energieoperators H (*Hamilton-Operator*) fixiert, zB $(H\psi)(x) = -\psi''(x) + V(x)\psi(x)$ für ψ in einem (dichten Unterraum von) $L^2(\mathbb{R})$ und V einer Potentialfunktion, um die Bewegung eines Teilches unter dem Kraftfeld $F = -V'$ zu modellieren.

Die Zeitentwicklung eines Zustandsvektors $\psi \in \mathcal{H}$ ist dann durch

$$\psi_t = e^{itH}\psi, \qquad t \in \mathbb{R},$$

beschrieben, wobei die Exponentialfunktion per Funktionalkalkül definiert ist. Da H selbstadjungiert ist, ist e^{itH} unitär (isometrisch), so dass $\|\psi_t\| = 1$ für alle Zeiten t (Erhaltung der Wahrscheinlichkeitsinterpretation).

In typischen Systemen lässt sich zeigen, dass die Dynamik im wesentlichen in zwei Klassen zerfällt: *Gebundene Zustände*, für die ψ_t für alle Zeiten im Wesentlichen in einem endlichen Bereich lokalisiert ist, und *Streuzustände*, für die ψ_t nach $\pm\infty$ läuft. Diese beiden Klassen entsprechen genau dem Punktspektrum von H (gebundene Zustände) und dem kontinuierlichen Spektrum von H (Streuzustände).

5 Kompakte Operatoren

Für lineare Abbildungen $A : X \to Y$ zwischen normierten Räumen X, Y sind die Begriffe von Stetigkeit und Beschränktheit äquivalent (Satz 1.12).

In diesem Abschnitt wenden wir uns einer Klasse von Operatoren zu, die stärkere Eigenschaften als Stetigkeit/Beschränktheit haben und kompakt genannt werden. Wir werden sehen, dass kompakte Operatoren dadurch charakterisitiert sind, dass sie in einem gewissen Sinne "fast endlich-dimensionales Bild" haben. Aufgrund dieser Nähe zu endlichdimensionalen Situationen sind kompakte Operatoren eine der am einfachsten handhabbaren Klassen von Operatoren, was sich auch in ihrer Spektralzerlegung zeigen wird.

5.1 Kompakte Mengen und kompakte Operatoren

Wir erinnern zunächst an einige bereits bekannte Tatsachen zu kompakten Mengen:

- Eine Teilmenge M eines metrischen Raumes ist genau dann kompakt, wenn jede Folge in M eine in M konvergente Teilfolge hat.
- Eine abgeschlossene Teilmenge einer kompakten Teilmenge eines metrischen Raums ist auch kompakt.
- Jede kompakte Teilmenge M eines metrischen Raumes ist abgeschlossen und beschränkt.
- Für Teilmengen von endlichdimensionalen Banachräumen gilt auch die Umkehrung: Die kompakten Mengen sind genau die abgeschlossenen beschränkten Mengen.
- Eine Teilmenge M eines vollständigen metrischen Raumes heißt *prä-kompakt*, wenn ihr Abschluss $\overline{M}$ kompakt ist.

In Beispiel 2.33 haben wir bereits gesehen, dass die abgeschlossene Einheitskugel in einem unendlichdimensionalen Hilbertraum nicht kompakt ist. Wir verallgemeinern diese Aussage nun.

Satz 5.1. *Sei X ein normierter Raum. Dann ist die abgeschlossene Einheitskugel $B := \{x \in X : \|x\| \leq 1\}$ genau dann kompakt, wenn* $\dim X < \infty$.

Beweis. Wir zeigen zuerst folgende Aussage: Ist $X \neq U \subset X$ ein abgeschlossener Unterraum und $0 < \delta < 1$, so existiert $x \in X$ mit $\|x\| = 1$ und

$$\|x - u\| \geq 1 - \delta, \qquad u \in U.$$

Sei $y \in X \backslash U$. Aufgrund der Abgeschlossenheit von U gilt $d := d(y, U) = \inf\{\|y-u\| : u \in U\} > 0$. Also gibt es $u_0 \in U$ mit $\|y - u_0\| \leq \frac{d}{1-\delta}$. Wir setzen

$$x := \frac{y - u_0}{\|y - u_0\|}.$$

Dann gilt $\|x\| = 1$, und für beliebiges $u \in U$

$$\begin{aligned} \|x - u\| &= \left\| \frac{y}{\|y - u_0\|} - \frac{u_0}{\|y - u_0\|} - u \right\| = \frac{1}{\|y - u_0\|} \Big\| y - (u_0 + \|y - u_0\| u) \Big\| \\ &\geq \frac{d}{\|y - u_0\|} > 1 - \delta. \end{aligned}$$

Dies beendet den Beweis der Hilfsaussage.

