Funktionalanalysis und Integralgleichungen

Steffen Roch

SS 2007

Funktionalanalysis + Integralgleichungen

Literatur	Heuser	Funktionalanalysis	(Teubner)
	Mathieu	Funktionalanalysis	(Spektrum)
	Schröder	Funktionalanalysis	(Akademie-Verlag)
	Hackbusch	Integralgleichungen	(Teubner)
	Kress	Linear Integral Equations	(Springer)
	Gohberg/Goldberg	Basic Operator Theory	(Birkhäuser)

0 Einleitung

Die Herausbildung der Funktionalanalysis als eigenständige mathematische Disziplin vollzog sich im wesentlichen in der ersten Hälfte des 20. Jahrhunderts im Zusammenhang mit dem Bemühen, eine Theorie der Integralgleichungen aufzubauen (Hilbert, Schmidt, Fredholm, Banach). Ihre Ursprünge reichen jedoch viel weiter zurück (Fourier, Euler). Heute ist die Funktionalanalysis eine breitgefächerte Theorie mit zahlreichen Verbindungen zu anderen Gebieten der Mathematik und mit vielfältigen Anwendungen, etwa in der Physik und der Mechanik. Einige dieser Richtungen, auf die wir hier nicht eingehen können, werden in weiterführenden Vorlesungen behandelt: Nichtlineare Funktionalanalysis, Topologische Räume und lokalkonvexe Räume, Sobolevräume und Distributionen, Spektraltheorie für beschränkte und unbeschränkte Operatoren, Banach- und C^*-Algebren, Darstellungstheorie, Anwendungen in der Numerischen Analysis, ...
Das Ziel dieser Vorlesung ist dagegen recht bescheiden. Wir wollen uns (bis auf den ersten Abschnitt) nur mit Grundbegriffen der *linearen* Funktionalanalysis befassen. Als Leitmotiv wird uns dabei immer wieder das Problem der Lösbarkeit von *linearen Gleichungen*

$$A\,x = y \tag{0.1}$$

dienen, wobei y ein gegebenes Element eines linearen Raumes X ist und A eine lineare Abbildung von X nach X. Solche Probleme kennen wir aus der linearen Algebra, wo etwa $y \in X = \mathbb{R}^n$ ist und A eine $n \times n$-Matrix. In diesem Fall kann man sich (0.1) als lineares Gleichungssystem von n Gleichungen für n Unbekannte vorstellen, über dessen Lösbarkeit wir gut Bescheid wissen. So ist (0.1) genau dann eindeutig lösbar, wenn das homogene System $Ax = 0$ nur die triviale Lösung $x = 0$ besitzt. Wenn dies der Fall ist, so ist die Matrix A invertierbar, und (0.1) ist für *jede rechte* Seite eindeutig lösbar: $x = A^{-1}y$. Wir werden solche Gleichungen in Situationen zu betrachten haben, wo unendlich viele Unbekannte vorliegen und man A etwa als unendliche Matrix auffassen kann, oder wo x, y Funktionen sind und A als Integraloperator wirkt, welcher einer gegebenen Funktion eine neue Funktion zuordnet. Beim Übergang von endlichen zu unendlichen linearen Gleichungssystemen gehen viele der bekannten Eigenschaften verloren.

Beispielsweise ist es im Falle unendlich vieler Unbekannter nicht mehr wahr, dass aus der eindeutigen Lösbarkeit von $A\,x = 0$ die Lösbarkeit von $A\,x = y$ für alle y folgt.

Beispiel 1: Wir betrachten die Menge X aller Vektoren (Folgen) $x = (x_1, x_2, \ldots)$ mit unendlich vielen Einträgen und auf dieser Menge den *Verschiebungsoperator*

$$V : X \to X, \quad (x_1, x_2, x_3, \ldots) \mapsto (0, x_1, x_2, x_3, \ldots).$$

Offenbar folgt aus $Vx = 0$, dass $x = 0$. Die Gleichung $Vx = (1, 0, 0, \ldots)$ besitzt jedoch keine Lösungen. ■

Fragen, die uns im Zusammenhang mit Gleichungen $Ax = y$ besonders interessieren, sind:

(A) Für welche y ist diese Gleichung lösbar?

(B) Ist die Lösung eindeutig bestimmt?

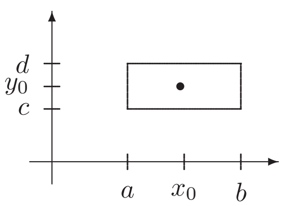
(C) Wie hängt die Lösung x von gegebener rechter Seite y ab? Führen kleine Änderungen von y auch nur zu kleinen Änderungen von x, d.h. ist diese Abhängigkeit *stetig?*

Wir betrachten nun einige Beispiele, welche ausgehend von einfachen Differentialgleichungen auf Lösbarkeitsprobleme für Integralgleichungen führen.

Beispiel 2: Anfangswertprobleme für gewöhnliche Differentialgleichungen. Wir betrachten das Anfangswertproblem

$$y'(x) = f\big(x, y(x)\big), \quad y(x_0) = y_0, \tag{0.2}$$

wobei f eine auf einem Rechteck $[a, b] \times [c, d]$ definierte und stetige Funktion ist und (x_0, y_0) ein Punkt im Inneren dieses Rechtecks. Das AWP (0.1) ist äquivalent zu einer Integralgleichung: Integration von $y' = f(x, y)$ liefert



$$\int_{x_0}^{x} y'(t)dt = \int_{x_0}^{x} f\big(t, y(t)\big)dt, \quad x \in [a, b],$$

und wegen

$$\int_{x_0}^{x} y'(t)dt = y(t)\Big|_{x_0}^{x} = y(x) - y(x_0) = y(x) - y_0$$

erhalten wir

$$y(x) = y_0 + \int_{x_0}^{x} f\big(t, y(t)\big)dt, \quad x \in [a, b]. \tag{0.3}$$

Jede Lösung y von (0.2) löst also die Integralgleichung (0.3) und umgekehrt: Ist y eine stetige Lösung von (0.3), so ist y sogar einmal stetig differenzierbar, und durch Ableiten von (0.3) nach x findet man, dass y auch (0.2) löst. Beide Probleme sind also äquivalent. Die Gleichung (0.3) läßt sich jedoch in mancher Hinsicht bequemer untersuchen. Führen wir nämlich den *Integraloperator*

$$(Ay)(x) := y_0 + \int_{x_0}^{x} f\big(t, y(t)\big)dt, \quad x \in [a,b]$$

ein, so läßt sich (0.3) als Fixpunktgleichung $Ay = y$ schreiben. Für Fixpunktgleichungen hat man Kriterien, die die Lösbarkeit unter geeigneten Voraussetzungen an A garantieren (siehe Abschnitt 1.3). ■

Beispiel 3: Anfangswertprobleme für lineare Differentialgleichungen 2. Ordnung. Seien a, b, f auf einem abgeschlossenen Intervall I stetige Funktionen und x_0 ein Punkt im Inneren von I. Wir betrachten das AWP

$$y''(x) + a(x)y'(x) + b(x)y(x) = f(x), \quad y(x_0) = y_0, \quad y'(x_0) = y_0'. \tag{0.4}$$

Um eine Integralgleichung für y'' zu finden, stellen wir y und y' durch y'' dar. Wie in Beispiel 2 ist

$$y'(x) = y'(x_0) + \int_{x_0}^{x} y''(t)dt, \tag{0.5}$$

und mit partieller Integration erhält man

$$\begin{aligned} \int_{x_0}^{x}(x-t)y''(t)dt &= (x-t)y'(t)\Big|_{x_0}^{x} - \int_{x_0}^{x}(-1)y'(t)dt \\ &= -(x-x_0)y'(x_0) + y(t)\Big|_{x_0}^{x} \\ &= -(x-x_0)y'(x_0) + y(x) - y(x_0), \end{aligned}$$

also

$$y(x) = y(x_0) + y'(x_0)(x-x_0) + \int_{x_0}^{x}(x-t)y''(t)dt \tag{0.6}$$

(was nichts anderes als der Satz von Taylor mit Restglied in Integralform ist). Einsetzen von (0.5) und (0.6) in die Ausgangsdifferentialgleichung liefert

$$\begin{aligned} &y''(x) + a(x)\left(y'(x_0) + \int_{x_0}^{x} y''(t)\,dt\right) \\ &\quad + b(x)\left(y(x_0) + y'(x_0)(x-x_0) + \int_{x_0}^{x}(x-t)y''(t)\,dt\right) = f(x) \end{aligned}$$

bzw.

$$\begin{aligned} &\underbrace{y''(x)}_{=:u(x)} + \int_{x_0}^{x} \underbrace{\big(a(x) + b(x)(x-t)\big)}_{=:k(x,t)} y''(t)\,dt \\ &\quad = \underbrace{f(x) - a(x)y'(x_0) - b(x)\big(y(x_0) + y'(x_0)(x-x_0)\big)}_{=:g(x)} \end{aligned}$$

und damit

$$u(x) + \int_{x_0}^{x} k(x,t)\, u(t)\, dt = g(x), \quad x \in [a,b]. \tag{0.7}$$

Ist umgekehrt u eine stetige Lösung von (0.7), so bilden wir

$$v(x) := y_0' + \int_{x_0}^{x} u(t)\, dt, \quad y(t) = y_0 + \int_{x_0}^{x} v(t)\, dt$$

und bekommen so eine Lösung von (0.4). ∎

Beispiel 4: Randwertprobleme für gewöhnliche Differentialgleichungen 2. Ordnung. Der Einfachheit halber betrachten wir nur Randwertprobleme (RWP) der Form

$$y''(x) = f\big(x, y(x)\big), \quad y(0) = y_0, \quad y(1) = y_1, \tag{0.8}$$

wobei f eine auf dem Rechteck $[0,1] \times [c,d]$ stetige Funktion ist (die erste Ableitung kommt also in unserer RWA nicht vor). Zweimaliges Integrieren der Differentialgleichung liefert

$$y(x) = \int_0^x \left(\int_0^t f\big(s, y(s)\big) ds \right) dt + c_0 + c_1 x \tag{0.9}$$

mit Konstanten c_0, c_1. Nach partieller Integration

$$\int_0^x \left(\int_0^t f\big(s, y(s)\big) ds \right) dt = t \int_0^t f\big(s, y(s)\big)\, ds \Big|_0^x - \int_0^x t\, f\big(t, y(t)\big)\, dt$$

geht (0.9) über in die Integralgleichung

$$y(x) - \int_0^x (x-t)\, f\big(t, y(t)\big)\, dt = c_0 + c_1 x. \tag{0.10}$$

Die Konstanten c_0, c_1 werden durch Einsetzen der Randbedingungen bestimmt:

$x = 0$: $\quad y(0) = c_0 \quad$ bzw. $\quad c_0 := y_0.$

$x = 1$: $\quad y(1) - \int_0^1 (1-t) f\big(t, y(t)\big) dt = c_0 + c_1 \quad$ bzw.

$$c_1 := y_1 - y_0 - \int_0^1 (1-t) f\big(t, y(t)\big) dt.$$

Mit dieser Wahl von c_0, c_1 geht (0.10) über in

$$y(x) - \int_0^x (x-t) f\big(t, y(t)\big) dt - \int_0^1 x(t-1) f\big(t, y(t)\big) dt = y_0 + (y_1 - y_0) x.$$

Mit Hilfe der Funktion

$$k(x,t) := \begin{cases} x - t + x(t-1) & \text{für } 0 \le t \le x \le 1 \\ x(t-1) & \text{für } 0 \le x < t \le 1 \end{cases} = \begin{cases} t(x-1) & \text{für } 0 \le t \le x \le 1 \\ x(t-1) & \text{für } 0 \le x < t \le 1 \end{cases}$$

können wir beide Integrale zu einem zusammenfassen und gelangen zur Integralgleichung

$$y(x) - \int_0^1 k(x,t) f\big(t, y(t)\big)\, dt = y_0 + (y_1 - y_0)x, \quad x \in [0,1]. \tag{0.11}$$

Gehen wir wie in Beispiel 3 von einem *linearen* RWP

$$y''(x) + a(x)y(x) = b(x), \quad y(0) = y_0, \quad y(1) = y_1,$$

mit stetigen Funktionen a und b aus, so gelangen wir von (0.11) zur Integralgleichung

$$y(x) + \int_0^1 \underbrace{k(x,t)a(t)}_{k^*(x,t)} y(t)\, dt = \underbrace{\int_0^1 k(x,t)b(t)\, dt + y_0 + (y_1 - y_0)x}_{f(x)}$$

bzw.

$$y(x) + \int_0^1 k^*(x,t)y(t)\, dt = f(x). \tag{0.12}$$

■

Integralgleichungen werden grob klassifiziert

(a) *nach ihrer Gestalt* in

- Gleichungen 1. Art: die gesuchte Funktion steht nur unter dem Integral
- Gleichungen 2. Art: die gesuchte Funktion steht auch außerhalb des Integrals

(b) *nach dem Integrationsgebiet:*

- Fredholmsche Integralgleichung: Integration erfolgt über einem festen Bereich
- Volterrasche Integralgleichung: Integrationsbereich hängt von der Variablen ab

(c) *nach der Art der Integralbildung:*

- reguläre Integralgleichung: Integral existiert als eigentliches Integral.

- schwach singuläre Gleichung: Integral existiert als uneigentliches Integral
- stark singuläre Integralgleichung: Integral existiert erst nach einer speziellen Regularisierung, z.B. im Sinne des Cauchyschen Hauptwertes.

Daneben unterscheidet man natürlich wie bei Differentialgleichungen auch zwischen linearen und nichtlinearen Gleichungen. So ist z.B.

(0.3) nichtlineare Volterra-Gleichung 2. Art

(0.7) lineare Volterra-Gleichung 2. Art

(0.11) nichtlineare Fredholm-Gleichung 2. Art

(0.12) lineare Fredholm-Gleichung 2. Art.

Die sog. *Abelsche Integralgleichung*

$$\int_0^x \frac{y(t)}{\sqrt{x-t}}\, dt = f(x), \quad x \in (0,1)$$

ist ein Beispiel für eine schwach singuläre lineare Volterra-Gleichung 1. Art, und die *singuläre Integralgleichung*

$$a(x)y(x) + \frac{b(x)}{\pi i} \int_{-1}^{1} \frac{y(t)}{t-x}\, dt = f(x), \quad x \in [-1,1],$$

bei der das Integral im Sinne des Cauchyschen Hauptwertes verstanden werden muss, eine singuläre lineare Fredholm-Gleichung 2. Art.
Diese Klassifikationen erscheinen im Moment recht formal; wir werden jedoch bald sehen, dass sich z.B. Gleichungen 1. und 2. Art in ihren Eigenschaften grundlegend unterscheiden.

1 Metrische Räume und Banachscher Fixpunktsatz

Um über die stetige Abhängigkeit der Lösung einer Integralgleichung von den Ausgangsdaten sprechen zu können oder um beurteilen zu können, wie gut eine numerisch gefundene Näherungslösung die exakte Lösung einer Integralgleichung approximiert, müssen wir Abstände zwischen Funktionen messen. Mit dem Begriff des metrischen Raumes stellen wir einen Rahmen bereit, in dem sich Begriffe wie Abstand oder Stetigkeit hinreichend allgemein diskutieren lassen.

1.1 Metrische Räume

Definition 1.1 *Gegeben sei eine Menge $M(\neq \emptyset)$ und eine Abbildung $d : M \times M \to \mathbb{R}$, d.h. je zwei Elementen $x, y \in M$ sei eine reelle Zahl $d(x, y)$ zugeordnet. Wenn die Abbildung d folgende Eigenschaften für alle $x, y, z \in M$ besitzt:*

(a) $d(x, y) \geq 0$ *und* $d(x, y) = 0$ *genau dann, wenn* $x = y$,

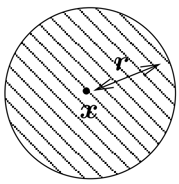
(b) $d(x, y) = d(y, x)$ *(Symmetrie),*

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

so heißt d eine Abstandsfunktion *oder* Metrik *auf M, die Zahl $d(x, y)$ heißt der* Abstand *der Punkte x, y, und das Paar (M, d) heißt* metrischer Raum.

Vertraute Beispiele für metrische Räume sind die Menge $M = \mathbb{R}$ der reellen Zahlen mit dem Abstand $d(x, y) = |x - y|$ und die Menge $M = \mathbb{R}^n$ aller Vektoren mit n Einträgen, versehen mit dem Euklidschen Abstand. Wie für $\mathbb{R}$ oder $\mathbb{R}^n$ führt man für allgemeine metrische Räume die folgenden Begriffe ein.

Definition 1.2 *Sei (M, d) ein metrischer Raum, $x \in M$ und $r > 0$. Unter der* Umgebung von x vom Radius r *oder kurz r-Umgebung von x versteht man die Menge $U_r(x) := \{y \in M : d(x, y) < r\}$.*



Definition 1.3 *Sei (M, d) ein metrischer Raum und $A \subseteq M$.*

(a) *Ein Punkt $a \in A$ heißt* innerer Punkt *von A, wenn es eine Umgebung von a gibt, die ganz in A liegt.*

(b) *Ein Punkt $a \in M$ heißt* Häufungspunkt *von A, wenn jede Umgebung von a einen Punkt aus A enthält, der von a verschieden ist.*

(c) *A heißt* offen, *wenn jeder Punkt von A ein innerer Punkt ist.*

(d) *A heißt* abgeschlossen*, wenn jeder Häufungspunkt von A zu A gehört.*

(e) *Die Menge aller inneren Punkte von A heißt das* Innere *von A (Bezeichnung:* int A*).*

(f) *Die Vereinigung von A mit der Menge aller Häufungspunkte von A heißt* Abschließung *von A (Bezeichnung:* $\overline{A}$ *oder* $\operatorname{clos} A$*).*

(g) *Der* Rand *von A ist die Menge* $\partial A := \operatorname{clos} A \setminus \operatorname{int} A$.

(h) *A heißt* beschränkt*, wenn es ein* $x \in M$ *und ein* $r > 0$ *gibt mit* $A \subseteq U_r(x)$.

Bevor wir uns Beispiele metrischer Räume ansehen, überlegen wir uns zwei wichtige Ungleichungen.

Lemma 1.4 (Hölder-Ungleichung) *Seien* $p, q > 1$ *reelle Zahlen mit* $\frac{1}{p} + \frac{1}{q} = 1$. *Dann gilt für beliebige Vektoren* $(a_1, \ldots, a_n)$, $(b_1, \ldots, b_n) \in \mathbb{R}^n$ *(oder* $\in \mathbb{C}^n$*):*

$$\Big| \sum_{j=1}^{n} a_j b_j \Big| \leq \big(\sum_{j=1}^{n} |a_j|^p \big)^{1/p} \big(\sum_{j=1}^{n} |b_j|^q \big)^{1/q}.$$

Beweis Ist die rechte Seite der zu beweisenden Ungleichung gleich 0, so sind alle a_j oder alle b_j gleich 0, und die Ungleichung ist offenbar richtig. Wir beweisen die Ungleichung nun für positive rechte Seiten. Für jedes fixierte $t \in (0, 1)$ betrachten wir die Funktion $f(x) = x^t - tx$ auf $[0, 1]$. Offenbar ist $f(0) = 0$ und $f(1) = 1 - t$, und wegen

$$f'(x) = tx^{t-1} - t = t(x^{t-1} - 1) > 0 \quad \text{für } x \in (0, 1)$$

ist f auf $(0, 1)$ streng monoton wachsend. Insbesondere gilt also

$$x^t - tx \leq 1 - t \quad \text{für alle } x \in [0, 1] \text{ und } t \in (0, 1).$$

Setzen wir in dieser Ungleichung $x := a/b$ mit $0 < a \leq b \leq 1$ und $t := 1/p$, so folgt

$$\big(\frac{a}{b}\big)^{1/p} - \frac{1}{p} \frac{a}{b} \leq 1 - \frac{1}{p} = \frac{1}{q}$$

bzw. nach Multiplikation mit b

$$a^{1/p} b^{1-1/p} \leq \frac{a}{p} + \frac{b}{q}, \quad \text{also} \quad a^{1/p} b^{1/q} \leq \frac{a}{p} + \frac{b}{q}.$$

Für jedes $k \in \{1, \ldots, n\}$ wenden wir diese Ungleichung an mit

$$a := |a_k|^p / \sum_{j=1}^{n} |a_j|^p \quad \text{und} \quad b := |b_k|^q / \sum_{j=1}^{n} |b_j|^q$$

und erhalten nach Aufsummieren über k von 1 bis n

$$\sum_{k=1}^{n} |a_k b_k| / \Big((\sum_{j=1}^{n} |a_j|^p)^{1/p} (\sum_{j=1}^{n} |b_j|^q)^{1/q} \Big) \le \frac{1}{p} + \frac{1}{q} = 1.$$

Multiplikation mit dem Nenner ergibt die Behauptung. ∎

Lemma 1.5 (Minkowski-Ungleichung) *Für $p \ge 1$ und beliebige Vektoren $(a_1, \dots, a_n)$, $(b_1, \dots, b_n) \in \mathbb{R}^n$ (bzw. $\in \mathbb{C}^n$) gilt:*

$$\big(\sum_{j=1}^{n} |a_j + b_j|^p\big)^{1/p} \le \big(\sum_{j=1}^{n} |a_j|^p\big)^{1/p} + \big(\sum_{j=1}^{n} |b_j|^p\big)^{1/p}.$$

Beweis Für $p = 1$ ist die Aussage klar. Für $p > 1$ haben wir

$$\begin{aligned}
\sum_{j=1}^{n} |a_j + b_j|^p &= \sum_{j=1}^{n} |a_j + b_j|\, |a_j + b_j|^{p-1} \\
&\le \sum_{j=1}^{n} |a_j|\, |a_j + b_j|^{p-1} + \sum_{j=1}^{n} |b_j|\, |a_j + b_j|^{p-1} \\
&\overset{\text{Lemma 1.4}}{\le} \big(\sum_{j=1}^{n} |a_j|^p\big)^{1/p} \big(\sum_{j=1}^{n} |a_j + b_j|^{(p-1)q}\big)^{1/q} \\
&\quad + \big(\sum_{j=1}^{n} |b_j|^p\big)^{1/p} \big(\sum_{j=1}^{n} |a_j + b_j|^{(p-1)q}\big)^{1/q} \\
&= \Big(\big(\sum_{j=1}^{n} |a_j|^p\big)^{1/p} + \big(\sum_{j=1}^{n} |b_j|^p\big)^{1/p}\Big) \big(\sum_{j=1}^{n} |a_j + b_j|^p\big)^{1/q}.
\end{aligned}$$

Im Fall $\sum_{j=1}^{n} |a_j + b_j|^p = 0$ ist die Aussage des Lemmas offenbar richtig. Wir können daher annehmen, dass diese Summe ungleich 0 ist und durch $\big(\sum_{j=1}^{n} |a_j + b_j|^p\big)^{1/q}$ dividieren. Dies ergibt unmittelbar die Behauptung. ∎

Es folgen einige Beispiele für metrische Räume, die uns durch diese Vorlesung begleiten werden.