Sei nun $\dim X = \infty$. Wir müssen zeigen, dass B nicht kompakt ist. Da B abgeschlossen ist, müssen wir eine Folge in B konstruieren, die keine konvergente Teilfolge hat. Dazu wählen wir $x_1 \in X$ mit $\|x_1\| = 1$ beliebig und wenden die Hilfsaussauge auf den eindimensionalen Raum $U_1 := \operatorname{span}\{x_1\}$ an. Damit finden wir $x_2 \in X$ mit $\|x_2\| = 1$ und $\|x_1 - x_2\| \geq \frac{1}{2}$. Nun betrachten wir den abgeschlossenen Unterraum $U_2 := \operatorname{span}\{x_1, x_2\}$ und wenden wieder die Hilfaussage an, was einen Vektor x_3 mit $\|x_3\| = 1$ und $\|x_1 - x_3\| \geq \frac{1}{2}$, $\|x_2 - x_3\| \geq \frac{1}{2}$ ergibt.

Induktiv erhalten wir so eine Folge von normierten Vektoren (x_n) mit $\|x_n - x_m\| \geq \frac{1}{2}$ für $n \neq m$. Diese Folge hat offensichtlich keine Cauchy-Teilfolge und deshalb erst recht keine konvergente Teilfolge. □

Definition 5.2. Seien X, Y Banachräume und $A : X \to Y$ linear. Dann heißt A *kompakt*, falls das Bild $A(X_1)$ des abgeschlossenen Einheitsballs in X präkompakt in Y ist, dh falls $\overline{A(X_1)} \subset Y$ kompakt ist.

Die Menge aller kompakten Abbildungen zwischen X und Y wird mit $\mathcal{K}(X, Y)$ bezeichnet, und wie üblich setzen wir $\mathcal{K}(X) := \mathcal{K}(X, X)$.

Bemerkungen:

- Ein Operator $A : X \to Y$ ist genau dann kompakt, wenn für jede beschränkte Menge $M \subset X$ das Bild $A(M) \subset Y$ präkompakt ist.
  Beweis: Diese Aussage impliziert sicherlich unsere Definition, da $X_1 \subset X$ beschränkt ist. Für die andere Implikation sei $M \subset X$ beschränkt. Per Skalierung dürfen wir $M \subset X_1$ annehmen. Also ist $\overline{A(M)}$ eine abgeschlossene Teilmenge der kompakten Menge $\overline{A(X_1)}$, also kompakt. □
- Kompakte Operatoren sind beschränkt, dh $\mathcal{K}(X, Y) \subset \mathcal{B}(X, Y)$, da $A \in \mathcal{K}(X, Y)$ den Einheitsball $X_1 = \{x \in X : \|x\|_X \leq 1\}$ auf eine beschränkte Menge abbildet, es also $C > 0$ mit $A(X_1) \subset \{y \in Y : \|y\|_Y \leq C\}$ gibt. Das impliziert $\|A\| \leq C < \infty$.
- Beschränkte Operatoren sind im Allgemeinen nicht kompakt, zB ist die Identität $\mathrm{id} : X \to X$ auf einem unendlichdimensionalen Banachraum nicht kompakt: Wäre id kompakt, so wäre auch $\mathrm{id}(X_1) = X_1$ kompakt, was nach Satz 5.1 nicht der Fall ist.

Für die nächste Definition erinnern wir uns daran, dass der *Rang* einer linearen Abbildung $A : X \to Y$ zwischen zwei Vektorräumen X, Y als

$$\operatorname{rank} A := \dim \operatorname{Ran} A$$

definiert ist. Wir definieren dann

$$\mathcal{F}(X, Y) := \{A \in \mathcal{B}(X, Y) : \operatorname{rank} A < \infty\}$$

als den Vektorraum aller Operatoren mit endlichem Rang.

Operatoren mit endlichem Rang sind kompakt, dh $\mathcal{F}(X, Y) \subset \mathcal{K}(X, Y)$. Denn in diesem Fall ist $\overline{A(X_1)}$ eine beschränkte abgeschlossene Teilmenge des endlichdimensionalen abgeschlossenen normierten Raums $\operatorname{Ran} X$, also kompakt.

Beispiel 5.3.

a) Sei X ein Banachraum $A \in \mathcal{B}(X)$ und $\lambda \in \mathbb{K}\backslash\{0\}$ so, dass $\dim\ker(A-\lambda) = \infty$. Dann ist A nicht kompakt.
Beweis: Sei $(x_n)_{n\in\mathbb{N}} \in \ker(A-\lambda)$ eine beschränkte Folge ohne konvergente Teilfolge. Eine solche Folge existiert wegen $\dim\ker(A-\lambda) = \infty$ (Satz 5.1). Dann hat die Folge $Ax_n = \lambda x_n$ auch keine konvergente Teilfolge. Also ist A nicht kompakt. □

b) Ist X oder Y endlichdimensional, so gilt $\mathcal{B}(X,Y) = \mathcal{K}(X,Y)$ (alle beschränkten Operatoren sind kompakt).
Beweis: In diesem Fall hat jeder Operator endlichen Rang. □

Der folgende Satz thematisiert zwei zentrale Eigenschaften von $\mathcal{K}(X,Y)$.

Satz 5.4. *Seien X, Y, Z Banachräume.*

a) $\mathcal{K}(X,Y) \subset \mathcal{B}(X,Y)$ ist ein abgeschlossener Unterraum in Operatornorm und damit ein Banachraum.

b) Für $A \in \mathcal{B}(Y,Z)$ und $B \in \mathcal{B}(X,Y)$ ist AB kompakt, falls A oder B kompakt ist.

Beweis. a) Wir zeigen zuerst, dass $\mathcal{K}(X,Y)$ ein Vektorraum ist. Es ist klar, dass mit $A \in \mathcal{K}(X,Y)$ auch $cA \in \mathcal{K}(X,Y)$ (mit $c \in \mathbb{C}$ beliebig). Seien nun $A, B \in \mathcal{K}(X,Y)$ und (x_n) eine beschränkte Folge in X. Aufgrund der Kompaktheit von A gibt es eine Teilfolge $(x_{n_k})_k$, so dass $(Ax_{n_k})_k$ konvergiert, und aufgrund der Kompaktheit von B gibt es eine Teilfolge $(x_{n_{k_l}})_l$, so dass $(Bx_{n_{k_l}})_l$ konvergiert. Da auch $(Ax_{n_{k_l}})_l$ konvergiert, sehen wir, dass $A+B$ kompakt ist. Also ist $\mathcal{K}(X,Y)$ ein Vektorraum.

Wir zeigen nun, dass $\mathcal{K}(X,Y)$ in Operatornorm abgeschlossen ist. Sei dazu $(A_n)_{n\in\mathbb{N}} \subset \mathcal{K}(X,Y)$ eine gegen $A \in \mathcal{B}(X,Y)$ konvergente Folge. Wir müssen $A \in \mathcal{K}(X,Y)$ zeigen.

Sei $(x_k)_{k\in\mathbb{N}} \subset X$ eine beschränkte Folge mit $\|x_k\| \le 1$ für alle $k \in \mathbb{N}$. Da A_1 kompakt ist, gibt es eine Teilfolge $x_n^{(1)} := x_{k_n}$, so dass $(A_1 x_n^{(1)})_n$ konvergiert. Da A_2 kompakt ist, gibt es eine Teilfolge $x_l^{(2)} := x_{n_l}^{(1)}$, so dass $(A_2 x_l^{(2)})_l$ konvergiert. Beachten Sie, dass auch $(A_1 x_l^{(2)})_l$ als Teilfolge der konvergenten Folge $(A_1 x_n^{(1)})_n$ konvergiert.

Induktiv erhalten wir so Folgen $(x_k^{(n)})_{k\in\mathbb{N}}$, so dass $(A_j x_k^{(n)})_{k\in\mathbb{N}}$ für alle $j \in \{1,\dots,n\}$ konvergieren. Sei nun $\xi_k := x_k^{(k)}$ die "Diagonalfolge". Per Konstruktion ist die Folge $(\xi_k)_k$ eine Teilfolge von $x^{(k)}$ ab Folgeglied k. Also konvergiert $(A_n \xi_k)_{k\in\mathbb{N}}$ für alle $n \in \mathbb{N}$.