Beispiel 1: Normierte Räume. Ein linearer Raum X über einem Körper $\mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$) heißt *normiert*, wenn jedem Element $x \in X$ eine Zahl $\|x\|$, die *Norm* von x, zugeordnet ist, so dass für alle $x, y \in X$ und $\alpha \in \mathbb{K}$ gilt

(a) $\|x\| \ge 0$ und $\|x\| = 0$ genau dann, wenn $x = 0$,

(b) $\|\alpha x\| = |\alpha|\, \|x\|$,

(c) $\|x + y\| \le \|x\| + \|y\|$. (Dreiecksungleichung)

Man nennt dann das Paar $(X, \|\cdot\|)$ einen *normierten Raum*. Wir überlegen uns, dass für jeden normierten Raum $(X, \|\cdot\|)$ durch

$$d(x, y) := \|x - y\| \tag{1.1}$$

eine Metrik auf X definiert wird. Das erste Metrikaxiom ist offenbar erfüllt:

$$d(x, y) = \|x - y\| \geq 0, \quad d(x, y) = \|x - y\| = 0 \Longleftrightarrow x - y = 0 \Longleftrightarrow x = y.$$

Die Symmetrie folgt aus

$$d(x, y) = \|x - y\| = \| - (y - x)\| \overset{(b)}{=} \|y - x\| = d(y, x),$$

und die Dreiecksungleichung aus

$$d(x, y) = \|x - y\| = \|(x - z) + (z - y)\| \overset{(c)}{\leq} \|x - z\| + \|z - y\| = d(x, z) + d(z, y).$$

Wir vereinbaren, normierte Räume immer als metrische Räume mit der Metrik (1.1) zu betrachten. ■

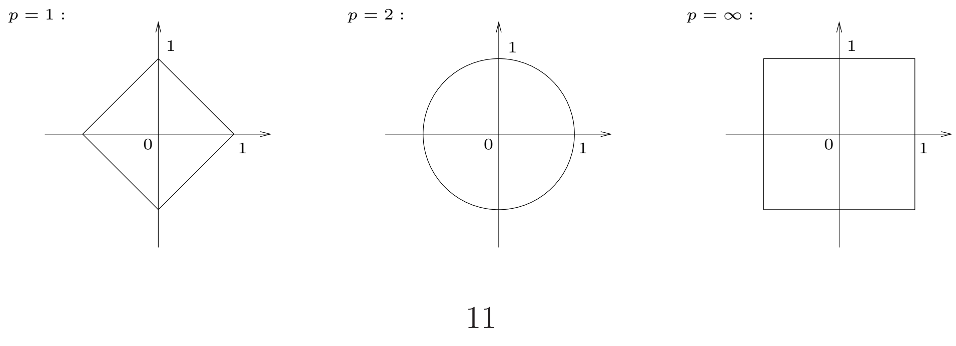
Beispiel 2: Die Räume $\mathbb{R}^n$ und $\mathbb{C}^n$. Für jedes $p \geq 1$ definieren wir die *p-Norm* eines Vektors $(a_1, \ldots, a_n) \in \mathbb{R}^n$ (oder $\in \mathbb{C}^n$) durch

$$\|(a_1, \ldots, a_n)\|_p := \big(\sum_{j=1}^{n} |a_j|^p\big)^{1/p}$$

und seine *Maximumnorm* durch

$$\|(a_1, \ldots, a_n)\|_\infty := \max_{j=1,\ldots,n} |a_j|.$$

Jede dieser Normen ist tatächlich eine Norm im Sinne von Beispiel 1. Dies ist offensichtlich mit Ausnahme der Dreiecksungleichung im Fall $p \in (1, \infty)$. Die Dreiecksungleichung in diesem Fall ist aber nichts anderes als die Minkowski-Ungleichung, die wir in Lemma 1.5 bewiesen haben. Für jedes $p \in [1, \infty]$ werden also durch $d_p(x, y) := \|x-y\|_p$ Abstände auf $\mathbb{R}^n$ bzw. $\mathbb{C}^n$ erklärt, die diese Menge zu metrischen Räumen machen. Für $p = 2$ erhält man gerade den Euklidschen Abstand. Die Einheitskreise $U_1(0)$ in $\mathbb{R}^2$ mit dem Abstand d_p haben die folgende Gestalt:



Beispiel 3: Räume von Zahlenfolgen. Die folgenden Definitionen beziehen sich auf Folgen reeller Zahlen. Für Folgen komplexer Zahlen kann man analoge Räume definieren.
Man kann die Menge $s := \{(x_n) : x_n \in \mathbb{R}\}$ aller Folgen reeller Zahlen zu einem metrischen Raum machen, indem man den Abstand zweier Folgen $x = (x_n)$, $y = (y_n)$ wie folgt erklärt:

$$d_s(x,y) := \sum_{j=1}^{\infty} \frac{1}{2^j} \frac{|x_j - y_j|}{1 + |x_j - y_j|}.$$

Dieser Abstand wird nicht durch eine Norm definiert (Beweise $\nearrow$ Übung).
Für beliebige Folgen $(x_n), (y_n) \in s$ und Zahlen $\alpha \in \mathbb{R}$ erklären wir Addition und skalare Multiplikation durch

$$(x_n) + (y_n) := (x_n + y_n) \quad \text{sowie} \quad \alpha(x_n) := (\alpha x_n).$$

Mit diesen Operationen wird s zu einem linearen Raum.
Mit l^∞, c bzw. c_0 bezeichnen wir die Menge aller beschränkten (d.h. $\sup |x_j| < \infty$), konvergenten bzw. gegen 0 konvergenten Zahlenfolgen aus s. Mit der Norm

$$\|(x_j)_{j=1}^\infty\|_\infty := \sup_j |x_j|$$

und den oben eingeführten Operationen wird jeder dieser Räume zu einem normierten Raum. Für $p \in [1, \infty)$ sei l^p die Menge aller Folgen $(x_j)_{j=1}^\infty$ mit der Eigenschaft, dass

$$\|(x_j)\|_p := \big(\sum_{j=1}^{\infty} |x_j|^p\big)^{1/p} < \infty. \tag{1.2}$$

Wir überlegen uns, dass $l^p\,(1 \leq p < \infty)$ tatsächlich ein linearer Raum ist und dass durch (1.2) eine Norm auf l^p definiert wird. Ist $(x_j), (y_j) \in l^p$, so gilt für jedes fixierte n die Minkowski-Ungleichung

$$\big(\sum_{j=1}^{n} |x_j + y_j|^p\big)^{1/p} \leq \big(\sum_{j=1}^{n} |x_j|^p\big)^{1/p} + \big(\sum_{j=1}^{n} |y_j|^p\big)^{1/p}. \tag{1.3}$$

Für $n \to \infty$ konvergiert die rechte Seite von (1.3) gegen $\|(x_j)\|_p + \|(y_j)\|_p < \infty$. Da jede monoton wachsende und nach oben beschränkte Folge konvergent ist, konvergiert auch die linke Seite von (1.3). Also ist $(x_j) + (y_j) = (x_j + y_j) \in l^p$, und die Dreiecksungleichung für $\|\cdot\|_p$ gilt. Die übrigen Normaxiome sind offenbar erfüllt. ■

Beispiel 4: Räume stetiger Funktionen. Wir betrachten in diesem Beispiel nur reellwertige Funktionen. Für komplexwertige Funktionen oder Funktionen mit Werten in $\mathbb{R}^n$ oder $\mathbb{C}^n$ kann man analoge Räume einführen.

Sei X eine beschränkte und abgeschlossene Menge im $\mathbb{R}^n$ und $C(X)$ die Menge aller stetigen Funktionen $f : X \to \mathbb{R}$. Mit den Operationen

$$(f+g)(x) := f(x) + g(x) \quad \text{und} \quad (\alpha f)(x) := \alpha f(x)$$

mit $\alpha \in \mathbb{R}$ wird $C(X)$ zu einem linearen Raum. Da jede stetige Funktion auf X ihr Maximum annimmt (Satz von Weierstraß), ist

$$\|f\|_\infty := \max_{x \in X} |f(x)| < \infty.$$

Die so definierte *Maximumnorm* ist eine Norm auf $C(X)$. Weiter: Sei X die Abschließung einer offenen beschränkten Menge $X_0 \subseteq \mathbb{R}^n$, und X sei Jordan-meßbar, d.h. die Funktion

$$\chi_X(x) := \begin{cases} 1 & \text{falls } x \in X \\ 0 & \text{falls } x \notin X \end{cases}$$

sei Riemann-integrierbar auf jedem Quader, der X umfaßt. Dann ist jede Funktion $f \in C(X)$ auf X Riemann-integrierbar, und folglich existiert für jedes $p \geq 1$ das Integral

$$\|f\|_p := \Big(\int_X |f(x)|^p \, dx \Big)^{1/p}$$

und ist endlich. Wir überlegen uns, dass $\|\cdot\|_p$ für jedes $p \geq 1$ ebenfalls eine Norm auf $C(X)$ ist.
Offenbar ist $\|f\|_p \geq 0$ und $\|\alpha f\|_p = |\alpha| \, \|f\|_p$ für jede Funktion $f \in C(X)$ und jede Zahl $\alpha \in \mathbb{R}$. Sei $\|f\|_p = 0$ für ein $f \in C(X)$. Angenommen, f ist nicht identisch 0. Dann gibt es einen Punkt $x_0 \in X_0$ mit $|f(x_0)| =: A > 0$. Wegen der Stetigkeit von f ist dann sogar $|f(x)| > A/2$ für alle x aus einer Umgebung $U_r(x_0) \subseteq X$. Dann ist aber

$$\Big(\int_X |f(x)|^p \, dx \Big)^{1/p} \geq \Big(\int_{\overline{U_r(x_0)}} |f(x)|^p \, dx \Big)^{1/p} \geq \frac{A}{2} \big(\text{Volumen von } U_r(x_0)\big)^{1/p} > 0$$

im Widerspruch zu $\|f\|_p = 0$. Die Dreiecksungleichung

$$\Big(\int_X |f(x) + g(x)|^p \, dx \Big)^{1/p} \leq \Big(\int_X |f(x)|^p \, dx \Big)^{1/p} + \Big(\int_X |g(x)|^p \, dx \Big)^{1/p}$$

für Funktionen $f, g \in C(X)$ kann man beweisen, indem man die Minkowski-Ungleichung für Riemannsummen aufschreibt und anschließend zur Grenze übergeht, oder indem man einfach die Beweise von Lemma 1.4, 1.5 entsprechend modifiziert ($\nearrow$ Übung). Es sei noch vermerkt, dass sich auch die Hölderungleichung auf Funktionen übertragen läßt. Für alle stetigen Funktionen f, g auf X und alle positiven reellen Zahlen p, q mit $\frac{1}{p} + \frac{1}{q} = 1$ ist

$$\Big| \int_X f(x) g(x) \, dx \Big| \leq \left(\int_X |f(x)|^p dx \right)^{1/p} \left(\int_X |g(x)|^q dx \right)^{1/q}.$$

■

1.2 Vervollständigung metrischer Räume

In diesem Abschnitt betrachten wir metrische Räume, in denen das Cauchykriterium für die Konvergenz einer Folge gilt.

Definition 1.6 *Eine Folge* (x_n) *in einem metrischen Raum* (M, d) konvergiert gegen $x \in M$, *wenn für jedes* $\varepsilon > 0$ *ein* N *existiert, so dass* $x_n \in U_\varepsilon(x)$ *(d.h.* $d(x, x_n) < \varepsilon$*) für alle* $n \geq N$.

Jede Folge kann gegen höchstens einen Punkt konvergieren. Angenommen, es gäbe eine Folge (x_n), welche gegen x^* und x^{**} konvergiert. Sei $r := d(x^*, x^{**}) > 0$. Nach der Definition der Konvergenz gibt es Zahlen N^* und N^{**} so, dass $d(x_n, x^*) < r/2$ für alle $n \geq N^*$ und $d(x_n, x^{**}) < r/2$ für alle $n \geq N^{**}$. Für alle $n \geq \max(N^*, N^{**})$ ist nach der Dreiecksungleichung

$$d(x^*, x^{**}) \leq d(x^*, x_n) + d(x_n, x^{**}) < r/2 + r/2 = d(x^*, x^{**}),$$

ein Widerspruch. Wenn eine Folge (x_n) gegen x konvergiert, so ist x also eindeutig bestimmt. Der Punkt x heißt *der Grenzwert* der Folge und wird mit $\lim_{x\to\infty} x_n$ bezeichnet. Es ist wünschenswert, ein Kriterium zu besitzen, womit man über die Konvergenz einer Folge entscheiden kann, ohne den Grenzwert kennen zu müssen.

Definition 1.7 (a) *Eine Folge* (x_n) *in einem metrischen Raum* (M, d) *heißt* Cauchy- oder Fundamentalfolge, *wenn zu jedem* $\varepsilon > 0$ *ein* N *existiert, so dass* $d(x_n, x_m) < \varepsilon$ *für alle* $n \geq N$.

(b) *Ein metrischer Raum heißt* vollständig, *wenn in ihm jede Cauchyfolge konvergiert.*

(c) *Ein normierter Raum, der bezüglich der durch die Norm induzierten Metrik vollständig ist, heißt* Banachraum.

Aus der Definition folgt leicht, dass jede Cauchyfolge beschränkt ist. Es ist auch klar, dass jede konvergente Folge eine Cauchyfolge ist: Ist nämlich $(x_n) \subseteq M$ eine Folge mit Grenzwert $x \in M$, so gibt es für jedes $\varepsilon > 0$ ein N so, dass $d(x, x_n) < \varepsilon/2$ für alle $n \geq N$. Für alle $m, n \geq N$ ist dann aber auch

$$d(x_n, x_m) \leq d(x, x_n) + d(x_m, x) < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Die umgekehrte Aussage gilt jedoch im allgemeinen nicht, was die Definitionen (b) und (c) rechtfertigt. Beispielsweise ist die Folge

$$1 \quad 1.4 \quad 1.41 \quad 1.414 \quad 1.4142 \quad \ldots$$

der abgeschnittenen Dezimalbruchentwicklung von $\sqrt{2}$ eine Cauchyfolge in der Menge $\mathbb{Q}$ der rationalen Zahlen. Wegen $\sqrt{2} \notin \mathbb{Q}$ konvergiert diese Folge jedoch nicht in $\mathbb{Q}$.

Beispiel 1: Vollständigkeit von $\mathbb{R}^n$. Wir wissen bereits aus der Analysisvorlesung, dass jeder der Räume $\mathbb{R}^n$ und $\mathbb{C}^n$ bezüglich jeder der Normen $\|\cdot\|_p$ mit $1 \le p \le \infty$ vollständig ist (Cauchysches Konvergenzkriterium). Wir wollen noch einen Schritt weiter gehen und führen dazu einen neuen Begriff ein:

Definition 1.8 *Zwei Normen $\|\cdot\|_*$ und $\|\cdot\|_{**}$ auf einem linearen Raum X heißen* äquivalent*, wenn es ein $C > 0$ gibt so, dass*

$$\|x\|_* \le C\|x\|_{**} \quad \textit{und} \quad \|x\|_{**} \le C\|x\|_* \quad \textit{für alle } x \in X.$$

In diesem Fall schreiben wir auch $\|\cdot\|_* \sim \|\cdot\|_{**}$.

Sind zwei Normen $\|\cdot\|_*$ und $\|\cdot\|_{**}$ auf X äquivalent, so ist jede Folge $(x_n) \subseteq X$, die bezüglich der Norm $\|\cdot\|_*$ konvergiert oder eine Cauchyfolge ist, auch bezüglich der Norm $\|\cdot\|_{**}$ konvergent bzw. eine Cauchyfolge. Äquivalente Normen erzeugen also den gleichen Konvergenzbegriff, und ist der lineare Raum X bezüglich $\|\cdot\|_*$ ein Banachraum, so auch bezüglich jeder zu $\|\cdot\|_*$ äquivalenten Norm. Die Äquivalenz $\sim$ von Normen auf X ist eine *Äquivalenzrelation*, d.h. man hat

(a) *Reflexivität* Jede Norm ist zu sich selbst äquivalent: $\|\cdot\| \sim \|\cdot\|$.

(b) *Symmetrie* $\|\cdot\|_* \sim \|\cdot\|_{**} \Rightarrow \|\cdot\|_{**} \sim \|\cdot\|_*$.

(c) *Transitivität* $\|\cdot\|_* \sim \|\cdot\|_{**}, \quad \|\cdot\|_{**} \sim \|\cdot\|_{***} \Rightarrow \|\cdot\|_* \sim \|\cdot\|_{***}$.

Die ersten beiden Eigenschaften sind klar. Für die dritte seien C_1, C_2 Konstanten mit

$$\|x\|_* \le C_1\|x\|_{**}, \quad \|x\|_{**} \le C_1\|x\|_*, \quad \|x\|_{**} \le C_2\|x\|_{***}, \quad \|x\|_{***} \le C_2\|x\|_{**}$$

für alle $x \in X$. Dann ist sicher auch

$$\|x\|_* \le C_1\|x\|_{**} \le C_1C_2\|x\|_{***} \quad \text{und} \quad \|x\|_{***} \le C_2\|x\|_{**} \le C_1C_2\|x\|_*$$

für alle $x \in X$, d.h. $\|\cdot\|_* \sim \|\cdot\|_{***}$.

Satz 1.9 *Zwei beliebige Normen auf $\mathbb{R}^n$ sind äquivalent.*

Beweis Wegen der Symmetrie und der Transitivität genügt es zu zeigen, dass jede Norm $\|\cdot\|$ auf $\mathbb{R}^n$ zur Norm $\|\cdot\|_1$ äquivalent ist. Ist $e_1, \ldots, e_n$ eine Basis in $\mathbb{R}^n$ und $x = \sum_{j=1}^n x_j e_j$ ein beliebiger Vektor aus $\mathbb{R}^n$, so gilt

$$\|x\| = \|\sum_{j=1}^n x_j e_j\| \le \sum_{j=1}^n |x_j| \, \|e_j\| \le \max_{1\le j\le n} \|e_j\| \cdot \sum_{j=1}^n |x_j| = C_1\|x\|_1$$

mit einer (von n abhängigen) Konstanten $C_1 = \max_j \|e_j\|$.

Für die umgekehrte Abschätzung vermerken wir zunächst, dass wegen

$$\big|\, \|x\| - \|y\| \,\big| \leq \|x - y\| \leq C_1 \|x - y\|_1$$

die Abbildung $x \mapsto \|x\|$ in der durch $\|\cdot\|_1$ bestimmten Metrik stetig ist. Außerdem ist die Menge $S_1 := \{x \in \mathbb{R}^n : \|x\|_1 = 1\}$ bezüglich dieser Metrik beschränkt und abgeschlossen. Nach dem Satz von Weierstraß nimmt die Funktion

$$f : S_1 \to [0, \infty), \quad x \mapsto \|x\|$$

auf S_1 ihr Minimum an. Dieses Minimum ist nicht 0, denn sonst würde es ja einen Punkt $x_0 \in S_1$ geben mit $\|x_0\| = 0$, was unmöglich ist. Das Minimum ist also positiv; wir bezeichnen es mit $1/C_2$. Dann ist

$$\frac{1}{C_2} \leq \|x\| \quad \text{bzw.} \quad 1 \leq C_2 \|x\| \quad \text{für alle } x \text{ mit } \|x\|_1 = 1.$$

Ist $y \in \mathbb{R}^n \backslash \{0\}$ beliebig, so ist $\Big\| \frac{y}{\|y\|_1} \Big\|_1 = 1$ und daher

$$1 \leq C_2 \Big\| \frac{y}{\|y\|_1} \Big\| \quad \text{bzw.} \quad \|y\|_1 \leq C_2 \|y\| \quad \text{für alle } y \in \mathbb{R}^n \backslash \{0\}.$$

Dies gilt natürlich auch für $y = 0$. Für $C := \max\{C_1, C_2\}$ gilt schließlich

$$\|x\| \leq C \|x\|_1, \quad \|x\|_1 \leq C \|x\|$$

für alle $x \in \mathbb{R}^n$, d.h. $\|\cdot\| \sim \|\cdot\|_1$. ■

Ein analoger Beweis läßt sich für $\mathbb{C}^n$ führen. Da außerdem jeder n-dimensionale Vektorraum über einem Körper $\mathbb{K}$ ($\mathbb{K} = \mathbb{R}, \mathbb{C}$) zu $\mathbb{K}^n$ isomorph ist, gilt die Aussage von Satz 1.9 für jeden *endlichdimensionalen* Vektorraum. Da die Konstante C von n abhängt (genauer: mit wachsendem n größer werden kann), ist zu erwarten, dass es auf unendlich-dimensionalen Räumen zueinander nicht äquivalente Normen und damit unterschiedliche Konvergenzbegriffe gibt.

Beispiel 2: Vollständigkeit der l^p-Räume.

Satz 1.10 *Für $p \in [1, \infty]$ ist $(l^p, \| \cdot \|_p)$ ein Banachraum.*

Beweis Wir führen den Beweis für $p < \infty$. Den Fall $p = \infty$ betrachten wir in der Übung. Für jedes $n \geq 1$ sei $x^{(n)} := (x_j^{(n)})_{j \geq 1}$ ein Element aus l^p, und $(x^{(n)})_{n \geq 1}$ sei eine Cauchyfolge in l^p. Es gilt also

$$\forall \varepsilon > 0 \; \exists \; N \quad \forall n, m \geq N : \quad \|x^{(n)} - x^{(m)}\|_p = \Big(\sum_{j=1}^{\infty} |x_j^{(n)} - x_j^{(m)}|^p \Big)^{1/p} < \varepsilon. \quad (1.4)$$

Dann gilt aber erst recht für jedes fixierte j

$$\forall \varepsilon > 0 \; \exists \; N \quad \forall n, m \geq N : \quad |x_j^{(n)} - x_j^{(m)}| < \varepsilon,$$

d.h. für jedes feste j ist $(x_j^{(n)})_{n\geq 1}$ eine Cauchyfolge in $\mathbb{R}$. Da $\mathbb{R}$ vollständig ist (Cauchysches Konvergenzkriterium), findet man ein $x_j \in \mathbb{R}$ so, dass $(x_j^{(n)})_{n\geq 1}$ gegen x_j konvergiert. Wir zeigen, dass das Element $x = (x_j)_{j\geq 1}$ zu l^p gehört und dass die Folge $(x^{(n)})_{n\geq 1}$ gegen dieses Element konvergiert.
Sind ε und N wie in (1.4), so gilt für jedes $K \geq 1$

$$\Big(\sum_{j=1}^{K} |x_j^{(n)} - x_j^{(m)}|^p\Big)^{1/p} < \varepsilon \quad \text{für } m, n \geq N.$$

Lassen wir in dieser endlichen Summe $m \to \infty$ streben, folgt

$$\Big(\sum_{j=1}^{K} |x_j^{(n)} - x_j|^p\Big)^{1/p} \leq \varepsilon \quad \text{für alle } n \geq N.$$

Da dies für jedes K gilt, erhalten wir

$$\forall \varepsilon > 0 \; \exists \; N : \quad \|x^{(n)} - x\|_p = \Big(\sum_{j=1}^{\infty} |x_j^{(n)} - x_j|^p\Big)^{1/p} \leq \varepsilon \quad \forall n \geq N.$$

Dies zeigt, dass $\|x^{(n)} - x\|_p \to 0$ für $n \to \infty$ (also Konvergenz) und außerdem, dass $x^{(n)} - x$ für große n zu l^p gehört. Da $x^{(n)}$ in l^p liegt und l^p ein linearer Raum ist, gehört auch $x = (x - x^{(n)}) + x^{(n)}$ zu l^p. ■

Aus der Vollständigkeit von $(l^\infty, \|\cdot\|_\infty)$ folgt leicht die von $(c, \|\cdot\|_\infty)$ und $(c_0, \|\cdot\|_\infty)$ (↗ Übung). ■

Beispiel 3: Vollständigkeit von Räumen stetiger Funktionen.

Satz 1.11 *$X \subseteq \mathbb{R}^n$ sei beschränkt und abgeschlossen. Dann ist $\big(C(X), \|\cdot\|_\infty\big)$ ein Banachraum.*

Beweis Der Beweis folgt dem gleichen Schema wie der Beweis des vorigen Satzes. Sei $(f_n)_{n\geq 1}$ eine Cauchyfolge in $C(X)$, d.h.

$$\forall \varepsilon > 0 \; \exists \; N \quad \forall n, m \geq N : \quad \|f_n - f_m\|_\infty = \max_{x \in X} |f_n(x) - f_m(x)| < \varepsilon. \tag{1.5}$$

Dann ist für jedes fixierte x die Folge $\big(f_n(x)\big)_{n\geq 1}$ eine Cauchyfolge in $\mathbb{R}$ und folglich konvergent. Sei $f(x) := \lim f_n(x)$. Wir zeigen, dass die so definierte Funktion f zu $C(X)$ gehört und dass sie Grenzwert der Folge $(f_n)_{n\geq 1}$ ist.

Sind ε und N wie in (1.5), so ist für jedes $x \in X$

$$|f_n(x) - f_m(x)| < \varepsilon \quad \forall n, m \geq N.$$

Grenzübergang $m \to \infty$ liefert

$$|f_n(x) - f(x)| \leq \varepsilon \quad \forall n \geq N,$$

und da $x \in X$ beliebig gewählt war und N nicht von x abhängt, folgt

$$\forall \varepsilon > 0 \;\exists\, N: \;\forall n \geq N \quad \sup_{x \in X} |f(x) - f_n(x)| \leq \varepsilon. \tag{1.6}$$

Die *stetigen* Funktionen f_n konvergieren auf X also *gleichmäßig* gegen f. Aus der Analysisvorlesung wissen wir, dass dann f selbst eine stetige Funktion auf X ist. Wir können also das sup in (1.6) durch ein max ersetzen. ■

Man beachte aber, dass die normierten Räume $\big(C(X), \|\cdot\|_p\big)$ für $1 \leq p < \infty$ i.Allg. *nicht vollständig* sind. Sei z.B. $p = 2$, $X = [0, 2]$, und für jedes $n \geq 1$ sei

$$f_n(x) = \begin{cases} x^n & \text{für } x \in [0,1] \\ 1 & \text{für } x \in (1,2]. \end{cases}$$

Dann ist (f_n) eine Cauchyfolge in $\big(C[0,2], \|\cdot\|_2\big)$:

$$\begin{aligned} \|f_n - f_m\|_2^2 &= \int_0^2 |f_n(x) - f_m(x)|^2\, dx = \int_0^1 (x^n - x^m)^2\, dx \\ &= \int_0^1 (x^{2n} - 2x^{m+n} + x^{2m})\, dx = \frac{1}{2n+1} - \frac{1}{m+n+1} + \frac{1}{2m+1}. \end{aligned}$$

Sei nun $\varepsilon > 0$ vorgegeben. Für $m, n \geq \frac{1}{\varepsilon} =: N$ ist

$$\frac{1}{2n+1} - \frac{1}{m+n+1} + \frac{1}{2m+1} < \frac{1}{2n} + \frac{1}{2m} \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Also ist (f_n) eine Cauchyfolge. Es existiert jedoch *keine* stetige Funktion auf $[0, 2]$, gegen die (f_n) im Sinne der Norm $\|\cdot\|_2$ konvergiert. Angenommen, es gäbe eine solche Funktion $f \in C[0, 2]$. Dann wäre einerseits für jedes $n \geq 1$

$$\int_1^2 |f(x) - 1|^2\, dx = \int_1^2 |f(x) - f_n(x)|^2\, dx \leq \int_0^2 |f(x) - f_n(x)|^2\, dx = \|f - f_n\|_2^2,$$

und aus $\|f_n - f\|_2 \to 0$ würde folgen: $\int_1^2 |f(x) - 1|^2\, dx = 0$, d.h. $f(x) = 1$ auf $[1, 2]$.
Andererseits wäre nach der Dreiecksungleichung für die $\|\cdot\|_2$-Norm auf $[0, 1]$:

$$\Big|\Big(\int_0^1 |f(x)|^2\, dx\Big)^{1/2} - \Big(\int_0^1 |f_n(x)|^2\, dx\Big)^{1/2}\Big| \leq \Big(\int_0^1 |f(x) - f_n(x)|^2\, dx\Big)^{1/2} \leq \|f - f_n\|_2.$$

Da $\|f - f_n\|_2 \to 0$ nach Annahme, und da

$$\int_0^1 |f_n(x)|^2 \, dx = \int_0^1 x^{2n} \, dx = \frac{1}{2n+1} \to 0 \text{ für } n \to \infty,$$

folgt hieraus, dass

$$\int_0^1 |f(x)|^2 \, dx = 0, \quad \text{d.h. } f(x) = 0 \text{ auf } [0,1].$$

Die Funktion f kann also nicht stetig sein, ein Widerspruch. Sofern nichts anderes gesagt wird, werden wir daher $C(X)$ *stets* mit der Maximumnorm versehen. ∎

Wir zeigen nun, dass man jeden nichtvollständigen metrischen Raum (M, d) durch Hinzunahme neuer Elemente vervollständigen kann. Dazu definieren wir:

Definition 1.12 *Seien (M_1, d_1) und (M_2, d_2) metrische Räume. Eine Abbildung $f : M_1 \to M_2$ heißt* Isometrie*, wenn $d_2\big(f(x), f(y)\big) = d_1(x, y)$ für alle $x, y \in M_1$. Die Räume (M_1, d_1) und (M_2, d_2) heißen zueinander* isometrisch*, wenn es eine bijektive Isometrie zwischen M_1 und M_2 gibt.*

Die Isometrie metrischer Räume ist eine Äquivalenzrelation. Zueinander isometrische Räume werden i.Allg. miteinander identifiziert.

Definition 1.13 *Eine Teilmenge M' eines metrischen Raumes heißt* dicht *in M, wenn* $\operatorname{clos} M' = M$.

Satz 1.14 *Zu jedem metrischen Raum (M, d) gibt es einen vollständigen metrischen Raum $(\tilde{M}, \tilde{d})$, der einen zu (M, d) isometrischen und in $(\tilde{M}, \tilde{d})$ dichten Teilraum $(\hat{M}, \tilde{d})$ enthält.*

Der Raum $(\tilde{M}, \tilde{d})$ ist bis auf Isometrie eindeutig durch (M, d) bestimmt und heißt die *Vervollständigung* von (M, d).