Um den Beweis abzuschließen, müssen wir die Konvergenz von $(A\xi_n)_{n\in\mathbb{N}}$ zeigen. Sei dazu $\varepsilon > 0$. Da $A_n \to A$ in Operatornorm, gibt es $n \in \mathbb{N}$ mit $\|A_n - A\| < \varepsilon$. Da $(A_n\xi_k)_k$ konvergiert, finden wir k_0 mit $\|A_n\xi_i - A_n\xi_j\| < \varepsilon$ für $i, j \ge k_0$. Für diese Indexwerte gilt dann

$$\|A\xi_i - A\xi_j\| \le \|A\xi_i - A_n\xi_i\| + \|A_n\xi_i - A_n\xi_j\| + \|A_n\xi_j - A\xi_j\| \le 3\varepsilon,$$

wobei wir $\|\xi_m\| \le 1$ benutzt haben.

b) Sei $A \in \mathcal{K}(Y, Z)$ kompakt und $B \in \mathcal{B}(X, Y)$ beschränkt. Dann ist für jede beschränkte Folge $(x_n) \subset X$ auch $(Bx_n)_n \subset Y$ eine beschränkte Folge. Wegen der Kompaktheit von A gibt es eine Teilfolge, so dass $(ABx_{n_k})_k$ konvergent ist. Also ist $AB \in \mathcal{K}(X, Z)$ kompakt.

Das Argument für den Fall, dass A beschränkt und B kompakt ist, verläuft analog. □

Die letzte Eigenschaft besagt insbesondere, dass $\mathcal{K}(X) \subset \mathcal{B}(X)$ ein beidseitiges abgeschlossenes *Ideal* ist[11].

5.2 Kompakte Operatoren auf Hilberträumen

Wir betrachten nun speziell kompakte Operatoren auf Hilberträumen, also die Menge $\mathcal{K}(\mathcal{H}) \subset \mathcal{B}(\mathcal{H})$ mit $\mathcal{H}$ einem komplexen Hilbertraum. Zuerst zeigen wir, dass selbstadjungierte kompakte Operatoren ein vergleichsweise einfaches Spektrum haben. Um die Sprechweise zu vereinfachen, schließen wir den trivialen endlichdimensionalen Fall $\dim \mathcal{H} < \infty$ aus.

Satz 5.5. *Sei $A \in \mathcal{K}(\mathcal{H})$ selbstadjungiert und kompakt auf einem unendlichdimensionalen Hilbertraum $\mathcal{H}$. Dann gibt es eine Folge $(\lambda_n)_{n\in\mathbb{N}} \subset \mathcal{H}$ mit $\lim_n \lambda_n = 0$, so dass*

$$\sigma(A) = \{\lambda_n : n \in \mathbb{N}\} \cup \{0\}.$$

Die Zahlen λ_n, $n \in \mathbb{N}$, liegen im Punktspektrum $\sigma_p(A)$, und es gilt

$$\dim \ker(A - \lambda_n) < \infty$$

für alle $n \in \mathbb{N}$.

Beweis. Sei $\lambda \in \sigma(A)$. Dann verschwinden die Projektionen $E_n := E^A_{(\lambda-1/n,\lambda+1/n)}$ für kein $n \in \mathbb{N}$ (Lemma 4.51 a)). Wir wählen zu jedem $n \in \mathbb{N}$ einen normierten Vektor $v_n \in E_n\mathcal{H}$. Dann gilt einerseits per Spektralkalkül

$$\|(A - \lambda)v_n\|^2 = \int_{\lambda-1/n}^{\lambda+1/n} |x - \lambda|^2 d\langle v_n, E^A(\lambda)v_n\rangle \leq \frac{1}{n^2},$$

also $(A - \lambda)v_n \to 0$. Andererseits konvergiert aufgrund der Kompaktheit von A die Folge $(Av_n)_n$ (möglicherweise nach Übergang zu einer Teilfolge). Also muss auch $(\lambda v_n)_n$ konvergieren. Falls $\lambda \neq 0$ impliziert dies die Konvergenz der Folge (v_n), dh $v_n \to v$, und wir sehen $Av = \lambda v$. Da $v \neq 0$ (die Vektoren v_n sind normiert), folgt $\lambda \in \sigma_p(A)$. Wir haben also gezeigt:

$$\sigma(A)\backslash\{0\} = \sigma_p(A)\backslash\{0\}.$$

Bis auf 0 (diese Zahl kann im Punktspektrum oder im kontinuierlichen Spektrum liegen) besteht $\sigma(A)$ also nur aus Eigenwerten.

Wir zeigen als nächstes, dass $\sigma(A)$ außer 0 keine Häufungspunkte hat, dh dass (λ_n) eine Nullfolge ist. Sei dazu λ ein Häufungspunkt der Folge (λ_n), dh $\lambda_{n_k} \to \lambda$ für $k \to \infty$. Aufgrund der Abgeschlossenheit von $\sigma(A)$ folgt $\lambda \in \sigma(A) = \sigma_p(A)$. Wiederholen wir nun das obige

[11] Eine Teilmenge $I \subset \mathcal{A}$ einer Algebra $\mathcal{A}$ heißt *beidseitiges Ideal*, falls I ein Vektorraum ist und falls gilt: $A \in I, B \in \mathcal{A} \Rightarrow AB \in I, BA \in I$.

Argument, so erhalten wir eine Folge von normierten Eigenvektoren $v_k \in \ker(A - \lambda_{n_k})$ mit $v_k \to v \in \ker(A - \lambda)$. Aber das ist unmöglich, da Eigenvektoren zu unterschiedlichen Eigenwerten aufeinander senkrecht stehen. Also ist 0 der einzige mögliche Häufungspunkt des Spektrums.

Wir wissen bereits aus Lemma 5.3 a), dass die Multiplizitäten $\dim \ker(A - \lambda_n)$ aufgrund der Kompaktheit von A endlich sind. Wäre $\sigma(A)$ endlich, müsste also $\dim \mathcal{H}$ endlich sein, im Widerspruch zu unserer Annahme. Also ist $\sigma(A)$ unendlich und muss deshalb eine Nullfolge sein. □

Korollar 5.6. *Ein selbstadjungierter Operator $A = A^* \in \mathcal{B}(\mathcal{H})$ ist kompakt genau dann, wenn es eine Folge $(P_n)_{n\in\mathbb{N}} \subset \mathcal{H}$ von aufeinander senkrecht stehenden orthogonalen Projektionen P_n von endlichem Rang mit $\sum_n P_n = 1$ und eine Nullfolge $(\lambda_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ gibt, so dass*

$$A = \sum_{n=1}^{\infty} \lambda_n P_n.$$

Insbesondere besitzt $\mathcal{H}$ eine aus Eigenvektoren von A bestehende Orthonormalbasis. Es gilt $\|A\| = \max\{|\lambda_n| : n \in \mathbb{N}\}$.

Beweis. Sei A selbstadjungiert und kompakt. Dann gibt es eine Nullfolge (λ_n) mit $\sigma(A) = \{\lambda_n : n \in \mathbb{N}\} \cup \{0\}$, und jedes λ_n ist ein Eigenwert von A.

Da das Spektralmaß E^A von A σ-additiv ist, gilt

$$1 = E^A_{\mathbb{R}} = \sum_n E^A_{\{\lambda_n\}} + E^A_{\{0\}}.$$

Entweder gilt $0 \in \sigma_p(A)$ (und dann $E^A_{\{0\}} \neq 0$), oder $0 \in \sigma_c(A)$ (und dann $E^A_{\{0\}} = 0$). In jedem Fall ergibt der Spektralsatz

$$A = \int \lambda \, dE^A(\lambda) = \sum_n \lambda_n E^A_{\{\lambda_n\}}.$$

Die $P_n := E^A_{\{\lambda_n\}}$ haben nach Satz 5.5 endlichen Rang.

Wählen wir in jedem Eigenraum eine Orthonormalbasis, so ergibt deren Vereinigung eine Orthonormalbasis von $\mathcal{H}$, die aus Eigenvektoren von A besteht. Die Behauptung über die Norm folgt aus, $v \in \mathcal{H}$,

$$\|Av\|^2 = \sum_n |\lambda_n|^2 \|P_n v\|^2 \leq \max\{|\lambda_n|^2 : n \in \mathbb{N}\} \cdot \|v\|^2.$$

Um zu zeigen, dass jeder Operator der angegebenen Form auch kompakt ist, sei $A = \sum_n \lambda_n P_n$. Wir sortieren die λ_n in dem Betrage nach absteigender Reihenfolge, $|\lambda_1| \geq |\lambda_2| \geq \ldots$ und definieren $A_N := \sum_{n=1}^N \lambda_n P_n$. Dann hat A_N endlichen Rang, und es gilt $\|A - A_N\| = \|\sum_{n>N} \lambda_n P_n\| = |\lambda_{N+1}|$, was für $N \to \infty$ gegen Null konvergiert. Also ist A ein Normlimes von Operatoren mit endlichem Rang und damit kompakt. □

Wir verallgemeinern diese Normalform nun auf allgemeine, nicht notwendigerweise selbstadjungierte, kompakte Operatoren.