Beweis Sei M_{CF} die Menge aller Cauchyfolgen in (M, d). Wir nennen zwei Folgen (x_j), $(y_j) \in M_{CF}$ äquivalent und schreiben dafür $(x_j) \sim (y_j)$, wenn

$$\lim_{j \to \infty} d(x_j, y_j) = 0.$$

Die Relation $\sim$ ist eine Äquivalenzrelation auf M_{CF} ($\nearrow$ Übung). Sei $\tilde{M}$ die Menge allerÄquivalenzklassen bezüglich $\sim$, und für jede Folge (x_j) bezeichne $(x_j)^\sim$ die Äquivalenzklasse, in der (x_j) liegt. Für je zwei Cauchyfolgen $(x_j), (y_j)$ folgt aus

$$|d(x_j, y_j) - d(x_k, y_k)| \le d(x_j, x_k) + d(y_j, y_k), \qquad (\nearrow \text{ Übung})$$

dass $\big(d(x_j, y_j)\big)_{j \ge 1}$ eine Cauchyfolge in $\mathbb{R}$ und damit eine konvergente Folge ist.

Der Grenzwert $\lim_{j\to\infty} d(x_j, y_j)$ hängt nicht von den Folgen $(x_j), (y_j)$, sondern nur von deren Äquivalenzklassen $(x_j)^\sim, (y_j)^\sim$ ab. Liegen nämlich $(\hat{x}_j) \in (x_j)^\sim$ und $(\hat{y}_j) \in (y_j)^\sim$ in den gleichen Äquivalenzklassen, so folgt aus

$$|d(x_j, y_j) - d(\hat{x}_j, \hat{y}_j)| \le d(x_j, \hat{x}_j) + d(y_j, \hat{y}_j) \to 0,$$

dass

$$\lim d(x_j, y_j) = \lim d(\hat{x}_j, \hat{y}_j).$$

Wir bezeichnen den Grenzwert $\lim d(x_j, y_j)$ mit $\tilde{d}\big((x_j)^\sim, (y_j)^\sim\big)$. Hierdurch wird eine Metrik $\tilde{d}$ auf $\tilde{M}$ erklärt ($\nearrow$ Übung).

Mit $\hat{M}$ bezeichnen wir die Menge aller Äquivalenzklassen $(x)^\sim$ von konstanten Folgen (x), $x \in M$. Die Abbildung

$$f : M \to \hat{M}, \quad x \mapsto (x)^\sim$$

ist nach Definition surjektiv, und aus

$$\tilde{d}\big(f(x), f(y)\big) = \tilde{d}\big((x)^\sim, (y)^\sim\big) = \lim_{j\to\infty} d(x, y) = d(x, y)$$

folgt, dass f eine Isometrie ist. Dann ist offenbar f eine Bijektion von (M, d) auf $(\hat{M}, \tilde{d})$, d.h. diese Räume sind isometrisch.

Die Menge $\hat{M}$ liegt dicht in $\tilde{M}$. Für jede Folge $(x_n)_{n\ge1} \in M_{CF}$ und jedes $\varepsilon > 0$ gibt es nämlich ein N so, dass $d(x_n, x_N) < \varepsilon$ für alle $n \ge N$. Die Äquivalenzklasse $(x_N)^\sim_{n\ge1}$ der konstanten Folge $(x_N)_{n\ge1}$ liegt in $\hat{M}$, und es gilt

$$\tilde{d}\big((x_n)^\sim, (x_N)^\sim\big) = \lim_{n\to\infty} d(x_n, x_N) \le \varepsilon.$$

Abschließend zeigen wir, dass $(\tilde{M}, \tilde{d})$ ein vollständiger metrischer Raum ist. Sei $\big((x_n^{(k)})^\sim_{n\ge1}\big)_{k\ge1}$ eine Cauchyfolge in $\tilde{M}$. Da $\hat{M}$ dicht in $\tilde{M}$ liegt, gibt es für jedes k ein Element $y^{(k)} \in M$ so, dass

$$\tilde{d}\big((x_n^{(k)})^\sim_{n\ge1},\, (y^{(k)})^\sim_{n\ge1}\big) < \frac{1}{k}\,.$$

Wegen

$$\begin{aligned}\tilde{d}\big((y^{(k)})^\sim_{n\ge1},\, (y^{(m)})^\sim_{n\ge1}\big) &\le \tilde{d}\big((y^{(k)})^\sim_{n\ge1}, (x_n^{(k)})^\sim_{n\ge1}\big) + \tilde{d}\big((x_n^{(k)})^\sim,\, (x_n^{(m)})^\sim\big) \\ &\quad + \tilde{d}\big((x_n^{(m)})^\sim,\, (y^{(m)})^\sim\big) \\ &\le 1/k + \tilde{d}\big((x_n^{(k)})^\sim,\, (x_n^{(m)})^\sim\big) + 1/m\end{aligned}$$

ist $\big((y^{(k)})^\sim_{n\ge1}\big)_{k\ge1}$ eine Fundamentalfolge in $\tilde{M}$. Wie wir oben gesehen haben, gilt aber

$$\tilde{d}\big((y^{(k)})^\sim,\, (y^{(m)})^\sim\big) = d(y^{(k)}, y^{(m)}).$$

Folglich ist $(y^{(k)})_{k\geq 1}$ eine Fundamentalfolge im Ausgangsraum M, d.h. $(y^{(k)})^{\sim}_{k\geq 1}$ ist ein Element von $\tilde{M}$. Dieses Element ist gerade der Grenzwert der Folge $\big((x_n^{(k)})^{\sim}_{n\geq 1}\big)_{k\geq 1}$, denn

$$\begin{aligned} \tilde{d}\big((x_n^{(k)})^{\sim}_{n\geq 1},\, (y^{(k)})^{\sim}_{k\geq 1}\big) &\leq \tilde{d}\big((x_n^{(k)})^{\sim}_{n\geq 1},\, (y^{(k)})^{\sim}_{n\geq 1}\big) + \tilde{d}\big((y^{(k)})^{\sim}_{n\geq 1},\, (y^{(k)})^{\sim}_{k\geq 1}\big) \\ &< \frac{1}{k} + \lim_{n\to\infty} d(y^{(k)}, y^{(n)}). \end{aligned} \tag{1.7}$$

Wählt man für gegebenes ε ein $N > \frac{2}{\varepsilon}$ so, dass $d(y^{(k)}, y^{(n)}) < \frac{\varepsilon}{2}$ für alle $k, n \geq N$, so wird die rechte Seite von (1.7) für alle $k \geq N$ kleiner als ε. ■

Einen Beweis, dass $(\tilde{M}, \tilde{d})$ im Wesentlichen eindeutig durch (M, d) bestimmt ist, findet man in Heuser, FA, S. 80.

Wenden wir diesen Vervollständigungsmechanismus auf einen (unvollständigen) normierten linearen Raum $(X, \|\cdot\|)$ über einem Körper $\mathbb{K}$ an, so erhalten wir als Vervollständigung zunächst nur einen *metrischen* Raum $(\tilde{X}, \tilde{d})$. Diesen kann man durch Einführung geeigneter Operationen und einer passenden Norm $\|\cdot\|_{\sim}$ zu einem *normierten* Raum über $\mathbb{K}$, also zu einem Banachraum machen. Genauer:

Satz 1.15 *Zu jedem (unvollständigen) normierten Raum $(X, \|\cdot\|)$ gibt es einen Banachraum $(\tilde{X}, \|\cdot\|_{\sim})$, der einen zu $(X, \|\cdot\|)$ isometrischen und in $(\tilde{X}, \|\cdot\|_{\sim})$ dichten linearen Teilraum $(\hat{X}, \|\cdot\|_{\sim})$ enthält.*

Der Raum $(\hat{X}, \|\cdot\|_{\sim})$ ist wieder bis auf Isometrie eindeutig bestimmt.

Beweis Mit (x_n) und (y_n) sind auch $(x_n + y_n)$ und (αx_n) Cauchyfolgen in X. Wir definieren daher die Addition bzw. skalare Multiplikation auf $\tilde{X}$ durch

$$(x_n)^{\sim} + (y_n)^{\sim} := (x_n + y_n)^{\sim}, \quad \alpha (x_n)^{\sim} := (\alpha x_n)^{\sim}. \tag{1.8}$$

Diese Definition hängt nicht von den gewählten Repräsentanten (x_n) der Äquivalenzklasse $(x_n)^{\sim}$ ab ($\nearrow$ Übung), und für die Äquivalenzklassen $(x)^{\sim}, (y)^{\sim} \in \hat{X}$ ist insbesondere

$$(x)^{\sim} + (y)^{\sim} = (x + y)^{\sim}, \quad \alpha (x)^{\sim} = (\alpha x)^{\sim},$$

d.h. die in (1.8) definierten Operationen stimmen auf $\hat{X}$ genau mit denen von X überein. Die Norm $\|\cdot\|_{\sim}$ auf $\tilde{X}$ erklären wir durch

$$\|(x_n)^{\sim}\|_{\sim} := \tilde{d}\big((x_n)^{\sim}, (0)^{\sim}\big). \tag{1.9}$$

Aus

$$\|(x_n)^{\sim}\|_{\sim} = \tilde{d}\big((x_n)^{\sim}, (0)^{\sim}\big) = \lim d(x_n, 0) = \lim \|x_n - 0\| = \lim \|x_n\|$$

folgt sofort

$$\|(x_n)^\sim - (y_n)^\sim\|_\sim = \lim \|x_n - y_n\| = \lim d(x_n, y_n) = \tilde{d}\big((x_n)^\sim, (y_n)^\sim\big). \tag{1.10}$$

Hieraus erhält man die Gültigkeit der Dreiecksungleichung für $\|\cdot\|_\sim$, und die übrigen Normaxiome sind ebenfalls leicht zu sehen ($\nearrow$ Übung). $\|\cdot\|_\sim$ ist also tatsächlich eine Norm auf $\tilde{X}$, die wegen (1.10) die Metrik $\tilde{d}$ erzeugt. Für jede konstante Folge (x) ist natürlich $\|(x)^\sim\|_\sim = \lim \|x\| = \|x\|$, d.h. auf $\hat{X}$ stimmt $\|\cdot\|_\sim$ mit der Ausgangsnorm überein. ■

Definition 1.16 *Sei X die Abschließung einer offenen beschränkten Menge $X_0 \subseteq \mathbb{R}^n$ und Jordanmessbar. Die Vervollständigung von $C(X)$ bezüglich der p-Norm $\|\cdot\|_p$ mit $1 \le p < \infty$ heißt* Lebesgueraum *und wird mit $L^p(X)$ bezeichnet. Die Norm in $L^p(X)$ bezeichnen wir wieder mit $\|\cdot\|_p$.*

Wie wir oben gesehen haben (Diskussion der Unvollständigkeit von $\big(C[0,2], \|\cdot\|_2)\big)$, sollte es in $L^2[0,2]$ ein Element geben, welches auf $[0,1]$ mit der Funktion 0 und auf $[1,2]$ mit der Funktion 1 übereinstimmt. Dies zeigt, dass man sich die Elemente der Räume $L^p(X)$ nicht als Funktionen vorstellen darf. Glücklicherweise kann man sie aber doch „fast“ mit Funktionen identifizieren. Genauer: in der Maßtheorie wird gezeigt, dass die Elemente von $L^p(X)$ Äquivalenzklassen von messbaren Funktionen sind, die sich nur auf einer Menge vom Lebesgue-Maß 0 (also z.B. in endlich vielen Punkten) unterscheiden. In unserem Beispiel würden wir also die Funktionen

$$f_1(x) = \begin{cases} 0 & x \in [0,1] \\ 1 & x \in (1,2] \end{cases}, \qquad f_2(x) = \begin{cases} 0 & x \in [0,1) \\ 1 & x \in [1,2] \end{cases}$$

nicht als verschieden ansehen.
Die Maßtheorie setzen wir hier nicht voraus. Wir werden jedoch der Einfachheit halber (wie üblich, aber nicht korrekt) die Elemente von $L^p(X)$ als Funktionen bezeichnen. Integrale über Funktionen aus $L^p(X)$ definieren wir als Grenzwerte. Ist z.B. $f \in L^1(X)$, so gibt es eine Folge von Funktionen $f_n \in C(X)$, die in der L^1-Norm gegen f konvergieren. Die Folge (f_n) ist dann eine Cauchyfolge in $\big(C(X), \|\cdot\|_1\big)$, und aus

$$\Big|\int_X f_n(x)\,dx - \int_X f_m(x)\,dx\Big| \le \int_X |f_n(x) - f_m(x)|\,dx = \|f_n - f_m\|_1$$

folgt, dass die Folge $\big(\int_X f_n(x)\,dx\big)$ eine Cauchyfolge und daher konvergent in $\mathbb{R}$ ist. Wir können also definieren:

$$\int_X f(x)\,dx := \lim_{n\to\infty} \int_X f_n(x)\,dx. \tag{1.11}$$

Weiter: Seien $f \in L^p(X)$, $g \in L^q(X)$ mit $p \in (1,\infty)$ und $\frac{1}{p} + \frac{1}{q} = 1$. Wir wählen Funktionen $f_n, g_n \in C(X)$, die in der L^p- bzw. L^q-Norm gegen f bzw.

g konvergieren. Dann sind (f_n) bzw. (g_n) Cauchyfolgen in $\big(C(X), \|\cdot\|_p\big)$ bzw. $\big(C(X), \|\cdot\|_q\big)$, und aus der Hölderungleichung für stetige Funktionen folgt

$$\begin{aligned} \|f_n g_n - f_m g_m\|_1 &\leq \|f_n(g_n - g_m)\|_1 + \|(f_n - f_m) g_m\|_1 \\ &\leq \|f_n\|_p \, \|g_n - g_m\|_q + \|f_n - f_m\|_p \, \|g_m\|_q. \end{aligned}$$

Da Cauchyfolgen beschränkt sind, folgt hieraus, dass $(f_n g_n)$ eine Cauchyfolge in $\big(C(X), \|\cdot\|_1\big)$ ist. Es ist naheliegend, den Grenzwert dieser Cauchyfolge mit $fg \in L^1(X)$ zu bezeichnen. Mit dieser Festlegung und mit der Integraldefinition (1.11) gelangen wir zur

Hölderungleichung Seien $f \in L^p(X)$, $g \in L^q(X)$ und $\frac{1}{p} + \frac{1}{q} = 1$ mit $p > 1$. Dann ist $fg \in L^1(X)$ und $\|fg\|_1 \leq \|f\|_p \, \|g\|_q$, d.h.

$$\int_X |f(x)g(x)| \, dx \leq \Big(\int_X |f(x)|^p \, dx \Big)^{1/p} \Big(\int_X |g(x)|^q \, dx \Big)^{1/q}.$$

Hieraus folgt wie in Abschnitt 1.1 die

Minkowski-Ungleichung oder Dreiecksungleichung Für $f, g \in L^p(X)$ ist $f + g \in L^p(X)$ und $\|f + g\|_p \leq \|f\|_p + \|g\|_p$, d.h.

$$\Big(\int_X |f(x) + g(x)|^p \, dx \Big)^{1/p} \leq \Big(\int_X |f(x)|^p \, dx \Big)^{1/p} + \Big(\int_X |g(x)|^p \, dx \Big)^{1/p}.$$

Dies gilt auch für $p = 1$.

1.3 Der Banachsche Fixpunktsatz

Wir kommen nun zu einem Problem, welches uns bereits in Beispiel 2 der Einleitung begegnete: Wann besitzt eine Abbildung A eines metrischen Raumes in sich einen *Fixpunkt*, d.h. wann existiert ein x mit $Ax = x$? Es gibt zahlreiche, z.T. recht tiefliegende Fixpunktsätze. Der folgende Banachsche Fixpunktsatz zeichnet sich durch seine Einfachheit und breite Anwendbarkeit aus.
Für jede Abbildung $A : M \to M$ eines metrischen Raumes in sich sei A^k die k-fache Hintereinanderausführung von A, und A^0 bezeichne die identische Abbildung.

Satz 1.17 (Fixpunktsatz von Banach-Weissinger) *Sei (M, d) ein vollständiger metrischer Raum und $A : M \to M$ eine Abbildung mit folgender Eigenschaft: Es gibt eine Folge $(\alpha_n) \in l^1$ von nichtnegativen Zahlen so, dass*

$$d(A^n x, A^n y) \leq \alpha_n d(x, y) \quad \textit{für alle } x, y \in M \textit{ und } n \in \mathbb{N}.$$

Dann besitzt die Gleichung $Ax = x$ genau eine Lösung $x^* \in M$. Diese kann iterativ gewonnen werden: Für beliebiges $x_0 \in M$ konvergiert die Folge $(A^n x_0)_{n\geq 1}$ gegen x^*, und es gilt die Fehlerabschätzung

$$d(x^*, A^n x_0) \leq \sum_{r=n}^{\infty} \alpha_r \, d(Ax_0, x_0). \tag{1.12}$$

Wir schreiben auch x_n statt $A^n x_0$. Iterativ heißt dann $x_{n+1} = Ax_n$, d.h. durch wiederholtes Anwenden von A nähert man sich dem Fixpunkt.

Beweis Wegen

$$d(x_{k+1}, x_k) = d(A^{k+1}x_0, A^k x_0) = d(A^k x_1, A^k x_0) \leq \alpha_k \, d(x_1, x_0)$$

ist

$$d(x_{n+k}, x_n) \leq d(x_{n+k}, x_{n+k-1}) + \ldots + d(x_{n+1}, x_n) \leq \sum_{r=n}^{n+k-1} \alpha_r \, d(x_1, x_0). \tag{1.13}$$

Da die Reihe $\sum_{r=1}^{\infty} \alpha_r$ nach Voraussetzung konvergiert, bilden die Partialsummen dieser Reihe eine Cauchyfolge. Die Abschätzung (1.13) zeigt, dass dann (x_n) eine Cauchyfolge ist, die wegen der Vollständigkeit von (M, d) konvergiert. Sei $x^* := \lim x_n$. Lassen wir in (1.13) $k \to \infty$ streben, folgt sofort die Abschätzung (1.12).
Wir zeigen, dass x^* ein Fixpunkt von A ist. Aus der Abschätzung

$$d(Ax, Ay) \leq \alpha_1 d(x, y)$$

folgt, dass A eine stetige Abbildung ist. Läßt man in $x_{n+1} = Ax_n$ n gegen Unendlich streben, folgt daher sofort $x^* = Ax^*$. Sei schließlich $x^{**} \in M$ ein weiterer Fixpunkt von A. Dann ist

$$d(x^*, x^{**}) = d(A^k x^*, A^k x^{**}) \leq \alpha_k \, d(x^*, x^{**}) \quad \text{für alle } k.$$

Grenzübergang $k \to \infty$ in dieser Ungleichung liefert $d(x^*, x^{**}) = 0$, d.h. $x^* = x^{**}$.
■

Der Fixpunktsatz von Banach wird oft für kontrahierende Abbildungen formuliert und bewiesen. Eine Abbildung $A : M \to M$ heißt *kontrahierend*, wenn es ein $q \in [0, 1)$ so gibt, dass

$$d(Ax, Ay) \leq q \, d(x, y) \quad \forall \, x, y \in M. \tag{1.14}$$

Für kontrahierende Abbildungen gilt offenbar

$$d(A^n x, A^n y) \leq q^n \, d(x, y) \quad \forall \, x, y \in M,$$

und da die geometrische Reihe $\sum_{n=0}^{\infty} q^n$ konvergiert, liegt die Folge (q^n) in l^1, und aus Satz 1.17 folgt sofort

Satz 1.18 (Banachscher Fixpunktsatz) *Sei (M, d) ein vollständiger metrischer Raum und $A : M \to M$ eine kontrahierende Abbildung. Dann besitzt die Abbildung A genau einen Fixpunkt $x^* \in M$. Für einen beliebigen Startpunkt $x_0 \in M$ konvergiert die Folge $(A^n x_0)_{n\geq 1}$ gegen x^*, und es gilt die Fehlerabschätzung*

$$d(x^*, A^n x_0) \leq \frac{q^n}{1-q}\, d(Ax_0, x_0). \tag{1.15}$$

Der Fixpunktsatz von Banach-Weissinger liefert nicht nur Existenz und Eindeutigkeit des Fixpunktes, sondern zugleich eine Konstruktionsmöglichkeit.

Beispiel 1: Lösbarkeit linearer Gleichungssysteme. Jedes System von n linearen Gleichungen mit n Unbekannten

$$\sum_{k=1}^{n} \alpha_{jk} x_k = b_j, \quad j = 1, \ldots, n$$

läßt sich (mit $a_{jk} := -\alpha_{jk}$ für $j \neq k$ und $a_{kk} := 1 - \alpha_{kk}$) in die folgende Form bringen

$$x_j = \sum_{k=1}^{n} a_{jk} x_k + b_j\,, \quad j = 1, \ldots, n. \tag{1.16}$$

Definieren wir $A : \mathbb{R}^n \to \mathbb{R}^n$ durch

$$A(x_1, \ldots, x_n) = (\sum_{k=1}^{n} a_{1k} x_k + b_1, \ldots, \sum_{k=1}^{n} a_{nk} x_k + b_n),$$

kann man (1.16) als Fixpunktproblem $Ax = x$ mit $x = (x_1, \ldots, x_n)$ schreiben. Ist $\mathbb{R}^n$ mit der Summennorm $\|\cdot\|_1$ versehen, so ist für beliebige $x = (x_1, \ldots, x_n)$ und $y = (y_1, \ldots, y_n)$

$$d(Ax, Ay) = \|Ax - Ay\|_1 = \sum_{j=1}^{n} \Big| \sum_{k=1}^{n} a_{jk}(x_k - y_k) \Big| \leq \sum_{k=1}^{n} |x_k - y_k| \sum_{j=1}^{n} |a_{jk}|,$$

d.h.

$$d(Ax, Ay) \leq \max_k \sum_{j=1}^{n} |a_{jk}|\, d(x, y).$$

Im Falle $\max_k \sum_j |a_{jk}| < 1$ ist A kontrahierend, und (1.16) besitzt eine eindeutige Lösung, welche iterativ gewonnen werden kann. Bei Verwendung anderer Normen auf $\mathbb{R}^n$ erhält man andere hinreichende Lösbarkeitskriterien. ∎

Beispiel 2: Lösbarkeit von Integralgleichungen 2. Art. Wir betrachten die Fredholmsche Integralgleichung 2. Art

$$u(x) - \int_a^b k(x,t)\, u(t)\, dt = f(x), \quad x \in [a, b] \tag{1.17}$$

und setzen voraus, dass f auf $[a,b]$ und k auf $[a,b]\times[a,b]$ stetig sind. Wir suchen Lösungen $u \in C[a,b]$. Definieren wir $A : C[a,b] \to C[a,b]$ durch

$$(Au)(x) = f(x) + \int_a^b k(x,t)\,u(t)\,dt\,, \quad x \in [a,b],$$

so können wir (1.17) als Fixpunktgleichung $Au = u$ betrachten. Da wir Vollständigkeit benötigen, versehen wir $C[a,b]$ mit der Supremumsnorm $\|\cdot\|_\infty$ und erhalten bezüglich dieser Norm die Abschätzung für $u, v \in C[a,b]$

$$\begin{aligned}\|Au - Av\|_\infty &= \max_{x\in[a,b]} \left| \int_a^b k(x,t)\big(u(t)-v(t)\big)\,dt \right| \\ &\le \max_{t\in[a,b]} |u(t)-v(t)| \max_{x\in[a,b]} \int_a^b |k(x,t)|\,dt.\end{aligned}$$

Im Falle

$$\max_{x\in[a,b]} \int_a^b |k(x,t)|\,dt < 1$$

besitzt (1.17) also eine eindeutig bestimmte Lösung. Auf ähnliche Weise zeigt man in der Vorlesung „Gewöhnliche Differentialgleichungen“ den Satz von Picard-Lindelöf für die eindeutige Lösbarkeit des AWP (0.2) aus der Einleitung. ■

1.4 Kompakte Mengen in metrischen Räumen

Eine Teilmenge eines metrischen Raumes heißt *beschränkt*, wenn sie in einer Kugel $U_r(x)$ mit endlichem Radius r liegt. Aus der Analysis 1 ist bekannt, dass jede unendliche und beschränkte Teilmenge des $\mathbb{R}^n$ einen Häufungspunkt besitzt (Satz von Bolzano/Weierstraß). In unendlich-dimensionalen Räumen gilt dies nicht mehr. Ist z.B. $e_i := (0,\ldots,0,1,0,\ldots) \in l^2$ mit der 1 an i-ter Stelle, so ist $\{e_i\}_{i\ge1}$ eine unendliche beschränkte Menge in l^2, die wegen $d(e_i,e_j) = \sqrt{2}$ für $i \ne j$ keinen Häufungspunkt besitzt. Wir definieren daher

Definition 1.19 (a) *Eine Teilmenge X eines metrischen Raumes heißt* relativ kompakt, *wenn jede Folge in X eine konvergente Teilfolge besitzt (deren Grenzwert jedoch nicht zu X gehören muß).*

(b) *X heißt* kompakt, *wenn X relativ kompakt und abgeschlossen ist.*

Kompakte Mengen X sind also dadurch ausgezeichnet, dass jede Folge in X eine konvergente Teilfolge besitzt, *deren Grenzwert wieder zu X gehört.*

Lemma 1.20 *Kompakte Mengen sind beschränkt und vollständig.*

Beweis $\nearrow$ Übung.

Kompakte Mengen sind also insbesondere beschränkt und abgeschlossen. In endlich-dimensionalen Räumen sind diese beiden Bedingungen auch hinreichend für Kompaktheit (Bolzano/Weierstraß), in unendlich-dimensionalen Räumen dagegen nicht (siehe oben). Wir benötigen daher Kompaktheitskriterien für unendlich-dimensionale Räume.

Sei X Teilmenge eines metrischen Raumes (M, d). Für $\varepsilon > 0$ heißt eine Teilmenge A von X ein *ε-Netz für X*, wenn es für jedes $x \in X$ ein $a \in A$ mit $d(x, a) < \varepsilon$ gibt. Beispielsweise ist die Menge $\mathbb{Z}$ der ganzen Zahlen ein 1-Netz für $\mathbb{R}$. Weiter heißt eine Familie (G_α) offener Mengen in M eine *Überdeckung* von X, wenn $X \subseteq \cup_\alpha G_\alpha$.

Lemma 1.21 *Der metrische Raum (X, d) besitze für jedes $\varepsilon > 0$ ein endliches ε-Netz. Dann besitzt auch jede Teilmenge $A \subseteq X$ für jedes $\varepsilon > 0$ ein endliches ε-Netz.*

Beweis Sei $\varepsilon > 0$ und $b_1, \ldots, b_n$ ein endliches $\varepsilon/2$-Netz für X. Wir betrachten die Mengen $A \cap U_{\varepsilon/2}(b_i)$, $i = 1, \ldots, n$. Diese Mengen überdecken A. Sei o.E.d.A. $A \cap U_{\varepsilon/2}(b_i) \neq \emptyset$ für $i = 1, \ldots, k$ und $= \emptyset$ für $i = k+1, \ldots, n$. Dann überdecken die Mengen $A \cap U_{\varepsilon/2}(b_i)$ mit $i = 1, \ldots, k$ offenbar ebenfalls A. Wir wählen

$$a_i \in A \cap U_{\varepsilon/2}(b_i) \quad \text{für } i = 1, \ldots, k$$

und zeigen:

$$U_{\varepsilon/2}(b_i) \subseteq U_\varepsilon(a_i).$$

Sei $a \in U_{\varepsilon/2}(b_i)$. Daraus folgt

$$d(a, a_i) \leq d(a, b_i) + d(b_i, a_i) < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Die Mengen $U_\varepsilon(a_i)$, $i = 1, \ldots, k$ (genommen im metrischen Raum $(A, d|_A)$) überdecken A. Also ist $\{a_1, \ldots, a_k\}$ ein endliches ε-Netz für A. ■

Satz 1.22 *Für einen metrischen Raum (X, d) sind folgende Aussagen äquivalent:*

(a) *X ist kompakt.*

(b) *X ist vollständig, und für jedes $\varepsilon > 0$ existiert ein* endliches *ε-Netz für X.*

(c) *Aus jeder offenen Überdeckung von X läßt sich eine endliche Überdeckung auswählen.*

Beweis (a) $\Rightarrow$ (b). Die Vollständigkeit haben wir in Lemma 1.20 bewiesen. Für die zweite Aussage von (b) nehmen wir an, es gäbe ein $\varepsilon_0 > 0$, für das kein endliches ε_0-Netz für X existiert. Wir wählen $a_1 \in X$ beliebig. Dann gibt es ein $a_2 \in X$ mit $d(a_1, a_2) \geq \varepsilon_0$ (sonst wäre ja $\{a_1\}$ ein endliches ε_0-Netz). Weiter können wir ein $a_3 \in X$ wählen so, dass $d(a_i, a_3) \geq \varepsilon_0$ für $i = 1, 2$ (sonst wäre ja $\{a_1, a_2\}$ ein endliches ε_0-Netz). Wir fahren auf diese Weise fort und konstruieren eine Folge $(a_n) \subseteq X$ mit der Eigenschaft, dass $d(a_m, a_n) \geq \varepsilon_0 > 0$ für alle

$m \neq n$. Aus dieser Folge läßt sich natürlich keine konvergente Teilfolge auswählen. Widerspruch.

(b) $\Rightarrow$ (c). Angenommen, (G_α) ist eine offene Überdeckung von X, aus der sich *keine* endliche Überdeckung auswählen läßt. Sei $\{x_1, \ldots, x_n\}$ endliches 1-Netz für X. Wenigstens eine der Kugeln $\overline{U_1(x_1)}, \ldots, \overline{U_1(x_n)}$ wird nicht durch endlich viele der G_α überdeckt (würde jede dieser n Kugeln eine endliche Überdeckung besitzen, hätte man auch eine endliche Überdeckung von ganz X). Sei z.B. $K_1 := \overline{U_1(a_1)}$ eine solche Kugel. Wir betrachten (K_1, d) als vollständigen metrischen Raum. Die Mengen $G_\alpha \cap K_1$ bilden dann eine offene Überdeckung von K_1, aus der sich keine endliche Überdeckung auswählen läßt.

Wir wiederholen obige Überlegungen, nun aber mit einem 1/2-Netz für K_1. Ein solches existiert nach Lemma 1.21. Das Resultat ist eine Kugel $K_2 := \overline{U_{1/2}(a_2)}$, die sich nicht durch endlich viele der $K_1 \cap G_\alpha$ überdecken läßt. Wir fahren so fort, indem wir jedesmal die Maschenweite des Netzes halbieren. Schließlich gelangen wir zu einer Folge (K_n) ineinandergeschachtelter (d.h. $K_n \subset K_{n-1}$) abgeschlossener Kugeln $K_n = \overline{U_{1/2^{n-1}}(a_n)}$, die sich jeweils nicht durch endlich viele der Mengen G_α überdecken lassen. Wir betrachten die Folge (a_n) der Kugelmittelpunkte. Diese ist eine Cauchyfolge, denn für alle $m \geq n$ ist ja $a_m \in K_n = \overline{U_{1/2^{n-1}}(a_n)}$ und folglich $d(a_n, a_m) \leq 1/2^{n-1}$. Da X vollständig ist, besitzt die Folge (a_n) einen Grenzwert a^*. Da von n ab alle Glieder der Folge in K_n liegen und da K_n abgeschlossen ist, liegt a^* in jeder der Kugeln K_n und folglich auch im Durchschnitt $\cap K_n$.

Da $X \subseteq \bigcup_\alpha G_\alpha$, ist a^* innerer Punkt wenigstens einer der Mengen G_α, etwa G_{α_0}. Mit a^* liegt auch eine Umgebung $U_\varepsilon(a^*)$ in G_{α_0}. Wählen wir n so groß, dass $1/2^{n-2} < \varepsilon$, so ist für alle $y \in K_n = \overline{U_{1/2^{n-1}}(a_n)}$

$$d(y, a^*) \leq d(y, a_n) + d(a_n, a^*) \leq 1/2^{n-1} + 1/2^{n-1} = 1/2^{n-2}$$

und daher $K_n \subset G_{\alpha_0}$. Also wird K_n doch durch eine endliche Anzahl (genauer durch eine) der Mengen G_α überdeckt im Widerspruch zu unserer Konstruktion.

(c) $\Rightarrow$ (a). Angenommen, X besitzt eine unendliche Teilmenge S ohne Häufungspunkt in X. Dann besitzt jeder Punkt $x \in X$ eine Umgebung $U_r(x)$ (mit einem Radius $r = r(x)$), welche außer eventuell x selbst keinen weiteren Punkt aus S enthält. Das System $\{U_r(x)\}_{x \in X}$ bildet eine offene Überdeckung von X. Jede Menge aus diesem System überdeckt aber höchstens einen Punkt aus S. Da S unendlich viele Punkte enthält, läßt sich aus $\{U_r(x)\}_{x \in X}$ keine endliche Überdeckung von X auswählen. ■

Man überlegt sich leicht, dass jedes ε-Netz für eine Teilmenge A von X auch ein 2ε-Netz für die Abschließung $\overline{A}$ von A in (X, d) ist. Hieraus und aus Satz 1.22 erhält man leicht

Folgerung 1.23 *Für jede Teilmenge A eines metrischen Raumes (X,d) sind folgende Aussagen äquivalent:*

(a) *A ist relativ kompakt.*

(b) *Für jedes $\varepsilon > 0$ gibt es ein endliches ε-Netz für A (d.h. A ist total-beschränkt).*

Es gibt zahlreiche weitere Charakterisierungen kompakter Mengen, z.B. über die Gültigkeit eines Kugelschachtelungssatzes ($\nearrow$ Literatur).
Sehen wir uns nun stetige Funktionen auf kompakten Mengen an und definieren dazu

Definition 1.24 *Seien (X,d_X), (Y,d_Y) metrische Räume und $f : X \to Y$.*

(a) *Die Funktion f heißt* stetig in $x \in X$, *wenn für jedes $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass $d_Y\big(f(x),f(y)\big) < \varepsilon$ für alle $y \in X$ mit $d_X(x,y) < \delta$.*

(b) *f heißt* auf X stetig, *wenn f in jedem Punkt von X stetig ist.*

(c) *f heißt* gleichmäßig stetig auf X, *wenn es für jedes $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass $d_Y\big(f(x),f(y)\big) < \varepsilon$ für alle $x,y \in X$ mit $d_X(x,y) < \delta$.*

Von den zahlreichen Eigenschaften stetiger Funktionen auf kompakten Mengen erwähnen wir nur zwei, die uns aus der Analysis (wenigstens für $X = \mathbb{R}^n$) bekannt sind.

Satz 1.25 *Seien X,Y metrische Räume, $K \subseteq X$ kompakt und $f : K \to Y$ stetig. Dann ist $f(K)$ kompakt.*

Folgerung 1.26 (Satz von Weierstraß) *Sei X ein metrischer Raum, $K \subseteq X$ kompakt und $f : K \to \mathbb{R}$ stetig. Dann gibt es Punkte $x_{min}, x_{max} \in K$ so, dass*

$$f(x_{min}) \leq f(x) \leq f(x_{max}) \quad \textit{für alle } x \in X.$$

Satz 1.27 *Seien X,Y metrische Räume, $K \subseteq X$ kompakt und $f : K \to Y$ stetig. Dann ist f auf K gleichmäßig stetig.*

Die Beweise lassen sich wie im Fall $X = \mathbb{R}$ führen ($\nearrow$ Übung).

Für einige konkrete Räume hat man leichter überprüfbare Kompaktheitskriterien. So wissen wir bereits, dass Teilmengen des $\mathbb{R}^n$ oder $\mathbb{C}^n$ genau dann kompakt sind, wenn sie beschränkt und abgeschlossen sind. Unser nächstes Ziel ist ein Kompaktheitskriterium für $\big(C(X), \|\cdot\|_\infty\big)$. Dabei lassen wir zu, dass (X,d) ein beliebiger kompakter metrischer Raum ist. $C(X)$ besteht also aus allen stetigen Funktionen $f : X \to \mathbb{R}$. Wegen Folgerung 1.26 ist $\|f\|_\infty := \max_{x\in X} |f(x)|$ wohldefiniert und eine Norm auf $C(X)$. Die Vollständigkeit von $\big(C(X), \|\cdot\|_\infty\big)$ zeigt man wie oben. Für das angestrebte Kompaktheitskriterium benötigen wir noch einige Begriffe.

Definition 1.28 *Eine Menge $F \subseteq C(X)$ heißt*

- gleichmäßig beschränkt, *wenn es ein $M > 0$ gibt so, dass $\|f\|_\infty = \max_{x \in X} |f(x)| \leq M$ für alle $f \in F$.*
- punktweise beschränkt, *wenn für jedes $x \in X$ ein $M(x)$ existiert, so dass $|f(x)| \leq M(x)$ für alle $f \in F$.*
- gleichgradig stetig, *wenn für jedes $\varepsilon > 0$ ein δ existiert, so dass für jede Funktion $f \in F$ und beliebige $x, y \in X$ mit $d(x,y) < \delta$ gilt: $|f(x) - f(y)| < \varepsilon$.*

Satz 1.29 (Arzelà-Ascoli) *Sei (X,d) ein kompakter metrischer Raum. Eine Teilmenge F von $\big(C(X), \|\cdot\|\big)$ ist genau dann relativ kompakt, wenn sie punktweise beschränkt und gleichgradig stetig ist. In diesem Fall ist F auch gleichmäßig beschränkt.*

Beweis Wir zeigen zuerst, dass aus der relativen Kompaktheit die punktweise Beschränktheit und die gleichgradige Stetigkeit folgen. Aus Lemma 1.20 folgt die gleichmäßige und daher erst recht die punktweise Beschränktheit. Die gleichgradige Stetigkeit überlegen wir uns so: Sei $\varepsilon > 0$ beliebig. Nach Satz 1.22 und Folgerung 1.23 gibt es ein endliches $\frac{\varepsilon}{3}$-Netz $\{f_1, \ldots, f_n\}$ in F. Jede der Funktionen f_i ist nach Satz 1.27 gleichmäßig stetig auf X. Für jedes i finden wir also ein δ_i so, dass

$$|f_i(x) - f_i(y)| < \varepsilon/3 \quad \text{für alle } x, y \text{ mit } d(x,y) < \delta_i.$$

Sei $\delta := \min_i \delta_i$ und seien $x, y \in X$ mit $d(x,y) < \delta$ und $f \in F$ beliebig. Wir wählen f_i so, dass $\|f - f_i\|_\infty < \varepsilon/3$ und erhalten

$$|f(x) - f(y)| \leq |f(x) - f_i(x)| + |f_i(x) - f_i(y)| + |f_i(y) - f(y)| < 3 \cdot \frac{\varepsilon}{3} = \varepsilon.$$

Also ist F gleichgradig stetig.
Sei umgekehrt F punktweise beschränkt und gleichgradig stetig, und sei (f_n) eine Folge in F. Wegen der gleichgradigen Stetigkeit gibt es für jedes $\varepsilon > 0$ ein δ so, dass

$$|f_n(x) - f_n(y)| < \frac{\varepsilon}{3} \quad \text{für alle } x, y \in X \text{ mit } d(x,y) < \delta \text{ und alle } n \in \mathbb{N}. \tag{1.18}$$

Sei $\{x_1, \ldots, x_n\}$ ein endliches δ-Netz von X. Da die Menge $\{f_n(x_1)\}_{n \in \mathbb{N}} \subseteq \mathbb{R}$ beschränkt ist, enthält sie eine konvergente Teilfolge (Bolzano/Weierstraß), d.h. es gibt eine Teilfolge $(f_{11}, f_{12}, \ldots)$ von (f_n) so, dass $\big(f_{1k}(x_1)\big)_{k \geq 1}$ konvergiert. Mit der gleichen Begründung wählen wir eine Teilfolge $\big(f_{21}, f_{22}, \ldots\big)$ von $(f_{11}, f_{12}, \ldots)$ so, dass auch $\big(f_{2k}(x_2)\big)_{k \geq 1}$ konvergiert. Wir fahren so fort und erhalten nach n Schritten eine Teilfolge $(g_1, g_2, \ldots)$ von (f_n) mit der Eigenschaft, dass die Folge $\big(g_n(x_i)\big)_{n \geq 1}$ für jedes x_i aus dem δ-Netz konvergiert.

Insbesondere gibt es ein N so, dass

$$|g_n(x_i) - g_m(x_i)| < \frac{\varepsilon}{3} \quad \text{für alle } m, n \geq N \text{ und alle } i. \tag{1.19}$$

Seien $m, n \geq N$ und $x \in X$ beliebig. Wir wählen x_i so, dass $d(x, x_i) < \delta$, und erhalten mit (1.18) und (1.19):

$$|g_m(x) - g_n(x)| \leq |g_m(x) - g_m(x_i)| + |g_m(x_i) - g_n(x_i)| + |g_n(x_i) - g_n(x)| < 3 \cdot \frac{\varepsilon}{3} = \varepsilon.$$

Wenn wir die Folge (g_n) erst ab dem Index N betrachten, können wir unser bisheriges Ergebnis wie folgt zusammenfassen:

Für jedes $\varepsilon > 0$ gibt es eine Teilfolge (g_n) von (f_n) mit $\|g_n - g_m\|_\infty < \varepsilon$ für alle $m, n \in \mathbb{N}$.

Wir wählen nun speziell $\varepsilon = 1$ und finden eine Teilfolge (g_{1n}) von (f_n) mit $\|g_{1n} - g_{1m}\|_\infty < 1$ für alle $m, n \in \mathbb{N}$. Dann wiederholen wir obige Überlegungen und finden eine Teilfolge (g_{2n}) von (g_{1n}) mit $\|g_{2n} - g_{2m}\|_\infty < \frac{1}{2}$ für alle $m, n \in \mathbb{N}$. Wir fahren so fort und finden für jedes $k \in \mathbb{N}$ eine Teilfolge $(g_{(k+1),n})_{n \geq 1}$ von $(g_{k,n})_{n \geq 1}$ mit

$$\|g_{(k+1),n} - g_{(k+1),m}\|_\infty < \frac{1}{k+1}$$

für alle $m, n \in \mathbb{N}$. Nun betrachten wir die Folge (h_n) mit $h_n := g_{n,n}$ (Cantorsches Diagonalverfahren). Dann ist (h_n) eine Teilfolge von (f_n), und aus der Konstruktion folgt

$$\|h_n - h_m\|_\infty \leq \max\left\{\frac{1}{n}, \frac{1}{m}\right\}.$$

Also ist (h_n) eine Cauchyfolge in $C(X)$.

Aus der Vollständigkeit von $C(X)$ folgt die Konvergenz von (h_n). Die Folge (h_n) ist also eine konvergente Teilfolge von (f_n). Somit ist F relativ kompakt. ∎

2 Banachräume und beschränkte lineare Operatoren

2.1 Beschränkte lineare Operatoren

Seien X, Y lineare Räume über einem Körper $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Eine Abbildung $A : X \to Y$ heißt ein *linearer Operator*, wenn $A(\alpha x + \beta y) = \alpha Ax + \beta Ay$ für alle $x, y \in X$ und $\alpha, \beta \in \mathbb{K}$. Für lineare Operatoren $A, B : X \to Y$ erklären wir ihre *Summe* $A + B$ durch $(A + B)x := Ax + Bx$ und ein *Produkt mit Zahlen* $\alpha \in \mathbb{K}$ durch $(\alpha A)x := \alpha(Ax)$. Das *Produkt* $AB : X \to Z$ zweier Operatoren $B : X \to Y$ und $A : Y \to Z$ ist ihre Hintereinanderausführung $(AB)x = A(Bx)$. Die Abbildungen $A + B$, αA und AB sind wieder lineare Operatoren. Für jede lineare Abbildung $A : X \to Y$ ist $A0 = 0$. Es ist nämlich

$$A0 = A(0 + 0) = A0 + A0.$$

Eine Menge M eines normierten Raumes X heißt *beschränkt*, wenn es ein $c > 0$ so gibt, dass $\|x\|_X \le c$ für alle $x \in M$.

Definition 2.1 *Seien X, Y normierte Räume (über $\mathbb{K}$). Ein linearer Operator $A : X \to Y$ heißt* beschränkt, *wenn er beschränkte Mengen in beschränkte Mengen überführt.*

Satz 2.2 *Seien X, Y normierte Räume und $A : X \to Y$ ein linearer Operator. Dann sind folgende Aussagen äquivalent:*

(a) *A ist beschränkt.*

(b) *A überführt die Einheitskugel $B_1 := \{x \in X : \|x\|_X \le 1\}$ in eine beschränkte Menge.*

(c) *Es gibt eine Zahl $C \ge 0$ mit*

$$\|Ax\|_Y \le C\,\|x\|_X \quad \text{für alle } x \in X. \tag{2.1}$$

(d) *A ist gleichmäßig stetig auf X.*

(e) *A ist stetig im Punkt $0 \in X$.*

Beweis (a) $\Rightarrow$ (b). Klar, da B_1 beschränkt ist.

(b) $\Rightarrow$ (c). Sei C eine Konstante so, dass $\|Ax\|_Y \le C$ für alle $x \in B_1$. Für jeden Vektor $y \in X \backslash \{0\}$ ist dann $\frac{1}{\|y\|_X} \cdot y \in B_1$ und folglich

$$\left\| A \, \frac{y}{\|y\|_X} \right\|_Y \le C \quad \text{bzw.} \quad \|Ay\|_Y \le C\,\|y\|_X.$$

Dies gilt auch für $y = 0$.

(c) $\Rightarrow$ (d). Aus (c) und der Linearität von A folgt für alle $x, y \in X$:

$$\|Ax - Ay\|_Y = \|A(x-y)\|_Y \leq C\|x-y\|_X,$$

d.h. A ist gleichmäßig stetig auf X.

(d) $\Rightarrow$ (e). Dies ist wieder trivial.

(e) $\Rightarrow$ (a). Wegen $A0 = 0$ ist A genau dann stetig in 0, wenn für jedes $\varepsilon > 0$ ein $\delta > 0$ existiert so, dass $\|Ax\| < \varepsilon$ für alle x mit $\|x\| < \delta$. Insbesondere gibt es ein $\delta > 0$ so, dass $\|Ax\| \leq 1$ für alle x mit $\|x\| \leq \delta$. Sei nun $M \subseteq X$ beschränkt, d.h. $\|x\| \leq C$ für alle $x \in M$. Dann ist sicher $\|\frac{\delta}{C}x\| \leq \delta$ für alle $x \in M$ und folglich $\|A\frac{\delta}{C}x\| \leq 1$ bzw. $\|Ax\| \leq \frac{C}{\delta}$ für alle $x \in M$. Also ist AM beschränkt. ∎

Die Menge aller linearen beschränkten Operatoren $A : X \to Y$ bezeichnen wir mit $L(X,Y)$. Statt $L(X,X)$ schreiben wir auch $L(X)$ und statt $L(X,\mathbb{K})$ schreiben wir X'. X' heißt der zu X *duale Raum* und seine Elemente heißen *lineare beschränkte Funktionale* auf X.

Definition 2.3 *Sei $A \in L(X,Y)$. Die kleinste aller Zahlen $C \geq 0$ mit $\|Ax\|_Y \leq C\|x\|_X$ für alle $x \in X$, heißt* Norm *von A und wird mit $\|A\|$ oder $\|A\|_{L(X,Y)}$ bezeichnet.*

Die Existenz einer solchen Zahl muß man natürlich beweisen ($\nearrow$ Übung).

Lemma 2.4 *Für jeden Operator $A \in L(X,Y)$ gilt*

$$\|A\|_{L(X,Y)} = \sup_{x \neq 0} \frac{\|Ax\|_Y}{\|x\|_X} = \sup_{x:\|x\|=1} \|Ax\| = \sup_{x:\|x\|\leq 1} \|Ax\|.$$

Beweis Sei $C^* := \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|}$. Für alle $x \neq 0$ folgt aus $\|Ax\| \leq \|A\|\,\|x\|$, dass $\frac{\|Ax\|}{\|x\|} \leq \|A\|$. Wir gehen links zum Supremum über und erhalten $C^* \leq \|A\|$ sein. Andererseits gilt natürlich für jedes $x \neq 0$ $\frac{\|Ax\|}{\|x\|} \leq C^*$ bzw. $\|Ax\| \leq C^*\|x\|$. Aus der Minimalitätseigenschaft von $\|A\|$ folgt nun $\|A\| \leq C^*$, also $\|A\| = C^*$. Rest $\nearrow$ Übung. ∎

Satz 2.5 *Seien X, Y normierte Räume. Dann ist $L(X,Y)$ ein linearer Raum, und $\|\cdot\|_{L(X,Y)}$ eine Norm auf $L(X,Y)$, d.h. $L(X,Y)$ ist ebenfalls ein normierter Raum. Ist Y vollständig, so ist auch $L(X,Y)$ vollständig, d.h. ein Banachraum. Insbesondere ist der duale Raum X' eines normierten Raumes X vollständig.*

Beweis Offenbar ist für $A \in L(X,Y)$ stets $\|A\| \geq 0$, und aus $\|A\| = 0$ folgt wegen Lemma 2.4 sofort $A = 0$. Weiter ist für jedes $\alpha \in \mathbb{K}$

$$\|\alpha A\| = \sup_{x\neq 0} \frac{\|\alpha Ax\|}{\|x\|} = |\alpha| \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|} = |\alpha|\,\|A\|.$$

Die Dreiecksungleichung folgt schließlich aus

$$\begin{aligned}\|A+B\| &= \sup_{\|x\|\le 1}\|(A+B)x\| = \sup_{\|x\|\le 1}\|Ax+Bx\| \\ &\le \sup_{\|x\|\le 1}\big(\|Ax\|+\|Bx\|\big) \le \|A\|+\|B\|.\end{aligned}$$

Damit ist $L(X,Y)$ ein linearer und normierter Raum. Sei nun Y vollständig und (A_n) eine Cauchyfolge in $L(X,Y)$. Wegen

$$\|A_nx - A_mx\| = \|(A_n - A_m)x\| \le \|A_n - A_m\|\,\|x\|$$

ist dann für jedes $x \in X$ die Folge $(A_nx)_{n\ge 1}$ eine Cauchyfolge in Y. Diese konvergiert wegen der Vollständigkeit von Y. Ihren Grenzwert bezeichnen wir mit Ax. Hierdurch wird ein linearer Operator $A : X \to Y$ festgelegt.
Die Beschränktheit von A überlegen wir uns so. Da (A_n) eine Cauchyfolge ist, gibt es ein N so, dass $\|A_n - A_m\| \le \varepsilon$ für alle $m, n \ge N$. Für alle $x \in X$ und $m, n \ge N$ ist also

$$\|A_n\,x - A_m\,x\| \le \|A_n - A_m\|\,\|x\| \le \varepsilon\|x\|.$$

Lassen wir hierin $m \to \infty$ streben, folgt $\|A_n\,x - Ax\| \le \varepsilon\|x\|$ für alle $x \in X$, d.h. der Operator $A_n - A$ ist beschränkt. Da A_n beschränkt ist, ist $A = A_n - (A_n - A)$ als Differenz zweier beschränkter Operatoren ebenfalls beschränkt, gehört also zu $L(X,Y)$. Außerdem ist, wie wir gerade gesehen haben, $\|A_n - A\| \le \varepsilon$ für alle $n \ge N$, d.h. die Folge (A_n) konvergiert in der Norm von $L(X,Y)$ gegen A. ■

Satz 2.6 *Sei X ein normierter Raum, X_0 ein linearer dichter Teilraum von X, Y ein Banachraum und $A \in L(X_0,Y)$. Dann gibt es einen eindeutig bestimmten Operator $\tilde{A} \in L(X,Y)$, der auf X_0 mit A übereinstimmt. Für diesen gilt $\|\tilde{A}\|_{L(X,Y)} = \|A\|_{L(X_0,Y)}$.*

Der Operator $\tilde{A}$ heißt die *stetige Fortsetzung* von A auf X.

Beweis Wir stellen $x \in X$ als Grenzwert einer Folge $(x_n) \subseteq X_0$ dar. Wegen

$$\|Ax_n - Ax_m\| = \|A(x_n - x_m)\| \le \|A\|_{L(X_0,Y)}\|x_n - x_m\|$$

ist (Ax_n) eine Cauchyfolge in Y und folglich konvergent (man beachte, dass die Folge (x_n) konvergiert und daher eine Cauchyfolge ist). Weiter: ist $(y_n) \subseteq X_0$ eine andere Folge, die ebenfalls gegen $x \in X$ konvergiert, so ist

$$\|Ax_n - Ay_n\| = \|A(x_n - y_n)\| \le \|A\|\,\|x_n - y_n\| \to 0.$$

Also ist der Grenzwert $\lim Ax_n$ unabhängig von der konkreten Wahl der Folge $(x_n) \to x$. Wir bezeichnen diesen Grenzwert mit $\tilde{A}x$. Der so erklärte Operator $\tilde{A} : X \to Y$ ist offenbar linear, stimmt auf X_0 mit A überein, und ist beschränkt:

$$\|\tilde{A}x\| = \lim_{n\to\infty}\|Ax_n\| \le \lim \|A\|\,\|x_n\| = \|A\|\,\|\lim x_n\| = \|A\|\,\|x\|.$$

Hieraus folgt außerdem $\|\tilde{A}\| \leq \|A\|$. Die umgekehrte Ungleichung $\|A\| \leq \|\tilde{A}\|$ ergibt sich aus

$$\|A\| = \sup_{\substack{x\in X_0\\ \|x\|\leq 1}} \|Ax\| = \sup_{\substack{x\in X_0\\ \|x\|\leq 1}} \|\tilde{A}x\| \leq \sup_{\substack{x\in X\\ \|x\|\leq 1}} \|\tilde{A}x\| = \|\tilde{A}\|$$

(auf der rechten Seite wird das Supremum über einer größeren Menge gebildet).