Satz 5.7 (Normalform von kompakten Operatoren).

a) Sei $A \in \mathcal{B}(\mathcal{H})$ kompakt. Dann gibt es zwei orthonormale Folgen (e_n) und (f_n) und eine Nullfolge von nicht-negativen reellen Zahlen (λ_n), so dass

$$Av = \sum_n \lambda_n \langle e_n, v\rangle f_n, \qquad v \in \mathcal{H}.$$

Weiterhin sind alle Operatoren dieser Form kompakt.

b) Falls A kompakt und normal ist, gibt es eine orthonormale Folge (e_n) und eine Nullfolge von komplexen Zahlen (ζ_n), so dass

$$Av = \sum_n \zeta_n \langle e_n, v\rangle e_n, \qquad v \in \mathcal{H}.$$

Weiterhin sind alle Operatoren dieser Form kompakt und normal.

Beweis. a) Wir betrachten die Polarzerlegung $A = V|A|$ von A. Wegen Satz 5.4 b) ist mit A auch A^*A kompakt, hat nach Satz 5.5 also die Form $A^*A = \sum_n \mu_n P_n$ mit $\mu_n \geq 0$. Demnach ist $|A| = \sqrt{A^*A} = \sum_n \sqrt{\mu_n} P_n$ auch kompakt.

Die partielle Isometrie V erfüllt $\ker V = \ker A$ und bildet $(\ker A)^\perp$ isometrisch auf den abgeschlossenen Unterraum $\operatorname{Ran} V$ ab. Nach Satz 5.5 gibt es eine Orthonormalbasis (e_n) von $\mathcal{H}$ mit $|A|e_n = \lambda_n e_n$ (hier sind die λ_n entsprechend Multiplizität wiederholt). Da V auf $(\ker A)^\perp$ eine Isometrie ist, bilden die Vektoren $\{f_n := Ve_n : n \in \mathbb{N}, \lambda_n \neq 0\}$, eine Orthonormalbasis von $\operatorname{Ran} V$.

Damit folgt $Ae_n = V|A|e_n = \lambda_n f_n$, also

$$Av = \sum_n \lambda_n \langle e_n, v\rangle f_n,$$

wie behauptet.

Um die Kompaktheit der angegebenen Operatoren einzusehen, argumentieren wir wie in Satz 5.5. Es gilt $\|A\| = \max\{\lambda_n\}$, und $A_N := \sum_{n=1}^N \lambda_n \langle e_n, v\rangle f_n$ (mit absteigend sortierten λ_n) ist eine Folge von Operatoren von endlichem Rang, die für $N \to \infty$ gegen A konvergieren.

b) Zuerst bemerken wir, dass der Operator in a) $Ae_k = \lambda_k \cdot f_k$ und

$$A^*v = \sum_n \lambda_n \langle f_n, v\rangle e_n$$

erfüllt. Das ergibt für jedes $k \in \mathbb{N}$

$$\lambda_k^2 = \langle e_k, A^*Ae_k\rangle = \langle e_k, AA^*e_k\rangle = \left\| \sum_n \lambda_n \langle f_n, e_k\rangle e_n \right\|^2 = \sum_n \lambda_n^2 |\langle f_n, e_k\rangle|^2. \qquad (\star)$$

Wir wählen die λ_k in absteigender Sortierung, $\lambda_1 \geq \lambda_2 \geq \ldots \geq 0$, und betrachten $k = 1$. Dann ergibt Abschätzung per Dreiecksungleichung $\lambda_1^2 = \sum_n \lambda_n^2 |\langle f_n, e_1 \rangle|^2 \leq \lambda_1^2$, da die Basisentwicklungskoeffizienten $\sum_n |\langle f_n, e_1 \rangle|^2 = 1$ erfüllen. Also $0 = \sum_n (\lambda_1^2 - \lambda_n^2) |\langle f_n, e_1 \rangle|^2$, was $\langle f_n, e_1 \rangle = 0$ für $n > 1$ und damit (Cauchy-Schwarz-Gleichung) $f_1 = \alpha_1 e_1$ mit einem $\alpha_1 \in \mathbb{C}$, $|\alpha_1| = 1$, bedeutet.