■

Sind X, Y, Z normierte Räume und $B : X \to Y$, $A : Y \to Z$ lineare beschränkte Operatoren, so ist $AB \in L(X, Z)$, und es gilt $\|AB\|_{L(X,Z)} \leq \|A\|_{L(Y,Z)}\|B\|_{L(X,Y)}$. Für beliebiges $x \in X$ ist nämlich

$$\|ABx\|_Z \leq \|A\|_{L(Y,Z)}\|Bx\|_Y \leq \|A\|_{L(Y,Z)}\|B\|_{L(X,Y)}\|x\|_X.$$

Beispiel 1: Beschränktheit von linearen Operatoren $A : \mathbb{K}^n \to \mathbb{K}^n$. Wir versehen den linearen Raum $\mathbb{K}^n$ ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$) mit irgendeiner Norm $\|\cdot\|$ und wählen als Basis in $\mathbb{K}^n$ die Vektoren $e_i = (0, \ldots 0, 1, 0, \ldots 0)^T$ mit der 1 an der i-ten Stelle. Für *jede* lineare Abbildung $A : \mathbb{K}^n \to \mathbb{K}^n$ und jedes $x = \sum_{i=1}^n x_i e_i \in \mathbb{K}^n$ gilt dann

$$\|Ax\| = \|A\sum_i x_i e_i\| = \|\sum_i x_i Ae_i\| \leq \sum_i |x_i|\,\|Ae_i\| \leq \max_i \|Ae_i\| \sum_{i=1}^n |x_i|.$$

$\sum_{i=1}^n |x_i|$ ist gerade die Summennorm $\|x\|_1$, und aus Satz 1.9 wissen wir, dass es ein $C > 0$ so gibt, dass $\|x\|_1 \leq C\|x\|$ für alle $x \in \mathbb{K}^n$. Es ist also

$$\|Ax\| \leq C \max_i \|Ae_i\|\,\|x\| \quad \text{für alle } x \in \mathbb{K}^n, \tag{2.2}$$

d.h. *jede lineare Abbildung $A : \mathbb{K}^n \to \mathbb{K}^n$ ist beschränkt.*

Für die Norm von A gewinnen wir aus (2.2) nur die Abschätzung $\|A\| \leq C \max_i \|Ae_i\|$. Die exakte Bestimmung der Matrixnorm $\|A\|$ für verschiedene Normen $\|\cdot\|$ auf $\mathbb{K}^n$ ist i.Allg. recht schwierig. In der linearen Algebra wird für $A = (a_{ij})_{i,j=1}^n \in \mathbb{K}^{n\times n}$ gezeigt:

$$\begin{aligned}
\|x\|_1 &= \sum_{i=1}^n |x_i| && \text{auf } \mathbb{K}^n \Rightarrow \|A\| = \max\nolimits_j \textstyle\sum_{i=1}^n |a_{ij}|,\\
\|x\|_\infty &= \max_i |x_i| && \text{auf } \mathbb{K}^n \Rightarrow \|A\| = \max\nolimits_i \textstyle\sum_{j=1}^n |a_{ij}|,\\
\|x\|_2 &= \sqrt{\sum_i |x_i|^2} && \text{auf } \mathbb{K}^n \Rightarrow \|A\| = (\text{größter Eigenwert von } A^*A)^{1/2}.
\end{aligned}$$

Diese Normen heißen die *Spaltensummennorm, Zeilensummennorm* bzw. *Spektralnorm* von A.

Beispiel 2: Multiplikationsoperatoren. Sei X ein kompakter metrischer Raum, $C(X)$ der Banachraum der stetigen reellwertigen Funktion auf X, versehen mit der max-Norm, und $f \in C(X)$. Der Multiplikationsoperator

$$A : C(X) \to C(X), \quad u \mapsto fu$$

ist dann beschränkt:

$$\|fu\|_\infty = \sup_{x\in X} |f(x)u(x)| \leq \sup_{x\in X} |f(x)| \sup_{x\in X} |u(x)| = \|f\|_\infty \|u\|_\infty,$$

und es gilt $\|A\| \leq \|f\|_\infty$. In diesem Fall ist sogar $\|A\| = \|f\|_\infty$. Setzt man nämlich in der nach Definition für alle $u \in C(X)$ gültigen Abschätzung $\|Au\|_\infty \leq \|A\| \|u\|_\infty$ bzw. $\|fu\|_\infty \leq \|A\| \|u\|_\infty$ für u die konstante Funktion $u_0(x) = 1$ ein, folgt $\|f\|_\infty \leq \|A\|$. ■

In diesem Beispiel haben wir einen Weg kennengelernt, die Norm eines Operators A zu berechnen. Aus der Abschätzung $\|Ax\| \leq C\|x\|$ für alle x folgt $\|A\| \leq C$. Hat man darüberhinaus ein $x_0 \neq 0$ mit $\|Ax_0\| = C\|x_0\|$, so folgt $C = \frac{\|Ax_0\|}{\|x_0\|} \leq \|A\|$. Beide Ungleichungen zusammen ergeben $\|A\| = C$. Ein solcher Vektor x_0 muß jedoch nicht existieren.

Beispiel 3: Fredholmsche Integraloperatoren auf $\big(C(X), \|\cdot\|_\infty\big)$. Sei X die Abschließung einer beschränkten offenen Menge im $\mathbb{R}^n$, und X sei Jordanmeßbar. Für eine reellwertige Funktion $k \in C(X\times X)$ betrachten wir den Integraloperator

$$(Au)(t) := \int_X k(t,s)u(s)\, ds, \quad t \in X.$$

Wir überlegen uns zuerst, dass A ein linearer Operator von $C(X)$ nach $C(X)$ ist. Die Linearität von A ist klar: Für $u, v \in C(X)$ und $\alpha, \beta \in \mathbb{R}$ ist nämlich

$$\begin{aligned}
\big(A(\alpha u + \beta v)\big)(t) &= \int_X k(t,s)\big(\alpha u(s) + \beta v(s)\big)\, ds \\
&= \alpha \int_X k(t,s)u(s)\, ds + \beta \int_X k(t,s)v(s)\, ds \\
&= \alpha(Au)(t) + \beta(Av)(t).
\end{aligned}$$

Da dies für jedes $t \in X$ gilt, ist $A(\alpha u + \beta v) = \alpha Au + \beta Av$.

Wir zeigen nun, dass für jedes $u \in C(X)$ auch Au wieder eine stetige Funktion auf X ist. Für $u = 0$ ist dies offenbar erfüllt, und wir können $\|u\|_\infty \neq 0$ annehmen. Da k als stetige Funktion auf dem Kompakt $X \times X$ gleichmäßig stetig ist, gibt es zu beliebig vorgegebenem $\varepsilon > 0$ ein δ mit der Eigenschaft, dass

$$|k(t_1, s) - k(t_2, s)| \leq \frac{\varepsilon}{\|u\|_\infty \cdot \text{Volumen von } X}$$

für alle (t_1, s), $(t_2, s) \in X \times X$ mit $|(t_1, s) - (t_2, s)| < \delta$ bzw. für alle $t_1, t_2 \in X$ mit $|t_1 - t_2| < \delta$ und für alle s. Für alle $t_1, t_2 \in X$ mit $|t_1 - t_2| < \delta$ ist dann

$$\begin{aligned}
|(Au)(t_1) - (Au)(t_2)| &= \left| \int_X k(t_1, s)u(s)\, ds - \int_X k(t_2, s)u(s)\, ds \right| \\
&= \left| \int_X \big(k(t_1, s) - k(t_2, s)\big)\, u(s)\, ds \right| \\
&\leq \int_X |k(t_1, s) - k(t_2, s)|\, |u(s)|\, ds \\
&\leq \|u\|_\infty \int_X |k(t_1, s) - k(t_2, s)|\, ds \\
&\leq \|u\|_\infty \frac{\varepsilon}{\|u\|_\infty \mathrm{Vol}\,(X)} \int_X ds = \varepsilon.
\end{aligned}$$

Es ist nun auch leicht einzusehen, dass A sogar ein beschränkter Operator von $C(X)$ in sich ist. Für alle $u \in C(X)$ ist nämlich

$$\begin{aligned}
\|Au\|_\infty = \sup_{t \in X} \left| \int_X k(t, s)u(s)\, ds \right| &\leq \sup_{t \in X} \int_X |k(t, s)|\, |u(s)|\, ds \\
&\leq \|u\|_\infty \sup_{t \in X} \int_X |k(t, s)|\, ds
\end{aligned}$$

und damit insbesondere $\|A\| \leq \sup_{t \in X} \int_X |k(t, s)|\, ds$. Wir überlegen uns, dass hier sogar Gleichheit gilt.

Sei $t_0 \in X$ ein Punkt, in dem das Supremum $\sup_{t \in X} \int_X |k(t, s)|\, ds$ angenommen wird. (Ein solcher Punkt existiert wegen der Stetigkeit von k.) Falls $k(t_0, s) \geq 0$ für alle $s \in X$, ist die Identität $\|A\| = \int_X |k(t_0, s)|\, ds = \int_X k(t_0, s)\, ds$ leicht zu sehen, da für die konstante Funktion $u(s) = 1$ gilt

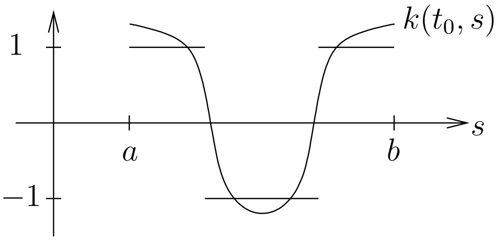
$$\begin{aligned}
\int_X |k(t_0, s)|\, ds &= \left| \int_X k(t_0, s)u(s)\, ds \right| \\
&\leq \sup_{t \in X} \left| \int_X k(t, s)u(s)\, ds \right| = \|Au\|_\infty \leq \|A\|\, \|u\|_\infty, \qquad (2.3)
\end{aligned}$$

d.h. $\int_X |k(t_0, s)| ds \leq \|A\|$.

Um auf ähnliche Weise die Normgleichheit $\|A\| = \int_X |k(t_0, s)|\, ds$ in dem Fall zu beweisen, wenn die Funktion $s \mapsto k(t_0, s)$ auf X ihr Vorzeichen wechselt, müßte man für u eine Funktion einsetzen, die gleich 1 ist wenn $k(t_0, s) > 0$ und gleich -1 wenn $k(t_0, s) < 0$. Für eine solche Funktion wäre nämlich

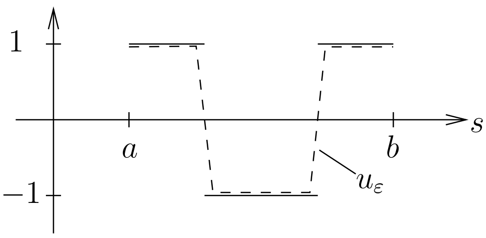
$$|k(t_0, s)| = k(t_0, s)\, u(s) \quad \text{für alle } s \in X,$$

und wir könnten wie in (2.3) weiterschließen. Wir müßten also $u(s) = \operatorname{sgn} k(t_0, s)$ wählen; dies können wir aber nicht, da diese Funktion i.Allg. unstetig ist. Man kann sich aber klarmachen, dass man die Funktion $s \mapsto \operatorname{sgn} k(t_0, s)$ durch eine stetige Funktion



u im Sinne der L^1-Norm approximieren kann, d.h. für jedes $\varepsilon > 0$ finden wir eine Funktion $u = u_\varepsilon \in C(X)$ so, dass $\|u_\varepsilon\|_\infty = 1$ und

$$\int_X |\operatorname{sgn} k(t_0, s) - u_\varepsilon(s)| \, ds < \varepsilon.$$



In einfachen Situationen ($\nearrow$ Skizze) ist dies sofort klar. Für einen allgemeinen Beweis $\nearrow$ Schröder S. 14.

Mit einer so gewählten Funktion $u = u_\varepsilon$ erhalten wir

$$\begin{aligned}
\int_X |k(t_0, s)| \, ds &= \left| \int_X k(t_0, s) \cdot \operatorname{sgn} k(t_0, s) \, ds \right| \\
&\le \left| \int_X k(t_0, s) u_\varepsilon(s) \, ds \right| + \left| \int_X k(t_0, s) \big(\operatorname{sgn} k(t_0, s) - u_\varepsilon(s) \big) \, ds \right| \\
&\le \sup_{t \in X} \left| \int_X k(t, s) u_\varepsilon(s) \, ds \right| + \sup_{s \in X} |k(t_0, s)| \int_X |\operatorname{sgn} k(t_0, s) - u_\varepsilon(s)| ds \\
&\le \|A\| \, \|u_\varepsilon\|_\infty + \sup_{s \in X} |k(t_0, s)| \cdot \varepsilon.
\end{aligned}$$

Da $\|u_\varepsilon\|_\infty = 1$ und da diese Abschätzung für beliebiges $\varepsilon > 0$ gilt, folgt $\int_X |k(t_0, s)| \, ds \le \|A\|$ und damit die Behauptung. ■

Man beachte die Analogie der Norm $\|A\| = \max_{t \in X} \int_X |k(t, s)| \, ds$ zur Zeilensummennorm $\max_i \sum_j |a_{ij}|$ einer Matrix $A = (a_{ij}) \in \mathbb{R}^n$.

2.2 Invertierbarkeit und Neumann-Reihe

Wir betrachten nun lineare beschränkte Operatoren, die einen Banachraum X in sich abbilden. Ein einfaches Beispiel eines solchen Operators ist der identische Operator $I = I_X$, der jedes Element $u \in X$ sich selbst zuordnet. Offenbar ist I beschränkt und $\|I\| = 1$. Die wesentlichen Definitionen dieses Abschnittes geben wir jedoch für den allgemeinen Fall. Eine Abbildung $A : X \to Y$ heißt *surjektiv* (oder Abbildung auf Y), wenn $AX = Y$, und *injektiv* (oder eineindeutig), wenn aus $Au_1 = Au_2$ folgt, dass $u_1 = u_2$. Für die Lösbarkeit der Gleichung $Au = f$ bedeutet Surjektivität, dass die Gleichung für *jede* rechte Seite $f \in Y$ lösbar ist, und Injektivität heißt, dass für jede rechte Seite *höchstens eine* Lösung existiert. Abbildungen, die surjektiv und injektiv sind, heißen *bijektiv* (oder invertierbar). Für jede bijektive Abbildung $A : X \to Y$ gibt es eine Abbildung $B : Y \to X$ mit

$$AB = I_Y, \quad BA = I_X. \tag{2.4}$$

Da nämlich die Gleichung $Au = f$ für jedes $f \in Y$ genau eine Lösung u besitzt, können wir B durch $Bf := u$ definieren. Dann gilt offenbar

$$ABf = Au = f \quad \text{und} \quad BAu = Bf = u.$$

Sind umgekehrt $A : X \to Y$, $B : Y \to X$ Abbildungen wie in (2.4), so ist A (und natürlich auch B) bijektiv. Für jedes $f \in Y$ ist nämlich Bf Lösung von $Au = f$ ($\Rightarrow$ Surjektivität), und aus $Au_1 = Au_2$ folgt $BAu_1 = BAu_2$ bzw. $u_1 = u_2$ ($\Rightarrow$ Injektivität). Weiter: Ist A invertierbar, so ist B wie in (2.4) eindeutig bestimmt. B heißt die zu A *inverse Abbildung* und wird mit A^{-1} bezeichnet.

Spezieller betrachten wir nun lineare Räume X, Y und lineare Abbildungen $A : X \to Y$. Mit $\operatorname{Im} A$ und $\operatorname{Ker} A$ bezeichnen wir das *Bild* und den *Kern* von A:

$$\operatorname{Im} A = \{y \in Y : \exists x \in X \text{ mit } Ax = y\}, \quad \operatorname{Ker} A = \{x \in X : \ Ax = 0\}.$$

$\operatorname{Im} A$ und $\operatorname{Ker} A$ sind lineare Teilräume von Y bzw. X ($\nearrow$ Übung). Offenbar ist $A : X \to Y$ genau dann surjektiv, wenn $\operatorname{Im} A = Y$. Dagegen ist A genau dann injektiv, wenn $\operatorname{Ker} A = \{0\}$. Aus der Injektivität folgt nämlich, dass die Gleichung $Au = 0$ höchstens eine Lösung besitzt. Da $u = 0$ eine Lösung ist, folgt $\operatorname{Ker} A = \{0\}$. Ist umgekehrt $\operatorname{Ker} A = \{0\}$ und $Au_1 = Au_2$ für $u_1, u_2 \in X$, so folgt aus der Linearität von A, dass $A(u_1 - u_2) = 0$, d.h. $u_1 - u_2 \in \operatorname{Ker} A$. Da $\operatorname{Ker} A = \{0\}$, ist $u_1 = u_2$.

Ist $A : X \to Y$ ein linearer invertierbarer Operator, so ist der inverse Operator A^{-1} wieder linear ($\nearrow$ Übung).

Schließlich betrachten wir den Fall, dass X, Y normierte Räume sind und $A \in L(X, Y)$. Dann ist zunächst $\operatorname{Ker} A$ *stets ein abgeschlossener Teilraum* von X ($\nearrow$ Übung), während $\operatorname{Im} A$ i.Allg. nicht abgeschlossen ist. Um ein einfaches Beispiel hierfür zu bekommen, betrachten wir neben dem Raum $C[0, 1]$ den Raum $\mathcal{P}$ aller Polynome auf $[0, 1]$, den wir ebenfalls mit der Supremumsnorm versehen. Ist $A : \mathcal{P} \to C[0, 1]$ der Operator der Einbettung (d.h. betrachten wir einfach jedes Polynom als Element von $C[0, 1]$), so ist offenbar $\operatorname{Im} A = \mathcal{P} \overset{\text{echt}}{\subset} C[0, 1]$, während $\operatorname{clos} \mathcal{P} = C[0, 1]$ auf Grund des Weierstraßschen Approximationsatzes.

Von besonderem Interesse ist die Frage, ob die Inverse A^{-1} eines linearen beschränkten Operators $A : X \to Y$ wieder beschränkt (also stetig) ist, d.h., ob die eindeutig bestimmte Lösung der Gleichung $Au = f$ *stetig* von der rechten Seite f abhängt. Bei der Anwendung von Näherungsverfahren ist man z.B. darauf angewiesen, dass kleine Störungen von f (etwa durch unvermeidbare Rundungsfehler) auch nur kleine Störungen der Lösung $u = A^{-1}f$ bewirken.

Beispiel 1: Auf $C[0,1]$ definieren wir den Operator A durch

$$(Au)(t) = \int_0^t u(s)ds, \quad t \in [0,1].$$

Wie wir aus der Analysis I wissen, bildet A den Raum $C[0,1]$ ab in den Raum C_0^1 der auf $[0,1]$ stetig differenzierbaren Funktionen f mit der Eigenschaft $f(0) = 0$. Diese Abbildung ist surjektiv: für jede Funktion $f \in C_0^1$ ist nämlich $f' \in C[0,1]$ und

$$(Af')(t) = \int_0^t f'(s)ds = f(t) - f(0) = f(t),$$

und A ist injektiv: aus $(Au)(t) = 0$ für alle t folgt ja

$$0 = \frac{d}{dt}\int_0^t u(s)ds = u(t) \quad \text{für alle } t \in [0,1].$$

Die Abbildung $A : C[0,1] \to C_0^1$ ist also invertierbar, und die inverse Abbildung A^{-1} ist offenbar gegeben durch

$$A^{-1} : C_0^1 \to C[0,1], \quad f \mapsto f'. \tag{2.5}$$

Versehen wir C_0^1 mit der Supremumsnorm, so ist $A : C[0,1] \to C_0^1$ beschränkt:

$$\|Au\|_\infty = \sup_{t\in[0,1]} \left| \int_0^t u(s)ds \right| \leq \|u\|_\infty \sup_{t\in[0,1]} \int_0^t ds = \|u\|_\infty.$$

Die inverse Abbildung A^{-1} (vgl. (2.5)) ist jedoch bezüglich dieser Norm auf C_0^1 *nicht beschränkt.* Aus ihrer Beschränktheit würde nämlich folgen, dass $\|u'\|_\infty \leq C\|u\|_\infty$ für alle $u \in C_0^1$. Mit $u(x) = x^n$ erhält man $n \leq C$ für alle $n \in \mathbb{N}$, ein Widerspruch. ■

Wir definieren daher: Sind X, Y normierte Räume und ist $A \in L(X,Y)$ ein invertierbarer Operator, so heißt A *stetig invertierbar*, falls A^{-1} stetig ist, d.h. falls $A^{-1} \in L(Y,X)$. Ein zentrales Resultat der Funktionalanalysis ist der folgende Satz.

Satz 2.7 (Satz von Banach über den inversen Operator) *Seien X und Y Banachräume und $A \in L(X,Y)$ ein invertierbarer Operator. Dann ist A stetig invertierbar.*

Wir werden diesen Satz erst später beweisen und untersuchen zunächst einfacher zu beweisende Bedingungen für die stetige Invertierbarkeit.

Satz 2.8 *Seien X, Y normierte Räume, X sei vollständig, und $A \in L(X,Y)$ sei ein Operator, dessen Bild in Y dicht liegt. Dann ist A genau dann stetig invertierbar, wenn es ein $C > 0$ gibt, so dass*

$$\|Ax\| \geq C\|x\| \quad \textit{für alle } x \in X. \tag{2.6}$$

Gilt für einen Operator $A \in L(X,Y)$ eine Abschätzung (2.6) mit $C > 0$, so heißt A auch *nach unten beschränkt*.

Beweis Wenn A stetig invertierbar ist, so gilt für alle $x \in X$

$$\|x\| = \|A^{-1}Ax\| \leq \|A^{-1}\| \, \|Ax\| \quad \text{bzw.} \quad \|Ax\| \geq \frac{1}{\|A^{-1}\|} \, \|x\|,$$

d.h. (2.6) gilt mit $C = \|A^{-1}\|^{-1}$. Sei umgekehrt A ein Operator mit dichtem Bild, und sei (2.6) erfüllt. Aus $Ax = 0$ und (2.6) folgt sofort $x = 0$, d.h. A ist injektiv. Wir zeigen, dass aus (2.6) auch die Abgeschlossenheit des Bildes von A folgt. Sei (Ax_n) eine Folge in $\operatorname{Im} A$, die gegen $y \in Y$ konvergiert. Dann ist (Ax_n) eine Cauchyfolge in Y, und aus

$$\|Ax_n - Ax_m\| = \|A(x_n - x_m)\| \geq C\|x_n - x_m\|$$

folgt, dass (x_n) eine Cauchyfolge in X ist. Da X vollständig ist, konvergiert die Folge (x_n) gegen ein $x \in X$. Wegen der Stetigkeit von A konvergiert dann Ax_n gegen Ax. Es ist also $Ax = y$ und $y \in \operatorname{Im} A$. Da das Bild von A abgeschlossen ist und zugleich dicht in Y liegt, ist $\operatorname{Im} A = Y$. Folglich ist A invertierbar. Die Inverse A^{-1} ist beschränkt. Für alle $y \in Y$ ist nämlich

$$\|A^{-1}y\| \overset{(2.6)}{\leq} \frac{1}{C} \, \|AA^{-1}y\| = \frac{1}{C} \, \|y\|,$$

wobei wir (2.6) auf das Element $x = A^{-1}y$ angewandt haben. Außerdem ist klar, dass $\|A^{-1}\| \leq 1/C$. ■

Wir beschäftigen uns nun mit dem Fall $X = Y$ genauer. In diesem Fall haben wir neben der Vektorraumstruktur noch eine weitere Operation in $L(X)$ erklärt: die Multiplikation oder Hintereinanderausführung AB zweier Operatoren $A, B \in L(X)$. Für diese Operation gilt für alle $A, B, C \in L(X)$ und alle $\alpha \in K$,

$$(AB)C = A(BC) \qquad \text{(Assoziativgesetz)} \tag{2.7}$$

$$(A + B)C = AC + BC, \quad A(B + C) = AB + AC \qquad \text{(Distributivgesetze)} \tag{2.8}$$

$$(\alpha A)B = A(\alpha B) = \alpha(AB) \qquad \text{und} \tag{2.9}$$

$$IA = AI = A, \tag{2.10}$$

und für die Norm gilt – wie wir bereits wissen –

$$\|AB\| \le \|A\|\,\|B\|. \tag{2.11}$$

Ein linearer Raum mit einer zusätzlichen Operation – der Multiplikation – heißt eine *Algebra*, wenn (2.7) – (2.9) gelten, und eine *Algebra mit Einselement*, wenn auch noch (2.10) gilt. Ein normierter linearer Raum, der eine Algebra ist, heißt *normierte Algebra*, wenn (2.11) erfüllt ist. Von einer *normierten Algebra mit Einselement* spricht man, wenn es ein Einselement I gibt, für das

$$\|I\| = 1 \tag{2.12}$$

gilt. Ist schließlich der einer normierten Algebra zugrunde liegende normierte Raum ein Banachraum, spricht man von einer *Banachalgebra*.
In diesem Sinn ist für jeden normierten Raum X die Menge $L(X)$ eine normierte Algebra mit Einselement, und $L(X)$ ist eine Banachalgebra, falls X ein Banachraum ist (Satz 2.5). Ebenso ist $\big(C(K),\ \|\cdot\|_\infty\big)$ für jeden Kompakt K eine Banachalgebra.
Da $L(X)$ eine Algebra ist, können wir insbesondere für jeden Operator $A \in L(X)$ seine Potenzen bilden: $A^0 := I$, $A^1 = A$, $A^k = AA^{k-1}$ für $k \ge 2$. Dies nutzen wir in dem folgenden Satz aus, der auch für beliebige Banachalgebren gilt.

Satz 2.9 (Neumann-Reihe) *Sei X ein Banachraum und $A \in L(X)$ ein Operator mit $\|A\| < 1$. Dann ist $I - A$ stetig invertierbar auf X, und es gilt die Abschätzung*

$$\|(I-A)^{-1}\| \le \frac{1}{1-\|A\|}\,. \tag{2.13}$$

Die Inverse $(I-A)^{-1}$ läßt sich in die Neumann-Reihe $(I-A)^{-1} = \sum_{k=0}^{\infty} A^k$ *entwickeln. Dabei gilt die Fehlerabschätzung*

$$\left\|(I-A)^{-1} - \sum_{k=0}^{n} A^k\right\| \le \frac{\|A\|^{n+1}}{1-\|A\|}\,. \tag{2.14}$$

Beweis Wegen

$$\left\|\sum_{k=m+1}^{n} A^k\right\| \le \sum_{k=m+1}^{n} \|A^k\| \le \sum_{k=m+1}^{n} \|A\|^k$$

ist $\sum_{k=0}^{\infty} A^k$ eine „Cauchyreihe“ in $L(X)$ (d.h. die Partialsummen dieser Reihe bilden eine Cauchyfolge). Da $L(X)$ vollständig ist, konvergiert diese Reihe gegen einen Operator $B \in L(X)$. Für diesen gilt

$$AB = A\sum_{k=0}^{\infty} A^k = \sum_{k=0}^{\infty} A^{k+1} = \sum_{k=1}^{\infty} A^k = B - I,$$

d.h. $B - AB = I$ oder $(I - A)B = I$. Analog sieht man auch, dass $B(I - A) = I$ ist (oder man beruft sich darauf, dass $AB = BA$). Also ist $I - A$ invertierbar, und $(I - A)^{-1} = B = \sum_{k=0}^{\infty} A^k$.
Die Beschränktheit von B und die Normabschätzung (2.13) folgen aus

$$\|B\| = \left\| \sum_{k=0}^{\infty} A^k \right\| \leq \sum_{k=0}^{\infty} \|A\|^k = \frac{1}{1 - \|A\|} \quad \text{(geometrische Reihe)},$$

und die Fehlerabschätzung (2.14) erhält man aus

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

■

Beispiel 2: Lösbarkeit Fredholmscher Integralgleichungen 2. Art. Wir betrachten wie in Beispiel 3 des vorigen Abschnitts den Operator

$$(Au)(t) = \int_X k(t,s)u(s)\,ds, \quad t \in X, \tag{2.15}$$

mit $k \in C(X \times X)$ reellwertig. Mit der in diesem Beispiel bestimmten Norm des Operators $A \in L\big(C(X)\big)$ erhalten wir aus Satz 2.9 :

Folgerung 2.10 *Sei $\|A\| = \sup_{t \in X} \int_X |k(t,s)|\,ds < 1$. Dann besitzt die Fredholmsche Integralgleichung 2. Art,*

$$\big((I - A)u\big)(t) = u(t) - \int_X k(t,s)u(s)\,ds = f(t), \quad t \in X, \tag{2.16}$$

für jede rechte Seite $f \in C(X)$ eine eindeutig bestimmte Lösung $u \in C(X)$.

Außerdem bietet uns Satz 2.9 eine Möglichkeit, die Lösung u dieser Gleichung näherungsweise zu bestimmen; es ist ja

$$u(t) = f(t) + (Af)(t) + (A^2 f)(t) + \dots .$$

Um diese Reihenentwicklung ausnutzen zu können, benötigen wir Potenzen des Operators (2.15). Für jedes $y \in C(X)$ ist

$$\begin{aligned} (A^2 y)(t) &= \int_X k(t,s)(Ay)(s)\,ds = \int_X k(t,s) \int_X k(s,r)y(r)\,dr\,ds \\ &= \int_X \left(\int_X k(t,s)k(s,r)\,ds \right) y(r)\,dr = \int_X k_2(t,r)y(r)\,dr \end{aligned}$$

mit $k_2(t,r) := \int_X k(t,s)k(s,r)\,ds$. Analog erhält man mit

$$k_n(t,r) := \int_X k(t,s)k_{n-1}(s,r)\,ds \quad \text{für } n \geq 3,$$

dass

$$(A^n y)(t) = \int_X k_n(t,r)y(r)\,dr$$

und folglich eine Darstellung der Lösung u von (2.16) mit Hilfe der sogenannten *iterierten Kerne* k_n

$$u(t) = f(t) + \sum_{n=1}^{\infty} \int_X k_n(t,s)f(s)\,ds, \quad t \in X.$$

∎

Die Invertierbarkeit von $I - A$ und die Konvergenz der Reihe $\sum_{n=0}^{\infty} A^n$ gegen $(I-A)^{-1}$ läßt sich auch unter schwächeren Voraussetzungen an A als in Satz 2.9 beweisen. Wir haben ja die Forderung $\|A\| < 1$ nur benutzt, um die Konvergenz von $\sum_{n=0}^{\infty} A^n$ zu garantieren. Nun kann man aber genauso wie für gewöhnliche Zahlenreihen auch für Reihen in Banachräumen das folgende *Wurzelkriterium* zeigen.

Satz 2.11 (Wurzelkriterium) *Ist X ein Banachraum und $x_n \in X$, so ist die Reihe $\sum_{n=1}^{\infty} x_n$ für*

$$\limsup \sqrt[n]{\|x_n\|} \begin{cases} < 1 & \textit{konvergent} \\ > 1 & \textit{divergent.} \end{cases}$$

Im Falle $\limsup \sqrt[n]{\|x_n\|} = 1$ *ist keine Aussage möglich.*

Um die Konvergenz der Neumannreihe $\sum_{n=0}^{\infty} A^n$ für einen Operator $A \in L(X)$ zu sichern, müssen wir also nur $\limsup \sqrt[n]{\|A^n\|} < 1$ verlangen (Wurzelkriterium im Banachraum $L(X)$). Wir gelangen so zu

Satz 2.12 (Neumann-Reihe) *Sei X ein Banachraum und $A \in L(X)$ ein Operator mit $\limsup_{n\to\infty} \sqrt[n]{\|A^n\|} < 1$. Dann ist $I - A$ stetig invertierbar, und es gilt*

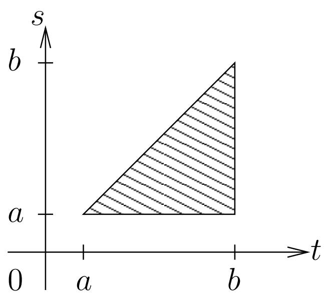
$$(I-A)^{-1} = \sum_{k=0}^{\infty} A^k.$$

Man kann sogar beweisen, dass für jeden Operator $A \in L(X)$ der *Grenzwert* $\lim_{n\to\infty} \sqrt[n]{\|A^n\|}$ existiert, d.h. der limes superior in Satz 2.12 kann durch einen limes ersetzt werden. Darauf gehen wir im Abschnitt 4.1 genauer ein.

Beispiel 3: Lösbarkeit von Volterra-Gleichungen 2. Art. Wir werden in diesem Beispiel sehen, dass die Konvergenzbedingung aus Satz 2.12 wesentlich schwächer ist als die aus Satz 2.9. Wir betrachten auf $C[a,b]$ (mit der sup-Norm) den Volterraschen Integraloperator

$$(Ku)(t) = \int_a^t k(t,s)u(s)\,ds, \quad t \in [a,b],$$

wobei k eine auf dem Dreieck $a \le s \le t \le b$ stetige Funktion ist.



Wie in Beispiel 3 aus 2.1 kann man zeigen, dass K ein linearer und beschränkter Operator von $C[a,b]$ in $C[a,b]$ ist. Wir suchen Lösungen der *Volterragleichung*

$$u(x) - \int_a^x k(x,t)u(t)\,dt = f(x), \quad x \in [a,b],\ f \in C[a,b], \tag{2.17}$$

die wieder in $C[a,b]$ liegen. Die Gleichung (2.17) können wir auch als

$$(I-K)u = f$$

schreiben. Für die Anwendbarkeit des Wurzelkriteriums untersuchen wir die iterierten K^n. Sei zur Abkürzung $M := \max_{a\le s\le t\le b} |k(t,s)|$. Dann ist für jede Funktion $u \in C[a,b]$ und jedes $x \in [a,b]$:

$$\begin{aligned}
|(Ku)(x)| &= \left|\int_a^x k(x,t)u(t)\,dt\right| \le \|u\|_\infty M(x-a),\\
|(K^2u)(x)| &= \left|\int_a^x k(x,t)(Ku)(t)\,dt\right| \le \int_a^x M\|u\|_\infty M(t-a)\,dt\\
&= M^2\|u\|_\infty \frac{(x-a)^2}{2},\\
|(K^3u)(x)| &= \left|\int_a^x k(x,t)(K^2u)(t)\,dt\right| \le \int_a^x M\cdot M^2\|u\|_\infty \frac{(t-a)^2}{2}\,dt\\
&= M^3\|u\|_\infty \frac{(x-a)^3}{6}
\end{aligned}$$

und allgemein

$$|(K^nu)(x)| \le M^n \frac{(x-a)^n}{n!}\|u\|_\infty.$$

Daher ist

$$\|K^nu\|_\infty \le M^n \frac{(b-a)^n}{n!}\|u\|_\infty$$

und schließlich

$$\|K^n\| \le M^n \frac{(b-a)^n}{n!}.$$

Wegen $\sqrt[n]{n!} \to \infty$ ist also $\limsup_{n\to\infty} \sqrt[n]{\|K^n\|} = 0$.

Folgerung 2.13 *Die Volterrasche Integralgleichung* (2.17) *besitzt für jede rechte Seite $f \in C[a,b]$ genau eine Lösung $u \in C[a,b]$.*

Diese Aussage gilt also unabhängig von der Größe von $\|K\|$. Man beachte, dass $\|K\|$ beliebig groß gemacht werden kann, indem die Kernfunktion k durch ein entsprechendes Vielfaches von k ersetzt wird. ■

2.3 Hilberträume und ihre Geometrie

Im $\mathbb{R}^n$ ist die Euklidsche Norm dadurch ausgezeichnet, dass sie von einem Skalarprodukt induziert wird, welches geometrische Beschreibungen wie den Begriff „orthogonale Transformation" oder den „Winkel" zwischen Vektoren ermöglicht. Eine ähnliche Struktur steht in speziellen unendlichdimensionalen Räumen zur Verfügung.

Definition 2.14 *Sei H ein linearer Raum über $\mathbb{K}$. Eine Abbildung $\langle\cdot,\cdot\rangle : H \times H \to \mathbb{K}$ heißt* Skalarprodukt *oder* inneres Produkt *auf H, wenn*

(a) $\langle \alpha_1 x_1 + \alpha_2 x_2, y\rangle = \alpha_1\langle x_1, y\rangle + \alpha_2\langle x_2, y\rangle$ *für alle $x_1, x_2, y \in H$ und $\alpha_1, \alpha_2 \in \mathbb{K}$ (Linearität),*

(b) $\langle x, y\rangle = \overline{\langle y, x\rangle}$ *für alle $x, y \in H$ (Symmetrie),*

(c) $\langle x, x\rangle \geq 0$ *für alle $x \in H$, und $\langle x, x\rangle = 0$ genau dann, wenn $x = 0$ (positive Definitheit).*

Das Paar $(H, \langle\cdot,\cdot\rangle)$ oder einfach H selbst heißt dann Prähilbertraum.

Aus diesen Bedingungen folgt, dass

$$\begin{aligned}\langle x, \alpha_1 y_1 + \alpha_2 y_2\rangle &= \overline{\langle \alpha_1 y_1 + \alpha_2 y_2, x\rangle} = \overline{\alpha_1\langle y_1, x\rangle + \alpha_2\langle y_2, x\rangle}\\ &= \overline{\alpha_1}\langle x, y_1\rangle + \overline{\alpha_2}\langle x, y_2\rangle.\end{aligned}$$

Im Falle $\mathbb{K} = \mathbb{C}$ ist also $\langle\cdot,\cdot\rangle$ *antilinear* bezüglich der zweiten Komponente.

Beispiele: Man überprüft leicht, dass die folgenden Räume Prähilberträume sind:

1. $H = \mathbb{R}^n$ mit $\langle (x_1,\ldots,x_n), (y_1,\ldots,y_n)\rangle := \sum_{i=1}^n x_i y_i$.

2. $H = \mathbb{C}^n$ mit $\langle (x_1,\ldots,x_n), (y_1,\ldots,y_n)\rangle := \sum_{i=1}^n x_i \overline{y_i}$.

3. $H = \ell^2$ mit $\langle (x_1, x_2,\ldots), (y_1, y_2,\ldots)\rangle := \sum_{i=1}^\infty x_i \overline{y_i}$ (aus der Hölderungleichung wissen wir, dass diese Reihe konvergiert).

4. $H = C(X)$ mit $\langle f, g\rangle := \int_X f(x)\overline{g(x)}\,dx$
(die Existenz des Integrals ist klar, da $f \cdot \overline{g}$ auf X wieder stetig ist).

5. $H = L^2(X)$ mit $\langle f, g\rangle := \int_X f(x)\overline{g(x)}\,dx$
(die Existenz des Integrals folgt aus der Hölderungleichung). ∎

Satz 2.15 (Cauchy-Schwarz Ungleichung) *In jedem Prähilbertraum H mit Skalarprodukt $\langle\cdot,\cdot\rangle$ gilt*

$$|\langle x, y\rangle|^2 \leq \langle x, x\rangle\langle y, y\rangle \quad \textit{für alle } x, y \in H.$$

Beweis Für alle $x, y \in H$ und $\alpha \in \mathbb{K}$ gilt

$$0 \leq \langle x + \alpha y, x + \alpha y\rangle = \langle x, x\rangle + \alpha\langle y, x\rangle + \overline{\alpha}\langle x, y\rangle + |\alpha|^2\langle y, y\rangle. \tag{2.18}$$

Falls $\langle x, x\rangle = \langle y, y\rangle = 0$, setzen wir in (2.18) $\alpha := -\langle x, y\rangle$ und erhalten

$$0 \leq -\langle x, y\rangle\langle y, x\rangle - \overline{\langle x, y\rangle}\langle x, y\rangle = -2\langle x, y\rangle\overline{\langle x, y\rangle} = -2|\langle x, y\rangle|^2$$

und hieraus die Behauptung $|\langle x, y\rangle|^2 \leq 0$.
Sei nun eine der Zahlen $\langle x, x\rangle$, $\langle y, y\rangle$ ungleich 0, etwa $\langle y, y\rangle \neq 0$. Dann setzen wir $\alpha := -\frac{\langle x,y\rangle}{\langle y,y\rangle}$ in (2.18) ein und erhalten

$$\begin{aligned} 0 &\leq \langle x, x\rangle - \frac{\langle x, y\rangle\langle y, x\rangle}{\langle y, y\rangle} - \frac{\overline{\langle x, y\rangle}\langle x, y\rangle}{\langle y, y\rangle} + \frac{|\langle x, y\rangle|^2\langle y, y\rangle}{\langle y, y\rangle^2} \\ &= \langle x, x\rangle - \frac{|\langle x, y\rangle|^2}{\langle y, y\rangle} \end{aligned}$$

und hieraus die Behauptung $|\langle x, y\rangle|^2 \leq \langle x, x\rangle\langle y, y\rangle$. ∎

Anmerkung Die Definitheit von $\langle\cdot,\cdot\rangle$ haben wir in diesem Beweis nicht benutzt.

Folgerung 2.16 *Jeder Prähilbertraum wird mit der Norm $\|x\| := \sqrt{\langle x, x\rangle}$ zu einem normierten Raum.*

Beweis Die ersten beiden Normaxiome sind klar; wir überlegen uns nur die Dreiecksungleichung. Im Fall $\mathbb{K} = \mathbb{R}$ ist für beliebige $x, y \in H$ wegen Satz 2.15

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y\rangle = \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \\ &\leq \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2. \end{aligned}$$

Für $\mathbb{K} = \mathbb{C}$ kann man wie folgt argumentieren:

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

∎

Wir fassen im weiteren Prähilberträume stets als normiert bezüglich der Norm $\|x\| = \sqrt{\langle x, x\rangle}$ auf. Die Cauchy-Schwarz-Ungleichung kann man dann auch schreiben als

$$|\langle x, y\rangle| \le \|x\| \, \|y\|.$$

Für die oben betrachteten Beispiele für Prähilberträume ist diese Cauchy-Schwarz-Ungleichung nichts anderes als die Hölderungleichung im Fall $p = 2$. Wir vermerken noch einige Beziehungen zwischen Skalarprodukt und Norm, die sich leicht nachrechnen und elementargeometrisch deuten lassen ($\nearrow$ Übung).

Satz 2.17 *Ist H Prähilbertraum, so gelten für alle $x, y \in H$*

(a) *die* Parallelogrammgleichung

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

(b) *die* Polarisationsformel

$$\begin{aligned} \langle x, y\rangle &= \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2) \qquad && \textit{falls } \mathbb{K} = \mathbb{R}, \\ \langle x, y\rangle &= \frac{1}{4}\sum_{k=0}^{3} i^k \|x + i^k y\|^2 && \textit{falls } \mathbb{K} = \mathbb{C}. \end{aligned}$$

Anmerkungen

1. Aussage (b) zeigt auch die Stetigkeit des Skalarprodukts.

2. Gilt in irgendeinem normierten Raum die Parallelogrammgleichung, so wird durch (b) ein Skalarprodukt definiert, das diese Norm erzeugt.

3. Der Raum $(C[a, b], \|\cdot\|_\infty)$ ist kein Prähilbertraum. Man findet nämlich leicht Funktionen $x, y \in C[a, b]$, die die Parallelogrammgleichung für $\|\cdot\|_\infty$ verletzen. ($\nearrow$ Übung)

Definition 2.18 *Ein Prähilbertraum heißt* Hilbertraum, *wenn er bezüglich der durch das Skalarprodukt definierten Norm vollständig ist.*

Von den oben erwähnten Prähilberträumen ist nur $(C[a, b], \|\cdot\|_2)$ kein Hilbertraum. Seine Vervollständigung $(L^2[a, b], \|\cdot\|_2)$ ist aber ein Hilbertraum. Man kann sich allgemein überlegen, dass man die Vervollständigung eines Prähilbertraumes zu einem Hilbertraum machen kann durch Einführung eines geeigneten Skalarprodukts, nämlich

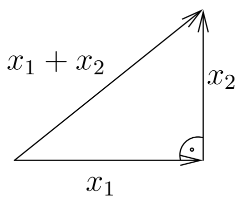
$$\langle (x_n)^\sim, (y_n)^\sim\rangle_\sim := \lim \langle x_n, y_n\rangle.$$

Zwei Vektoren im $\mathbb{R}^n$ stehen genau dann senkrecht aufeinander, wenn ihr Skalarprodukt gleich 0 ist. Wir definieren daher allgemein

Definition 2.19 *Zwei Elemente eines Prähilbertraumes* $(H, \langle\cdot,\cdot\rangle)$ *heißen* zueinander orthogonal*, wenn ihr Skalarprodukt gleich* 0 *ist. Steht* x *zu* y *orthogonal, schreiben wir* $x \perp y$*. Zwei Teilmengen* $M_1, M_2 \subseteq H$ *heißen* zueinander orthogonal*, wenn* $x_1 \perp x_2$ *für alle* $x_1 \in M_1$ *und* $x_2 \in M_2$*. Wir schreiben dann* $M_1 \perp M_2$*.*

Satz 2.20 (Pythagoras) *Sind* $x_1, \ldots, x_n$ *Elemente eines Prähilbertraumes mit* $x_j \perp x_k$ *für alle* $j \neq k$*, so ist*

$$\|x_1 + \ldots + x_n\|^2 = \|x_1\|^2 + \ldots + \|x_n\|^2.$$

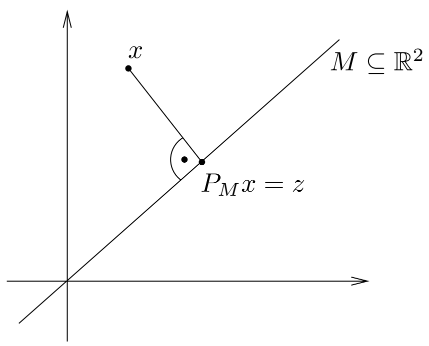


Beweis Wir überlegen uns die Aussage nur für zwei Vektoren $x, y \in H$ mit $x \perp y$. Der allgemeine Fall folgt mit vollständiger Induktion. Wir haben

$$\begin{aligned} \|x+y\|^2 &= \langle x+y, x+y\rangle = \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \\ &= \|x\|^2 + \underbrace{\langle x,y\rangle + \overline{\langle x,y\rangle}}_{=0} + \|y\|^2 = \|x\|^2 + \|y\|^2. \end{aligned}$$

■

Wir wenden uns nun einem Problem zu, welches aus der Approximationstheorie stammt. Wir gehen aus von einem Hilbertraum H und einem abgeschlossenen Teilraum M von H und betrachten für $x \in H$ folgendes Problem: Wie gut läßt sich x durch Elemente aus M approximieren? Besonders gut wäre es, Elemente in M zu finden, für die der Abstand zu x minimal wird. Gibt es solche Elemente? Wie viele? Im Fall $H = \mathbb{R}^2$ ist die Antwort klar: Wir fällen einfach das Lot von x auf M. Der Fußpunkt des Lotes ist dann das eindeutig bestimmte Element von M mit minimalem Abstand von x.



Satz 2.21 *Sei* H *ein Hilbertraum und* $M \subseteq H$ *ein abgeschlossener linearer Teilraum. Dann existiert für jedes* $x \in H$ *genau ein* $z = P_M x \in M$ *so, dass*

$$\|x - z\| = \inf_{y \in M} \|x - y\|.$$

Das Element z *heißt* Element der besten Approximation *zu* x*. Es gilt* $x - z \perp M$*.*

Beweis Wir überlegen uns zuerst, dass es ein Element $z \in M$ gibt, welches den Abstand zu x minimiert. Sei $d := \inf_{y \in M} \|x-y\|$. Wir wählen eine Folge $(y_n) \subseteq M$ mit $\|x - y_n\| \to d$.
Wir zeigen, dass (y_n) eine Cauchyfolge in M ist. Aus der Parallelogrammgleichung $\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2$ (Satz 2.17 (a)) erhalten wir mit $u = x - y_n$ und $v = x - y_m$:

$$\|y_m - y_n\|^2 = 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - 4\left\|x - \frac{y_n + y_m}{2}\right\|^2.$$

Wegen $\frac{1}{2}(y_n + y_m) \in M$ ist $\|x - \frac{1}{2}(y_n + y_m)\| \geq d$ und daher

$$0 \leq \|y_n - y_m\|^2 \leq 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - 4d^2.$$

Wegen $\|x - y_n\|^2 \to d^2$ erkennen wir, dass (y_n) tatsächlich eine Cauchyfolge ist. Da H vollständig und M abgeschlossen ist, konvergiert (y_n) gegen ein $y \in M$. Für dieses Element y ist

$$\|x - y\| = \lim_{n \to \infty} \|x - y_n\| = d,$$

d.h. y ist ein Element der besten Approximation.

Orthogonalität. Sei $x \in X$, und $z \in M$ sei ein Element mit $\|x - z\| = \inf_{y \in M} \|x - y\|$. Angenommen, es gibt ein $y \in M$, welches nicht orthogonal zu $x - z$ ist, d.h. für das $\langle x - z, y \rangle =: c \neq 0$ ist. Dann ist $y \neq 0$, und wir setzen $y' := z + \frac{c}{\|y\|^2} y \in M$. Für dieses Element gilt

$$\begin{aligned}
\|x - y'\|^2 &= \left\langle x - z - \frac{c}{\|y\|^2} y,\ x - z - \frac{c}{\|y\|^2} y \right\rangle \\
&= \|x - z\|^2 - \frac{c}{\|y\|^2} \underbrace{\langle y, x - z \rangle}_{=\overline{c}} - \frac{\overline{c}}{\|y\|^2} \underbrace{\langle x - z, y \rangle}_{=c} + \frac{|c|^2}{\|y\|^4} \langle y, y \rangle \\
&= \|x - z\|^2 - 2\frac{|c|^2}{\|y\|^2} + \frac{|c|^2}{\|y\|^2} = \|x - z\|^2 - \frac{|c|^2}{\|y\|^2} < \|x - z\|^2
\end{aligned}$$

im Widerspruch zum Minimaleigenschaft. Also ist $x - z \perp M$.

Eindeutigkeit. Seien $z_1, z_2 \in M$ Elemente der besten Approximation für x. Dann ist $z_2 - z_1 \in M$. Andererseits wissen wir aus dem vorigen Schritt, dass $z_2 - z_1 = (x - z_1) - (x - z_2) \perp M$. Insbesondere ist also $z_2 - z_1 \perp z_2 - z_1$, d.h. $\langle z_2 - z_1,\ z_2 - z_1 \rangle = 0$. Dann muss aber $z_1 = z_2$ sein. ■

Anmerkung Das Problem der besten Approximation kann (und muss) man auch in allgemeineren Räumen studieren. Beispielsweise gilt folgendes Resultat:

Satz 2.22 *Sei X ein normierter Raum und M ein* endlich-dimensionaler *Teilraum von X. Dann gibt es für jedes $x \in X$ ein $z \in M$ so, dass $\|x - z\| = \inf_{y \in M} \|x - y\|$.*

($\nearrow$ Übung) Diese Aussage stimmt jedoch i.Allg. nicht mehr, wenn M unendlich-dimensional ist (selbst wenn wir von M Abgeschlossenheit verlangen und von X Vollständigkeit.) Beispielsweise gilt für $X := c_0$ und

$$M := \{(y_n) \in c_0 : \sum_{n=1}^{\infty} y_n 2^{-n} = 0\}$$

diese Aussage nicht. Ist nämlich $x = (x_n) \in c_0$ eine Folge mit $\sum_{n=1}^{\infty} x_n 2^{-n} =: \alpha > 0$, so ist $\|x - y\| > \alpha$ für alle $y \in M$:

$$\begin{aligned} \|x - y\|_\infty &= \max_{k \geq 1} |x_k - y_k| = \sum_{n=1}^{\infty} \max_k |x_k - y_k| 2^{-n} \\ &> \sum_{n=1}^{\infty} |x_n - y_n| 2^{-n} \geq \Big| \sum_{n=1}^{\infty} (x_n - y_n) 2^{-n} \Big| = \alpha \end{aligned}$$

(man beachte, dass $x_n - y_n \to 0$). Andererseits ist aber $\inf_{y \in M} \|x - y\| = \alpha$. Für jedes $\varepsilon > 0$ gibt es nämlich ein $y_\varepsilon \in M$ so, dass $\|x - y_\varepsilon\| < \alpha + \varepsilon$. Ein solches y_ε läßt sich konkret angeben: wir wählen n so, dass $\frac{2^n}{2^n - 1} \alpha < \alpha + \varepsilon$ und setzen

$$y_\varepsilon := -\frac{2^n}{2^n - 1} (\underbrace{\alpha, \dots, \alpha}_{n \text{ mal}}, 0, \dots) + x.$$

Das Element y_ε gehört zu M:

$$\begin{aligned} \sum_{k=1}^{\infty} (y_\varepsilon)_k 2^{-k} &= \sum_{k=1}^{n} \Big(x_k - \frac{2^n}{2^n - 1} \alpha \Big) 2^{-k} + \sum_{k=n+1}^{\infty} x_k 2^{-k} \\ &= \sum_{k=1}^{\infty} x_k 2^{-k} - \frac{2^n}{2^n - 1} \alpha \sum_{k=1}^{n} 2^{-k} \\ &= \alpha - \frac{2^n}{2^n - 1} \alpha \Big(1 - \frac{1}{2^n} \Big) = 0, \end{aligned}$$

und es ist

$$\begin{aligned} \|x - y_\varepsilon\|_\infty &= \max \Big\{ \max_{k \leq n} \Big| x_k + \frac{2^n}{2^n - 1} \alpha - x_k |, \quad \max_{k > n} |x_k - x_k| \Big\} \\ &= \frac{2^n}{2^n - 1} \alpha < \alpha + \varepsilon, \end{aligned}$$

wie behauptet. ∎

Auch wenn ein Element der besten Approximation existiert, wird es i.Allg. nicht eindeutig bestimmt sein. Um die Eindeutigkeit zu garantieren, benötigt man in der Regel stärkere Voraussetzungen an die Norm wie etwa die gleichmäßige Konvexität. Ein normierter Raum heißt *gleichmäßig konvex*, wenn für jedes $\varepsilon > 0$ ein $\delta > 0$ existiert so, dass für alle x, y mit $\|x\| = \|y\| = 1$ und $\|x - y\| \geq \varepsilon$ folgt: $\|\frac{1}{2}(x + y)\| \leq 1 - \delta$. Die Räume $L^p(X)$ und ℓ^p haben für $1 < p < \infty$ diese Eigenschaft. (Details $\nearrow$ Schröder, Funktionalanalysis, S. 19). ■

Zurück zum Hilbertraum. Wir haben in Satz 2.21 gesehen, dass für jeden Hilbertraum H und jeden abgeschlossenen Teilraum M von H für jedes $x \in H$ ein Element $z = P_M x \in H$ der besten Approximation existiert und eindeutig bestimmt ist. Hierdurch wird eine Abbildung $P_M : H \to M$ festgelegt, die so genannte *orthogonale Projektion* von H auf M. Diese ist das Analogon des „Lotfällens“ von einem Punkt auf eine Gerade oder Ebene.