Betrachten wir nun λ_2 und wenden A^* auf e_2 an, so können wir dieses Argument iterieren – der Term $n = 1$ in $(\star)$ trägt wegen $f_1 \perp e_2$ nicht bei. Dies zeigt, dass es für jedes $k \in \mathbb{N}$ ein $\alpha_k \in \mathbb{C}$, $|\alpha_k| = 1$, gibt, so dass $f_k = \alpha_k e_k$. Dies zeigt die Behauptung mit $\zeta_n = \alpha_n \lambda_n$. $\square$

Die Zahlen λ_n in obigem Satz sind die Eigenwerte von $|A|$. Sie werden *Singulärwerte* von A genannt.

Korollar 5.8. *Sei $\mathcal{H}$ ein Hilbertraum.*

a) Es gilt $\mathcal{K}(\mathcal{H}) = \overline{\mathcal{F}(\mathcal{H})}$, dh jeder kompakte Operator ist ein Normlimes von Operatoren von endlichem Rang.

b) $A \in \mathcal{B}(\mathcal{H})$ ist genau dann kompakt, wenn A^* kompakt ist.

Der Beweis ergibt sich direkt aus dem vorigen Satz. Für allgemeine Banachräume ist die erste Aussage falsch! Die zweite Aussage hat eine Verallgemeinerung auf Banachräume, wenn A^* durch ein Banachraum-Adjungiertes ersetzt wird.

5.3 Die Fredholmsche Alternative

Kompakte Operatoren haben viele Gemeinsamkeiten mit Operatoren auf endlichdimensionalen Räumen. Zum Beispiel stimmt für $A \in \mathcal{B}(X)$ mit einem endlichdimensionalen Raum X die folgende Aussage (mit fixiertem $\lambda \in \mathbb{C}$): Entweder gibt es ein $x \in X \backslash \{0\}$ mit $Ax = \lambda x$, oder das Inverse $(A - \lambda)^{-1}$ existiert.

Wir werden nun zeigen, dass für kompaktes A auf einem Banachraum diese Aussage für $\lambda \neq 0$ richtig bleibt. Für allgemeine Operatoren ist die Aussage sicher nicht richtig (betrachte kontinuierliches Spektrum) und für kompakte *normale* Operatoren auf einem Hilbertraum ist sie sicher richtig, da wir bereits wissen, dass diese Operatoren außerhalb von $\{0\}$ nur Punktspektrum haben.

Im Folgenden wollen wir aber allgemeine nicht notwendigerweise normale kompakte Operatoren betrachten, so dass wir anders argumentieren müssen.

Satz 5.9 (Die Fredholmsche Alternative). *Sei X ein Banachraum, $\lambda \in \mathbb{C} \backslash \{0\}$, und $A \in \mathcal{K}(X)$ kompakt. Dann ist genau eine der folgenden beiden Aussagen wahr:*

a) Es gibt $x \in X \backslash \{0\}$ mit $Ax = \lambda x$.

b) $A - \lambda$ ist invertierbar.

Beweis. Wir müssen zeigen: Falls $\lambda \neq 0$ kein Eigenwert von A ist (dh wenn $A - \lambda$ injektiv ist), so ist $(A - \lambda)$ invertierbar (also bijektiv). Wir nehmen also $\lambda \neq 0$ als Nicht-Eigenwert an, und nehmen auch $\mathrm{Ran}(A - \lambda) \neq X$ an (Widerspruchsbeweis).

Als ersten Schritt behaupten wir: Es gibt $c > 0$, so dass

$$\|(A - \lambda)x\| \geq c\|x\|, \qquad x \in X.$$

Es genügt, $\|x\| = 1$ zu betrachten. Angenommen, die Behauptung wäre falsch: Dann gäbe es normierte Vektoren x_n mit $\|(A - \lambda)x_n\| \leq \frac{1}{n}$. Da A kompakt ist, konvergiert die Folge $(Ax_n)_n$ (eventuell nach Übergang zu einer Teilfolge). Das impliziert (wegen $\lambda \neq 0$), dass die Folge (x_n) konvergiert. Sei $x := \lim_n x_n$. Dann ist $x \neq 0$ (da $\|x_n\| = 1$) und $Ax = \lambda x$, ein Widerspruch. Wir haben bereits in Lemma 4.9 gezeigt, dass diese untere Schranke impliziert, dass $\operatorname{Ran}(A - \lambda)$ abgeschlossen ist.