Definition 2.23 *Für jede nichtleere Teilmenge M eine Prähilbertraumes X heißt $M^\perp := \{x \in X : \langle x, y\rangle = 0 \ \forall y \in M\}$ das* orthogonale Komplement *zu M.*

Für beliebiges $\emptyset \neq M \subseteq X$ ist $M^\perp$ ein abgeschlossener linearer Teilraum von X. Die Linearität von $M^\perp$ sieht man wie folgt

$$\begin{aligned} x_1, x_2 \in M^\perp \quad \Rightarrow \quad & \langle \alpha_1 x_1 + \alpha_2 x_2, y\rangle = \alpha_1\langle x_1, y\rangle + \alpha_2\langle x_2, y\rangle = 0 \\ & \text{für alle } \alpha_1, \alpha_2 \in \mathbb{K} \text{ und für alle } y \in M, \end{aligned}$$

und die Abgeschlossenheit folgt aus der Stetigkeit des Skalarproduktes:

$$x_n \in M^\perp,\ x_n \to x \Rightarrow\ 0 = \langle x_n, y\rangle \to \langle x, y\rangle = 0 \quad \text{für alle } y \in M.$$

Beispiel: Für $X = \mathbb{R}^3$ und $M = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_2 = x_3 = 0\}$ ist $M^\perp = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 = 0\}$. In diesem speziellen Fall gilt offenbar $M + M^\perp = \mathbb{R}^3$. ■

Wir werden nun sehen, dass die letzte Aussage dieses Beispiels auch allgemein gilt. Als Schreibweise vereinbaren wir: Sind M, N abgeschlossene lineare Unterräume eines Prähilbertraumes H mit $M \perp N$ und $M + N = H$, so schreiben wir

$$H = M \oplus N \quad \text{(orthogonale Summe)}.$$

Satz 2.24 *Sei H ein Hilbertraum und M ein abgeschlossener nichtleerer linearer Teilraum von H. Dann gilt*

(a) $M \oplus M^\perp = H$, $M \cap M^\perp = \{0\}$,

(b) *P_M ist ein linearer und beschränkter Operator auf H mit $\|P_M\| = 1$ falls $M \neq \{0\}$ (für $M = \{0\}$ ist natürlich $P_M = 0$ und $\|P_M\| = 0$). Außerdem ist $P_M|_M = I|_M$, $P_M|_{M^\perp} = 0|_{M^\perp}$ und $P_M^2 = P_M$.*

Beweis

(a) Jedes $x \in H$ läßt sich schreiben als $x = P_M x + (x - P_M x)$ mit $P_M x \in M$ und $x - P_M x \in M^\perp$ (Satz 2.21). Also ist $H = M + M^\perp$, und die Orthogonalität $M \perp M^\perp$ ist klar. Weiter: da M und $M^\perp$ lineare Räume sind, ist $0 \in M \cap M^\perp$. Weitere Elemente gibt es in $M \cap M^\perp$ nicht, denn aus $x \in M \cap M^\perp$ folgt $\langle x, x \rangle = 0$, also $x = 0$.

(b) Wir zeigen zuerst die Additivität von P_M. Seien $x_1, x_2 \in X$. Dann sind nach Satz 2.21 $x_1 - P_M x_1$ und $x_2 - P_M x_2$ aus $M^\perp$, und für jedes $y \in M$, $y \neq 0$, gilt nach Pythagoras

$$\begin{aligned} \|x_1 + x_2 - (P_M x_1 + P_M x_2) - y\|^2 &= \| \underbrace{x_1 - P_M x_1 + x_2 - P_M x_2}_{\in M^\perp} \underbrace{-y}_{\in M} \|^2 \\ = \quad \|x_1 + x_2 - (P_M x_1 + P_M x_2)\|^2 + \|y\|^2 &> \|x_1 + x_2 - (P_M x_1 + P_M x_2)\|^2. \end{aligned}$$

Also ist $P_M x_1 + P_M x_2 \in M$ das Element der besten Approximation für $x_1 + x_2$. Nach Definition ist daher

$$P_M x_1 + P_M x_2 = P_M(x_1 + x_2).$$

Ähnlich zeigt man $P_M(\alpha x) = \alpha P_M x$ für alle $x \in H$ und $\alpha \in \mathbb{K}$. Also ist P_M ein linearer Operator. Weiter gilt nach Satz 2.21 und Pythagoras für jedes $x \in M$

$$\|P_M x\|^2 \leq \|P_M x\|^2 + \|x - P_M x\|^2 = \| \underbrace{P_M x}_{\in M} + \underbrace{x - P_M x}_{\in M^\perp} \|^2 = \|x\|^2,$$

also $P_M \in L(H)$ und $\|P_M\| \leq 1$. Für jedes $z \in M$ gilt aber $P_M z = z$ und daher $\|P_M z\| = \|z\|$. Wählen wir ein $z \in M$ mit Norm 1 (was im Falle $M \neq \{0\}$ möglich ist), folgt $\|P_M\| = 1$. Die übrigen Aussagen sind klar. ∎

Anmerkung 2.25 Man kann den ersten Teil von Satz 2.24 auch so lesen:
Für jeden abgeschlossenen Teilraum M eines Hilbertraumes H gibt es einen abgeschlossenen Teilraum N von H so, dass

$$M + N = H, \quad M \cap N = \{0\}.$$

(Man braucht ja nur $N = M^\perp$ zu wählen.) Ein solcher Teilraum N heißt auch ein *direktes Komplement* von M. Für Banachräume gilt diese Aussage nicht mehr uneingeschränkt. Sie bleibt aber wenigstens dann noch richtig, wenn M endlichdimensional ist. Darüber hinaus gilt der folgende Satz von Lindenstrauss und Tzafrin: *Hat in einem Banachraum jeder abgeschlossene Teilraum ein direktes Komplement, so ist der Banachraum zu einem Hilbertraum isomorph.*

2.4 Orthonormalsysteme in separablen Hilberträumen

Unser nächstes Ziel ist ein Ersatz für den aus der linearen Algebra bekannten Begriff einer Basis. Dabei betrachten wir nur separable Hilberträume. Allgemein heißt ein metrischer Raum *separabel*, wenn es in ihm eine abzählbare („durchnummerierbare“) dichte Teilmenge gibt.

Beispiele

1. $\mathbb{R}$ ist separabel, da $\mathbb{Q}$ dicht in $\mathbb{R}$ liegt und abzählbar ist. Die Abzählbarkeit von $\mathbb{Q}$ beweist man mit dem *Cantorschen Diagonalverfahren*, welches rechts für die positiven rationalen Zahlen angedeutet ist. Dieses Verfahren hilft auch bei den folgenden Beispielen.

$$
\begin{array}{cccccccc}
\frac{1}{1} & \rightarrow & \frac{2}{1} & & \frac{3}{1} & \rightarrow & \frac{4}{1} & \cdots \\
 & \swarrow & & \nearrow & & \swarrow & & \\
\frac{1}{2} & & \frac{2}{2} & & \frac{3}{2} & & \frac{4}{2} & \cdots \\
\downarrow & \nearrow & & \swarrow & & & & \\
\frac{1}{3} & & \frac{2}{3} & & \frac{3}{3} & & \frac{4}{3} & \cdots \\
 & \swarrow & & & & & & \\
\vdots & & \vdots & & \vdots & & \vdots &
\end{array}
$$

2. $\mathbb{R}^n$ und $\mathbb{C}^n$ sind separabel.

3. Die Räume $(\ell^p, \|\cdot\|_p)$ mit $1 \le p < \infty$ sind separabel. Eine abzählbare und dichte Teilmenge ist jeweils gegeben durch die endlichen Folgen mit rationalen Einträgen. Dagegen ist $(\ell^\infty, \|\cdot\|_\infty)$ nicht separabel. ($\nearrow$ Übung)

4. Die Menge der Polynome mit rationalen Koeffizienten ist abzählbar und liegt dicht in der Menge aller Polynome auf $[0,1]$ mit der $\|\cdot\|_\infty$-Norm.

5. Nach Weierstraß liegen die Polynome dicht in $\big(C[0,1], \|\cdot\|_\infty\big)$. Zusammen mit 4. zeigt dies die Separabilität von $\big(C[0,1], \|\cdot\|_\infty\big)$.

6. Wegen

$$\left(\int_0^1 |f(x) - g(x)|^2 \, dx\right)^{1/2} \le \max_{x \in [0,1]} |f(x) - g(x)| = \|f - g\|_\infty$$

liegen die Polynome mit rationalen Koeffizienten auch dicht in $(C[0,1], \|\cdot\|_2)$. Da $L^2[0,1]$ die Vervollständigung dieses Raumes ist, ist $L^2[0,1]$ separabel. Für $p \in [1, \infty)$ ist auch $(L^p[0,1], \|\cdot\|_p)$ separabel.

Definition 2.26 *Eine Teilmenge S eines Prähilbertraumes H heißt ein* Orthogonalsystem, *wenn $\langle x, y\rangle = 0$ für alle $x, y \in S$ mit $x \neq y$, und S heißt ein* Orthonormalsystem *(ONS), wenn zusätzlich $\langle x, x\rangle = \|x\|^2 = 1$ für alle $x \in S$.*

Lemma 2.27 *Sei S ein ONS in einem Prähilbertraum H. Dann sind die Elemente von S linear unabhängig.*

Beweis Seien $x_1, \ldots, x_n \in S$, $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ und $\alpha_1 x_1 + \ldots + \alpha_n x_n = 0$. Wir multiplizieren diese Gleichung skalar mit x_i $(1 \leq i \leq n)$ und erhalten

$$\alpha_1 \langle x_1, x_i \rangle + \ldots + \alpha_{i-1} \langle x_{i-1}, x_i \rangle + \alpha_i \langle x_i, x_i \rangle + \ldots + \alpha_n \langle x_n, x_i \rangle = \langle 0, x_i \rangle = 0.$$

Diese Gleichung reduziert sich auf $\alpha_i \langle x_i, x_i \rangle = \alpha_i = 0$. Also sind die Vektoren $x_1, \ldots, x_n$ linear unabhängig. ■

Für jede Teilmenge M eines linearen Raumes bezeichnen wir mit $\operatorname{span} M$ die Menge aller (endlichen) Linearkombinationen von Elementen aus M. Die Menge $\operatorname{span} M$ ist also der kleinste lineare Raum, der alle Elemente aus M enthält. Er heißt auch die *lineare Hülle* von M.
Ausgehend von einer beliebigen höchstens abzählbaren linear unabhängigen Teilmenge $M = \{x_n : n \leq N \in \mathbb{N} \cup \{\infty\}\}$ läßt sich ein ONS $S = \{e_n : n \leq N\}$ konstruieren, welches dieselbe lineare Hülle wie M besitzt. Diese Konstruktion heißt

Schmidtsches Orthogonalisierungsverfahren. Da M eine linear unabhängige Menge ist, ist $x_1 \neq 0$, und wir können $e_1 := x_1 / \|x_1\|$ setzen. Dann ist $\|e_1\| = 1$ und $\operatorname{span}\{x_1\} = \operatorname{span}\{e_1\}$.
Wir nehmen nun an, $e_1, \ldots, e_{n-1}$ seien bereits so konstruiert, dass sie ein ONS bilden und die Bedingung

$$\operatorname{span}\{x_1, \ldots, x_{n-1}\} = \operatorname{span}\{e_1, \ldots, e_{n-1}\} \tag{2.19}$$

erfüllen. Für das n-te Element von S definieren wir zunächst

$$f_n := x_n - \sum_{k=1}^{n-1} \langle x_n, e_k \rangle e_k. \tag{2.20}$$

Wegen der linearen Unabhängigkeit von M ist $x_n \notin \operatorname{span}\{x_1, \ldots, x_{n-1}\}$ und damit $x_n \notin \operatorname{span}\{e_1, \ldots, e_{n-1}\}$ wegen (2.19). Folglich ist $f_n \neq 0$, und $e_n := f_n / \|f_n\|$ ist korrekt definiert und hat die Norm 1. Es ist auch klar, dass

$$\operatorname{span}\{x_1, \ldots, x_n\} = \operatorname{span}\{x_1, \ldots, x_{n-1}, f_n\} = \operatorname{span}\{e_1, \ldots, e_{n-1}, e_n\}.$$

Wir überlegen uns noch, dass $\{e_1, \ldots, e_n\}$ ein ONS bildet. Da $\{e_1, \ldots, e_{n-1}\}$ ein ONS bildet, müssen wir nur noch zeigen, dass $\langle e_j, e_n \rangle = 0$ für alle $j \leq n-1$ bzw. $\langle e_j, f_n \rangle = 0$ für alle $j \leq n-1$. Nun ist aber für $j \leq n-1$

$$\begin{aligned} \langle f_n, e_j \rangle &= \left\langle x_n - \sum_{k=1}^{n-1} \langle x_n, e_k \rangle e_k, e_j \right\rangle = \langle x_n, e_j \rangle - \sum_{k=1}^{n-1} \langle \langle x_n, e_k \rangle e_k, e_j \rangle \\ &= \langle x_n, e_j \rangle - \sum_{k=1}^{n-1} \langle x_n, e_k \rangle \langle e_k, e_j \rangle = \langle x_n, e_j \rangle - \langle x_n, e_j \rangle = 0. \end{aligned}$$ ■

Anmerkung Das Element $\sum_{k=1}^{n-1}\langle x_n, e_k\rangle e_k$ in (2.20) ist die Orthoprojektion von x_n auf $\operatorname{span}\{e_1, \ldots, e_{n-1}\}$, d.h. die beste Approximation von x_n durch ein Element aus $\operatorname{span}\{e_1, \ldots, e_{n-1}\}$. ($\nearrow$ Satz 2.35).

Satz 2.28 (a) *Jedes Orthonormalsystem in einem separablen Hilbertraum besitzt höchstens abzählbar viele Elemente.*

(b) *Jeder separable Hilbertraum H besitzt ein höchstens abzählbares ONS S mit* $\operatorname{clos}\operatorname{span} S = H$.

Beweis (a) Sind $x, y \in H$ Elemente mit $\|x\| = \|y\| = 1$ und $\langle x, y\rangle = 0$, so gilt

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

Der Abstand je zweier verschiedener Elemente eines ONS beträgt also $\sqrt{2}$. Ist nun (x_n) eine abzählbare dichte Teilmenge von H, so liegt jedes Element von H in einer der Kugeln $U_{1/2}(x_n)$. Für je zwei Elemente u, v einer solchen Kugel $U_{1/2}(x_n)$ gilt aber

$$\|u - v\| \leq \|u - x_n\| + \|v - x_n\| \leq \frac{1}{2} + \frac{1}{2} = 1 < \sqrt{2}.$$

Jede der Kugeln $U_{1/2}(x_n)$ enthält also höchstens eines der Elemente eines vorgegebenen ONS. Ein ONS in H kann also nicht mehr Elemente enthalten als es Kugeln $U_{1/2}(x_n)$ gibt, also höchstens abzählbar viele.

(b) Sei $(x_n) \subseteq H$ eine abzählbare dichte Teilmenge von H. Dann gilt $\operatorname{clos}\{x_n\} = H$, also erst recht $\operatorname{clos}\operatorname{span}\{x_n\} = H$. Wir wählen

$$y_1 := x_1 \quad \text{falls } x_1 \neq 0 \quad \text{und} \quad y_1 = x_2 \quad \text{sonst.}$$

Weiter sei für jedes $k \geq 2$ y_k das erste Element der Folge (x_n), welches nicht in $\operatorname{span}\{y_1, \ldots, y_{k-1}\}$ liegt. Auf diese Weise erhalten wir eine Teilfolge (y_n) von (x_n), deren Elemente linear unabhängig sind und für die gilt

$$\operatorname{clos}\operatorname{span}\{y_n\} = \operatorname{clos}\operatorname{span}\{x_n\} = H.$$

Außerdem ist klar, dass die Folge (y_n) höchstens abzählbar viele Einträge enthält. Mit dem Schmidtschen Orthogonalisierungsverfahren erzeugen wir aus (y_n) ein höchstens abzählbares ONS (z_n) von H mit

$$\operatorname{clos}\operatorname{span}\{z_n\} = \operatorname{clos}\operatorname{span}\{y_n\} = H. \qquad \blacksquare$$

Definition 2.29 *Ist H separabler Hilbertraum und S ein ONS in H mit* $\operatorname{clos}\operatorname{span} S = H$, *so heißt S* Hilbertbasis *oder* Orthonormalbasis *(ONB) von H oder* vollständiges ONS.

Nach Satz 2.28(b) besitzt jeder separable Hilbertraum eine ONB. Man beachte, dass eine Hilbertbasis *keine* Basis im Sinne der linearen Algebra ist.

Beispiel 1: Die Folgen $e_n := (0, \ldots, 0, 1, 0, \ldots)$ (mit der 1 an der n-ten Stelle) bilden eine Hilbertbasis im ℓ^2. Die Menge $(e_n)_{n \geq 1}$ ist nämlich ein ONS in ℓ^2, und die Menge aller endlichen Linearkombinationen $\sum_{k=1}^n a_k e_k$ liegt dicht in ℓ^2. ■

Beispiel 2: Die Funktionen $e_n(t) := e^{2\pi int} = \cos 2\pi nt + i \sin 2\pi nt$ mit $n \in \mathbb{Z}$ bilden ein vollständiges ONS in $L^2[0,1]$, wie wir aus der Analysis wissen. Dieses ONS bildet die Grundlage der Theorie der klassischen Fourierreihen. ■

Beispiel 3: Wie wir bereits vermerkt haben, liegen die Polynome dicht in $L^2[-1,1]$. Die Menge $S = \{1 = t^0, t^1, t^2, t^3, \ldots\}$ ist also eine linear unabhängige Teilmenge von $L^2[-1,1]$ mit $\operatorname{clos}\operatorname{span} S = L^2[-1,1]$. Anwendung des Schmidtschen Orthogonalisierungsverfahrens auf die Menge S liefert eine Hilbertbasis B von H. Die Funktionen in B haben die Gestalt $\sqrt{\frac{2n+1}{2}} P_n$ $(n = 0, 1, 2, \ldots)$, wobei P_n ein Polynom vom Grad n ist. Die Polynome P_n heißen *Legendre-Polynome*. Sie lassen sich auch aus

$$P_n(t) = \frac{1}{2^n n!} \frac{d^n}{dt^n} \big((t^2 - 1)^n\big)$$

berechnen. ($\nearrow$ Übung) ■

Ist $(e_i)_{i=1}^n$ eine orthonormale Basis in einem endlich-dimensionalen Hilbertraum H, so kann man jedes Element $x \in H$ eindeutig schreiben als $x = \sum_{i=1}^n x_i e_i$, wobei $x_i = \langle x, e_i \rangle$. Wir werden nun sehen, dass eine ähnliche Aussage für unendlich-dimensionale Hilberträume gilt: Ist $(e_i)_{i=1}^\infty$ eine Hilbertbasis von H, so läßt sich jedes Element $x \in H$ eindeutig schreiben als $x = \sum_{i=1}^\infty x_i e_i$ mit $x_i = \langle x, e_i \rangle$. Wir sehen uns zunächst das Konvergenzverhalten solcher Reihen an. Sind die x_i Elemente eines Prähilbertraumes (oder eines beliebigen normierten Raumes) H, so heißt wie üblich die Reihe $\sum_{i=1}^\infty x_i$ *konvergent gegen* $x \in H$, wenn

$$\lim_{n\to\infty} \sum_{i=1}^n x_i = x \quad \text{bzw.} \quad \lim_{n\to\infty} \Big\| \sum_{i=1}^n x_i - x \Big\| = 0.$$

Satz 2.30 *Sei $(e_i)_{i=1}^\infty$ ein ONS in einem Hilbertraum H über $\mathbb{K}$ und seien α_i, $i \in \mathbb{N}$ Zahlen aus $\mathbb{K}$. Dann konvergiert die Reihe $\sum_{i=1}^\infty \alpha_i e_i$ genau dann in H, wenn die Reihe $\sum_{i=1}^\infty |\alpha_i|^2$ in $\mathbb{R}$ konvergiert.*

Beweis Da (e_i) ein ONS ist, gilt nach Pythagoras für beliebige $n, r \in \mathbb{N}$

$$\Big\| \sum_{i=n+1}^{n+r} \alpha_i e_i \Big\|^2 = \sum_{i=n+1}^{n+r} \|\alpha_i e_i\|^2 = \sum_{i=n+1}^{n+r} |\alpha_i|^2. \tag{2.21}$$

Also ist die Folge der Partialsummen der Reihe $\sum \alpha_i e_i$ genau dann eine Cauchyfolge, wenn die Folge der Partialsummen der Reihe $\sum |\alpha_i|^2$ eine Cauchyfolge bildet. Da sowohl H (lt. Voraussetzung) als auch $\mathbb{R}$ (Cauchysches Konvergenzkriterium) vollständige Räume sind, folgt sofort die Behauptung. ■

Folgerung 2.31 *Die Voraussetzungen seien wie in Satz 2.30. Wenn die Reihe* $\sum_{i=1}^{\infty} \alpha_i e_i$ *in* H *konvergiert, so gilt*

$$\Big\| \sum_{i=1}^{\infty} \alpha_i e_i \Big\|_H = \Big(\sum_{i=1}^{\infty} |\alpha_i|^2 \Big)^{1/2} = \|(\alpha_i)\|_{\ell^2}. \tag{2.22}$$

Beweis Aus (2.21) folgt für alle $n \in \mathbb{N}$

$$\Big\| \sum_{i=1}^{n} \alpha_i e_i \Big\|^2 = \sum_{i=1}^{n} |\alpha_i|^2.$$

Der Grenzübergang $n \to \infty$ liefert wegen der Stetigkeit der Norm die Behauptung. ■

Folgerung 2.32 *Die Voraussetzungen seien wie in Satz 2.30. Wenn die Reihe* $\sum_{i=1}^{\infty} \alpha_i e_i$ *in* H *konvergiert, so konvergiert auch jede Umordnung dieser Reihe.*

Beweis Die Reihe $\sum_{i=1}^{\infty} |\alpha_i|^2$ ist absolut konvergent. Also konvergiert auch jede Umordnung dieser Reihe (↗ Vorlesung Analysis). Nach Satz 2.30 konvergiert dann auch die zugehörige Umordnung der Reihe $\sum_{i=1}^{\infty} \alpha_i e_i$. ■

Für das Weitere benötigen wir noch die *Besselsche Ungleichung.*

Satz 2.33 (Besselsche Ungleichung) *Sei* $(e_i)_{i=1}^{\infty}$ *ein ONS in einem Hilbertraum* H *über* $\mathbb{K}$ *und sei* $x \in H$. *Dann gilt*

$$\sum_{i=1}^{\infty} |\langle x, e_i \rangle|^2 \le \|x\|^2. \tag{2.23}$$

Die Konvergenz der Reihe ist Teil der Behauptung.

Beweis Es ist für jedes $k \in \mathbb{N}$

$$\begin{aligned}
0 &\le \Big\| x - \sum_{i=1}^{k} \langle x, e_i \rangle e_i \Big\|^2 = \Big\langle x - \sum_{i=1}^{k} \langle x, e_i \rangle e_i, x - \sum_{i=1}^{k} \langle x, e_i \rangle e_i \Big\rangle \\
&= \|x\|^2 - \sum_{i=1}^{k} \langle x, e_i \rangle \langle e_i, x \rangle - \sum_{i=1}^{k} \overline{\langle x, e_i \rangle} \langle x, e_i \rangle + \Big\langle \sum_{i=1}^{k} \langle x, e_i \rangle e_i, \sum_{j=1}^{k} \langle x, e_j \rangle e_j \Big\rangle \\
&= \|x\|^2 - 2 \sum_{i=1}^{k} |\langle x, e_i \rangle|^2 + \sum_{i=1}^{k} \langle x, e_i \rangle \overline{\langle x, e_i \rangle} \\
&= \|x\|^2 - \sum_{i=1}^{k} |\langle x, e_i \rangle|^2.
\end{aligned} \tag{2.24}$$

Also ist
$$\sum_{i=1}^{k} |\langle x, e_i\rangle|^2 \le \|x\|^2$$
für jedes $k \in \mathbb{N}$, woraus die Konvergenz der Reihe und die Besselsche Ungleichung (2.23) folgen. ■

Insbesondere folgt aus (2.23), dass für jedes $x \in H$ die Reihe $\sum_{i=1}^{\infty} |\langle x, e_i\rangle|^2$ konvergiert. Zusammen mit Satz 2.30 folgt hieraus, dass für jedes $x \in H$ die Reihe $\sum_{i=1}^{\infty} \langle x, e_i\rangle e_i$ in H konvergiert. Wir nennen diese Reihe die *Fourierreihe* und die Zahlen $\langle x, e_i\rangle$ die *Fourierkoeffizienten* von x bezüglich des ONS (e_i). Die Frage ist, unter welchen Voraussetzungen an (e_i) die Fourierreihe von x gegen x konvergiert.