Insbesondere ist $A - \lambda : X \to \operatorname{Ran}(A - \lambda)$ ein Isomorphismus von Banachräumen. Wir betrachten nun die Räume $R_n := \operatorname{Ran}((A - \lambda)^n)$, die offenbar $R_{n+1} \subset R_n$ für alle $n \in \mathbb{N}$ erfüllen. R_1 ist abgeschlossen, und mit demselben Argument wie oben sehen wir, dass alle R_n abgeschlossen sind. Wir behaupten $R_{n+1} \neq R_n$ für alle n. Dazu wählen wir $x \notin \operatorname{Ran}(A - \lambda)$ (möglich nach Annahme). Dann gilt $(A - \lambda)^n x \in R_n$. Falls $(A - \lambda)^n x = (A - \lambda)^{n+1} y$ für ein $y \in X$, führt die Injektivität von $(A - \lambda)$ auf $x = (A - \lambda)y$, ein Widerspruch.

Wir benutzen nun die Hilfaussage aus dem Beweis von Satz 5.1 (Rieszsches Lemma), angewandt auf die abgeschlossenen nichtrivialen Unterräume $R_{n+1} \subset R_n \subset \ldots \subset X$. Dies gibt uns Vektoren $x_n \in R_n$ mit $d(x_n, R_{n+1}) \geq \frac{1}{2}$ und $\|x_n\| = 1$.

Dies gibt eine beschränkte Folge (x_n), so dass, evtl nach Übergang zu einer Teilfolge, $(Ax_n)_n$ konvergiert. Für $n > m$ gilt per Konstruktion, dass $x_n, (A - \lambda)x_n, (A - \lambda)x_m$ alle in R_{m+1} liegen, und deshalb $Ax_n - Ax_m \in \lambda x_m + R_{m+1}$. Da x_m mindestens Abstand $\frac{1}{2}$ von R_{m+1} hat, folgt

$$\|Ax_n - Ax_m\| \geq \frac{|\lambda|}{2} > 0, \qquad n > m.$$

Aber dies ist ein Widerspruch zur Konvergenz von $(Ax_n)_n$. □

Korollar 5.10. *Sei X ein Banachraum und $A \in \mathcal{K}(X)$ kompakt. Dann ist jedes $\lambda \in \sigma(A) \backslash \{0\}$ ein Eigenwert von A mit endlicher Multiplizität.*

Beispiel 5.11. Viele Integraloperatoren sind kompakt. Zum Beispiel gilt für $K \in L^2(\mathbb{R}^2)$, dass der Operator

$$A : L^2(\mathbb{R}) \to L^2(\mathbb{R}), \qquad (Af)(x) = \int_{\mathbb{R}} K(x, y) f(y) dy,$$

kompakt ist (sogar Hilbert-Schmidt). Ein weiteres Beispiel sind Integraloperatoren von Volterra-Typ, dh $L \in C([0, 1])^2$

$$A : C([0, 1]) \to C([0, 1]), \qquad (Af)(x) = \int_0^x L(x, y) f(y) dy.$$

Auch dieser Operator ist kompakt (hier ohne Beweis).

Die Fredholmsche Alternative besagt in diesem Kontext: Falls die Gleichung (mit $\lambda \neq 0$)

$$Af = \lambda f$$

nur die triviale Lösung $f = 0$ hat (also λ kein Eigenwert von A ist), dann hat die inhomogene Gleichung

$$Af - \lambda f = g$$

für alle $g \in X$ eine eindeutige Lösung. Falls die homogene Gleichung $Af - \lambda f = 0$ hingegen von Null verschiedene Lösungen hat, so ist die inhomogene Gleichung $Af - \lambda f = g$ nicht eindeutig lösbar.

Aus der Eindeutigkeit der Lösung des homogenen Problems folgt also die Existenz einer eindeutigen Lösung des inhomogenen Problems. Mit Hilfe von Neumannreihen lässt sich diese Lösung oft als Grenzwert einer Reihe angeben.