Satz 2.34 *Sei $(e_i)_{i=1}^{\infty}$ ein ONS im Hilbertraum H über $\mathbb{K}$. Die folgenden Aussagen sind äquivalent:*

(a) *Für jedes $x \in H$ konvergiert die Reihe $\sum_{i=1}^{\infty} \langle x, e_i\rangle e_i$ gegen x.*

(b) *Für jedes $x \in H$ gilt die* Parsevalsche Gleichung
$$\sum_{i=1}^{\infty} |\langle x, e_i\rangle|^2 = \|x\|^2.$$

(c) *Das orthogonale Komplement der Menge $(e_i)_{i=1}^{\infty}$ ist $\{0\}$.*

(d) *$(e_i)_{i=1}^{\infty}$ ist ein vollständiges ONS, d.h. eine ONB.*

Beweis (a) $\Rightarrow$ (b) In (2.24) haben wir gesehen, dass
$$\left\| x - \sum_{i=1}^{k} \langle x, e_i\rangle e_i \right\|^2 = \|x\|^2 - \sum_{i=1}^{k} |\langle x, e_i\rangle|^2$$
für jedes ONS gilt. Ist x durch seine Fourierreihe darstellbar, so konvergiert die linke Seite dieser Identität gegen 0 für $k \to \infty$. Dann muß aber auch die rechte Seite für $k \to \infty$ gegen 0 konvergieren.

(b) $\Rightarrow$ (c) Ist x in $(e_i)^{\perp}$, so ist $\langle x, e_i\rangle = 0$ für alle i. Nach der Parsevalschen Gleichung ist $x = 0$.

(c) $\Rightarrow$ (a) Sei $x \in H$ und $z := \sum_{i=1}^{\infty} \langle x, e_i\rangle e_i$. Dann gilt für jedes j
$$\langle z - x, e_j\rangle = \left\langle \sum_{i=1}^{\infty} \langle x, e_i\rangle e_i, e_j \right\rangle - \langle x, e_j\rangle = \langle x, e_j\rangle - \langle x, e_j\rangle = 0,$$
d.h. $z - x \in (e_i)^{\perp}$. Nach Annahme (c) ist $z - x = 0$, also $x = z$.

(a) $\Rightarrow$ (d) Klar.

(d) $\Rightarrow$ (c) Sei $x \in H$ ein Element mit $\langle x, e_i \rangle = 0$ für alle i. Dann steht x auch senkrecht zu $\operatorname{span}(e_i)$ und wegen der Stetigkeit des Skalarproduktes auch zu $\operatorname{clos}\operatorname{span}(e_i) = H$. Es ist also $\langle x, y \rangle = 0$ für alle $y \in H$; insbesondere für $y = x$. Aus $\langle x, x \rangle = 0$ folgt $x = 0$. ■

Wir beenden diesen Abschnitt mit zwei einfachen Konsequenzen der obigen Überlegungen.

Satz 2.35 (Beste Approximation) *Sei (e_i) ein ONS im Hilbertraum H über $\mathbb{K}$ und sei $x \in H$. Dann gilt für alle $c_1, \ldots, c_n \in \mathbb{K}$*

$$\left\| x - \sum_{i=1}^{n} c_i e_i \right\| \geq \left\| x - \sum_{i=1}^{n} \langle x, e_i \rangle e_i \right\|,$$

und das Gleichheitszeichen wird genau dann angenommen, wenn $c_i = \langle x, e_i \rangle$.

Die n-te Partialsumme der Fourierreihe von x liefert also die beste Approximation von x durch Elemente aus $\operatorname{span}\{e_1, \ldots, e_n\}$, d.h. es ist für alle $x \in H$

$$P_{\operatorname{span}\{e_1,\ldots,e_n\}} x = \sum_{i=1}^{n} \langle x, e_i \rangle e_i.$$

Beweis Sei $x_i := \langle x, e_i \rangle$. Dann ist

$$\begin{aligned}
\left\| x - \sum_{i=1}^{n} c_i e_i \right\|^2 &= \left\langle x - \sum_{i=1}^{n} c_i e_i, x - \sum_{i=1}^{n} c_i e_i \right\rangle \\
&= \|x\|^2 - \sum_{i=1}^{n} c_i \langle e_i, x \rangle - \sum_{i=1}^{n} \overline{c_i} \langle x, e_i \rangle + \sum_{i=1}^{n} c_i \overline{c_i} \\
&= \|x\|^2 - \sum_{i=1}^{n} c_i \overline{x_i} - \sum_{i=1}^{n} \overline{c_i} x_i + \sum_{i=1}^{n} c_i \overline{c_i} + \sum_{i=1}^{n} x_i \overline{x_i} - \sum_{i=1}^{n} |x_i|^2 \\
&= \|x\|^2 - \sum_{i=1}^{n} |x_i|^2 + \sum_{i=1}^{n} (x_i - c_i)(\overline{x_i} - \overline{c_i}) \\
&= \|x\|^2 - \sum_{i=1}^{n} |x_i|^2 + \sum_{i=1}^{n} |x_i - c_i|^2 \\
&\overset{(2.24)}{=} \left\| x - \sum_{i=1}^{n} x_i e_i \right\|^2 + \sum_{i=1}^{n} |x_i - c_i|^2.
\end{aligned}$$

Hieraus folgt wegen $\sum_{i=1}^{n} |x_i - c_i|^2 \geq 0$ die Behauptung. ■

Satz 2.36 (Riesz-Fischer) *Jeder separable unendlich-dimensionale Hilbertraum H ist isometrisch isomorph zum Hilbertraum ℓ^2 (über dem gleichen Körper $\mathbb{K}$).*

Beweis Sei $(e_i)_{i=1}^{\infty}$ eine ONB von H (diese existiert nach Satz 2.28 (b)). Jedes Element $x \in H$ läßt sich eindeutig als Fourierreihe $\sum_{i=1}^{\infty} \langle x, e_i \rangle e_i$ darstellen, und aus der Parsevalschen Gleichung folgt, dass die Folge $(\langle x, e_i \rangle)_{i=1}^{\infty}$ der Fourierkoeffizienten von x zum Raum ℓ^2 gehört und dass die Gleichung

$$\|x\|_H = \sqrt{\sum_{i=1}^{\infty} |\langle x, e_i \rangle|^2} = \|(\langle x, e_i \rangle)\|_{\ell^2}$$

besteht. Die Abbildung

$$J : H \to \ell^2, \quad x = \sum_{i=1}^{\infty} \langle x, e_i \rangle e_i \mapsto (\langle x, e_i \rangle)_{i=1}^{\infty} \tag{2.25}$$

ist also eine Isometrie von H *in* ℓ^2. Wir zeigen, dass J den Raum H *auf* ℓ^2 abbildet. Sei dazu (α_i) eine beliebige Folge aus ℓ^2, d.h. die Reihe $\sum |\alpha_i|^2$ soll konvergieren. Nach Satz 2.30 konvergiert dann auch die Reihe $\sum \alpha_i e_i$ in H; ihren Grenzwert bezeichnen wir mit x. Wieder wegen der Stetigkeit des Skalarproduktes gilt

$$\langle x, e_j \rangle = \left\langle \sum_{i=1}^{\infty} \alpha_i e_i, e_j \right\rangle = \sum_{i=1}^{\infty} \alpha_i \langle e_i, e_j \rangle = \alpha_j.$$

Also ist α_j der j-te Fourierkoeffizient von x, d.h. J bildet x auf $(\alpha_j)_{j=1}^{\infty}$ ab.

Es ist klar, dass die in (2.25) definierte Abbildung J die Operationen respektiert, d.h. dass $J(x + y) = Jx + Jy$ und $J(\alpha x) = \alpha Jx$ ist. J ist also nicht nur eine *isometrische*, sondern auch eine lineare Abbildung. ∎

Man kann sich also beim Rechnen in einem separablen unendlich-dimensionalen Hilbertraum immer auf das Rechnen mit Folgen in ℓ^2 zurückziehen. Man beachte aber, dass es keinen *natürlichen* Isomorphismus von H auf ℓ^2 gibt. Die Konstruktion von J setzt nämlich die *Wahl* einer Basis voraus.

Folgerung 2.37 *Jeder separable Hilbertraum über $\mathbb{K}$ ist entweder isometrisch isomorph zu einem der Räume $\mathbb{K}^n$ $(n \in \mathbb{N})$ oder zum Raum ℓ^2 (über dem Körper $\mathbb{K}$).*

Damit haben wir zugleich (bis auf Isometrie) alle möglichen separablen Hilberträume beschrieben.

2.5 Lineare Operatoren auf Hilberträumen

Wir beginnen mit einem Resultat, welches einen Grundpfeiler der Hilbertraumtheorie darstellt. Dazu erinnern wir uns, dass für jeden normierten Raum X über einem Körper $\mathbb{K}$ der Raum $X' := L(X, \mathbb{K})$ der *duale Raum* von X heißt und die Elemente von X' *lineare beschränkte Funktionale* heißen. Ist H ein Hilbertraum, so definiert jedes Element $y \in H$ ein lineares beschränktes Funktional f auf H durch

$$f(x) := \langle x, y \rangle \quad \text{für alle } x \in H. \tag{2.26}$$

Die Linearität von f ist klar, und die Beschränktheit folgt sofort aus der Cauchy-Schwarzschen Ungleichung:

$$|f(x)| = |\langle x, y \rangle| \leq \|x\| \, \|y\|, \quad \text{d.h.} \quad \|f\|_{H'} \leq \|y\|_H.$$

Außerdem ist für $y \neq 0$

$$\|f\| = \sup_{\|x\|=1} |f(x)| \geq \left| f\Big(\frac{y}{\|y\|}\Big) \right| = \frac{|\langle y, y \rangle|}{\|y\|} = \frac{\|y\|^2}{\|y\|} = \|y\|,$$

und die so erhaltene Ungleichung $\|f\| \geq \|y\|$ gilt auch für $y = 0$. Also ist sogar

$$\|f\|_{H'} = \|y\|_H. \tag{2.27}$$

Der folgende Satz zeigt, dass alle Elemente von H' von der Gestalt (2.26) sind.

Satz 2.38 (Riesz'scher Darstellungssatz) *Sei H ein Hilbertraum und $f \in H'$. Dann gibt es genau ein Element $y \in H$ so, dass $f(x) = \langle x, y \rangle$ für alle $x \in H$.*

Beweis Wir zeigen zuerst, dass es für jedes $f \in H'$ ein $y \in H$ mit $f(x) = \langle x, y \rangle$ für alle $x \in H$ gibt. Da f stetig ist, ist der Kern von f, d.h. die Menge $L := \{x \in H : f(x) = 0\}$, ein abgeschlossener Unterraum von H. Ist $L = H$, so ist f das Nullfunktional, und wir können $y = 0$ wählen. Sei also L ein echter Teilraum von H.

Sei z ein Element aus $L^\perp$ mit $f(z) = 1$, welches wir fixieren, und sei x ein beliebiges Element aus H. Für das Element $x - f(x)z$ gilt

$$f(x - f(x)z) = f(x) - f(x)\, f(z) = 0, \quad \text{d.h.} \quad x - f(x)z \in L.$$

Dann ist

$$\langle x - f(x)z, z \rangle = 0 \quad \text{bzw.} \quad \langle x, z \rangle = f(x) \langle z, z \rangle \quad \text{bzw.} \quad f(x) = \langle x, \frac{z}{\|z\|^2} \rangle,$$

d.h. wir können $y = z/\|z\|^2$ wählen.

Um die Eindeutigkeitsaussage zu zeigen, nehmen wir an, es gäbe zwei Elemente $y_1, y_2 \in H$ so, dass $\langle x, y_1 \rangle = \langle x, y_2 \rangle$ für alle $x \in H$. Dann ist $\langle x, y_1 - y_2 \rangle = 0$ für alle $x \in H$, insbesondere für $x := y_1 - y_2$. Hieraus folgt aber $y_1 - y_2 = 0$. ■

Beispielsweise hat jedes lineare beschränkte Funktional f auf ℓ^2 die Gestalt

$$f(x) = f\big((x_i)\big) = \sum_{i=1}^{\infty} x_i \overline{y_i} \quad \text{mit einer fixierten Folge } (y_i) \in \ell^2,$$

und jedes lineare beschränkte Funktional f auf $L^2(X)$ ist von der Form

$$f(x) = \int_X x(t)\,\overline{y(t)}dt \quad \text{mit einer fixierten Funktion } y \in L^2(X).$$

Als erste Konsequenz von Satz 2.38 vermerken wir

Satz 2.39 *Sei H ein Hilbertraum und $T \in L(H)$. Dann existiert genau ein Operator $T^* \in L(H)$ mit*

$$\langle Tx, y\rangle = \langle x, T^*y\rangle \quad \textit{für alle } x, y \in H.$$

Beweis Für jedes feste $y \in H$ wird durch $x \mapsto \langle Tx, y\rangle$ ein lineares beschränktes Funktional f_y definiert: die Linearität ist klar, und die Beschränktheit sieht man so:

$$|f_y(x)| = |\langle Tx, y\rangle| \leq \|Tx\|\,\|y\| \leq \|T\|\,\|x\|\,\|y\|. \tag{2.28}$$

Nach Satz 2.38 gibt es ein eindeutig bestimmtes Element aus H, welches wir mit T^*y bezeichnen, so dass

$$f_y(x) = \langle Tx, y\rangle = \langle x, T^*y\rangle \quad \text{für alle } x \in H.$$

Hierdurch wird eine Abbildung $T^* : H \to H$ definiert. Diese ist linear: Für alle $x, y_1, y_2 \in H$, $\alpha_1, \alpha_2 \in \mathbb{K}$ gilt nämlich

$$\begin{aligned}
&\langle x, T^*(\alpha_1 y_1 + \alpha_2 y_2) - \alpha_1 T^* y_1 - \alpha_2 T^* y_2\rangle \\
&\quad = \langle x, T^*(\alpha_1 y_1 + \alpha_2 y_2)\rangle - \overline{\alpha}_1\langle x, T^* y_1\rangle - \overline{\alpha}_2\langle x, T^* y_2\rangle \\
&\quad = \langle Tx, \alpha_1 y_2 + \alpha_2 y_2\rangle - \langle Tx, \alpha_1 y_1\rangle - \langle Tx, \alpha_2 y_2\rangle = 0.
\end{aligned}$$

Da dies für alle $x \in H$ gilt, ist $T^*(\alpha_1 y_1 + \alpha_2 y_2) - \alpha_1 T^* y_1 - \alpha_2 T^* y_2 = 0$, also

$$T^*(\alpha_1 y_1 + \alpha_2 y_2) = \alpha_1 T^* y_1 + \alpha_2 T^* y_2.$$

Außerdem ist T^* beschränkt. Wegen (2.27) ist nämlich $\|T^*y\|_H = \|f_y\|_{H'}$, und aus (2.28) folgt weiter

$$\|T^*y\|_H = \|f_y\|_{H'} \leq \|T\|_{L(H)}\,\|y\|_H,$$

d.h. insbesondere

$$\|T^*\|_{L(H)} \leq \|T\|_{L(H)}. \tag{2.29}$$

■

Definition 2.40 *Der Operator T^* heißt die* Hilbertraum-Adjungierte *von T.*

Satz 2.41 (a) *Für beliebige Operatoren $S, T \in L(H)$ und Zahlen $\alpha, \beta \in \mathbb{K}$ gilt:*

- $(\alpha S + \beta T)^* = \overline{\alpha} S^* + \overline{\beta} T^*$.
- $(ST)^* = T^* S^*$.
- $(T^*)^* = T$.
- *Ist T stetig invertierbar, so ist auch T^* stetig invertierbar, und es gilt $(T^*)^{-1} = (T^{-1})^*$.*

(b) *Für jeden Operator $T \in L(H)$ gilt*

$$\|T\| = \|T^*\| = \|T^*T\|^{1/2}.$$

Beweis (a) $\nearrow$ Übung.

(b) Sei $x \in H$ mit $\|x\| = 1$. Dann gilt

$$\|Tx\|^2 = \langle Tx, Tx\rangle = \langle T^*Tx, x\rangle \leq \|T^*Tx\| \, \|x\| \leq \|T^*T\|.$$

Gehen wir auf der linken Seite von $\|Tx\|^2 \leq \|T^*T\|$ zum Supremum über alle x mit $\|x\| = 1$ über, so folgt $\|T\|^2 \leq \|T^*T\| \leq \|T^*\| \, \|T\|$. Es ist also $\|T\| \leq \|T^*\|$, und analog (wegen $T^{**} = T$ oder auch wegen (2.29)) $\|T^*\| \leq \|T\|$. Also ist $\|T\| = \|T^*\|$, und nun ist auch klar, dass $\|T\|^2 = \|T^*T\|$. ■

Beispiel 1: Die Adjungierte eines Operators $A \in L(\mathbb{K}^n)$. Sei A bezüglich der Standardbasis von $\mathbb{K}^n$ gegeben durch die Matrix $(a_{ij})_{i,j=1}^n$, und seien $x = (x_1, \ldots, x_n)^T$, $y = (y_1, \ldots, y_n)^T \in \mathbb{K}^n$ (ebenfalls bzgl. dieser Basis). Dann ist

$$\langle Ax, y\rangle = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_j \cdot \overline{y}_i = \sum_{j=1}^n \sum_{i=1}^n x_j \overline{\overline{a_{ij}} y_i} = \langle x, A^* y\rangle,$$

wobei A^* durch die Matrix $\left(\overline{a_{ji}}\right)_{i,j=1}^n$ beschrieben wird (adjungierte Matrix). ■

Beispiel 2: Adjungierte des Verschiebungsoperators. Sei $V : \ell^2 \to \ell^2$, $(x_1, x_2, \ldots) \mapsto (0, x_1, x_2, \ldots)$ der Operator der „Vorwärtsverschiebung“ (forward shift). Für jedes $(y_1, y_2, \ldots) \in \ell^2$ ist

$$\begin{aligned} \langle V(x_1, x_2, \ldots), (y_1, y_2, \ldots)\rangle &= \langle (0, x_1, x_2, \ldots), (y_1, y_2, \ldots)\rangle \\ &= x_1 \overline{y}_2 + x_2 \overline{y}_3 + \ldots = \langle (x_1, x_2, \ldots), (y_2, y_3, \ldots)\rangle. \end{aligned}$$

Also ist

$$V^* : \ell^2 \to \ell^2, \quad (y_1, y_2, y_3, \ldots) \mapsto (y_2, y_3, y_4, \ldots)$$

der Operator der „Rückwärtsverschiebung“ (backward shift). ■

Beispiel 3: Adjungierte von Integraloperatoren. Für $k \in C(X \times X)$ oder $k \in L^2(X \times X)$ ist der Integraloperator

$$(Au)(t) = \int_X k(t,s)u(s)ds$$

beschränkt auf $L^2(X)$. Wir überlegen uns die Beschränktheit unter den Voraussetzungen, dass k messbar auf $X \times X$ ist und dass

$$\int_X |k(t,s)|ds \le c_1, \quad \int_X |k(t,s)|dt \le c_2 \quad \text{fast überall auf } X.$$

Für jedes $u \in L^2(X)$ ist

$$\begin{aligned} |(Au)(t)| &\le \int_X |k(t,s)|\,|u(s)|ds = \int_X |k(t,s)|^{1/2}\,|k(t,s)|^{1/2}\,|u(s)|ds \\ &\le \Big(\int_X |k(t,s)|ds\Big)^{1/2}\Big(\int_X |k(t,s)|\,|u(s)|^2 ds\Big)^{1/2} \quad \text{(Hölder)} \\ &\le c_1^{1/2}\Big(\int_X |k(t,s)|\,|u(s)|^2 ds\Big)^{1/2}, \end{aligned}$$

also

$$\begin{aligned} \|Au\|_2^2 &= \int_X |(Au)(t)|^2 dt \le c_1 \int_X\int_X |k(t,s)|\,|u(s)|^2 ds\,dt \\ &= c_1 \int_X\int_X |k(t,s)|\,|u(s)|^2 dt\,ds = c_1 \int_X |u(s)|^2 \int_X |k(t,s)|dt\,ds \\ &\le c_1c_2 \int_X |u(s)|^2 ds = c_1c_2\|u\|_2^2, \end{aligned}$$

d.h. A ist beschränkt und $\|A\| \le \sqrt{c_1c_2}$. Hier haben wir den Satz von Fubini benutzt, wonach

$$\int_X\int_X f(s,t)\,dsdt = \int_X\int_X f(s,t)\,dtds$$

für jede Funktion $f \in L^1(X \times X)$.

Seien nun $u, v \in L^2(X)$. Wegen

$$
\begin{aligned}
&\Big(\int_{X\times X} |k(t,s)|\,|u(s)|\,|v(t)|\,d(s,t)\Big)^2 \\
&\quad \le \Big(\int_{X\times X} |k(t,s)|\,|u(s)|^2\,d(s,t)\Big)\Big(\int_{X\times X} |k(t,s)|\,|v(t)|^2\,d(s,t)\Big) \\
&\quad = \Big(\int_X |u(s)|^2 \int_X |k(t,s)|\,dt\,ds\Big)\Big(\int_X |v(t)|^2 \int_X |k(t,s)|\,ds\,dt\Big) \\
&\quad \le c_2 \int_X |u(s)|^2\,ds \cdot c_1 \int_X |v(t)|^2\,dt \\
&\quad \le c_1 c_2 \|u\|^2\,\|v\|^2 < \infty.
\end{aligned}
$$

dürfen wir Fubini anwenden und erhalten

$$
\begin{aligned}
\langle Au, v\rangle &= \int_X (Au)(t)\,\overline{v(t)}dt = \int_X \int_X k(t,s)u(s)ds\,\overline{v(t)}dt \\
&= \int_X u(s) \int_X k(t,s)\,\overline{v(t)}dt\,ds = \int_X u(s)\overline{\int_X \overline{k(t,s)}\,v(t)dt\,ds} \\
&= \int_X u(s)\,\overline{(A^*v)(s)}ds = \langle u, A^*v\rangle
\end{aligned}
$$

mit

$$
(A^*v)(s) = \int_X \overline{k(t,s)}\,v(t)dt \quad \text{für } s \in X.
$$

Der adjungierte Operator zu A ist also wieder ein Integraloperator, dessen Kernfunktion k^* durch $k^*(t,s) = \overline{k(s,t)}$ gegeben ist.

Man beachte die Analogie zur Adjungierten $(\overline{a_{ji}})$ einer Matrix (a_{ij}). Diese Analogie zeigt sich auch bei der Produktbildung. Sind A, B Integraloperatoren mit Kernfunktionen a, b, so ist $C = AB$ der Integraloperator mit der Kernfunktion

$$
c(t,s) = \int_X a(t,r)b(r,s)dr.
$$

Es ist nämlich

$$
\begin{aligned}
(ABu)(t) &= \int_X a(t,r)(Bu)(r)dr = \int_X a(t,r) \int_X b(r,s)u(s)ds\,dr \\
&= \int_X \Big(\int_X a(t,r)b(r,s)dr\Big)u(s)ds = \int_X c(t,s)u(s)ds
\end{aligned}
$$

in Analogie zu $(c_{ij}) = (a_{ij})(b_{ij}) = \big(\sum_{r=1}^n a_{ir}b_{rj}\big)_{i,j=1}^n$. ■

Definition 2.42 *Ein Operator $T \in L(H)$ heißt*

- selbstadjungiert, *wenn* $T^* = T$.
- normal, *wenn* $TT^* = T^*T$.
- unitär, *wenn T invertierbar und* $T^{-1} = T^*$.

In Abschnitt 2.3 haben wir Orthoprojektoren *geometrisch* eingeführt. Wir können diese Operatoren nun auch *algebraisch* charakterisieren.

Satz 2.43 *Für $P \in L(H)$ sind folgende Aussagen äquivalent:*

(a) *P ist Orthoprojektor, d.h. $P = P_M$ für einen abgeschlossenen Unterraum M von H.*

(b) *$P^* = P$ und $P = P^2$ (Selbstadjungiertheit und Idempotenz).*

Beweis (a) $\Rightarrow$ (b). Wir wissen aus Satz 2.24, dass $P_M = P_M^2$ und $y - P_M y \perp M$ für alle $y \in H$. Hieraus folgt für beliebige $x, y \in H$

$$\begin{aligned}\langle P_M x, y\rangle &= \langle \underbrace{P_M x}_{\in M}, P_M y + \underbrace{y - P_M y}_{\in M^\perp}\rangle = \langle P_M x, P_M y\rangle \\ &= \langle x + \underbrace{P_M x - x}_{\in M^\perp}, \underbrace{P_M y}_{\in M}\rangle = \langle x, P_M y\rangle,\end{aligned}$$

also $P_M^* = P_M$.

(b) $\Rightarrow$ (a). Sei $M := \operatorname{Im} P = \{x \in H : \exists y \in H \text{ mit } Py = x\}$. M ist ein linearer Teilraum von H. Wir überlegen uns die Abgeschlossenheit von M und zeigen dazu vorab, dass

$$x \in \operatorname{Im} P \quad \Leftrightarrow \quad x = Px. \tag{2.30}$$

Ist $x \in \operatorname{Im} P$, so gibt es ein $y \in H$ mit $Py = x$. Also ist $Px = P^2 y = Py = x$. Umgekehrt folgt aus $x = Px$ natürlich $x \in \operatorname{Im} P$. Dann ist (2.30) gezeigt.

Ist nun (x_n) eine Folge in M mit Grenzwert $x \in H$, so gilt nach (2.30) $x_n = Px_n$ für alle n. Grenzübergang $n \to \infty$ liefert $x = Px$, also $x \in M$ wiederum nach (2.30). Wir zeigen nun, dass für jedes $x \in H$ das Element $Px \in M$ das Element der besten Näherung für x in M ist. Zunächst ist $x - Px \perp M$. Für jedes $y \in H$ ist nämlich

$$\langle x - Px, Py\rangle = \langle Px - P^2 x, y\rangle = \langle Px - Px, y\rangle = 0.$$

Daher gilt für jedes $z \in M$

$$\|x - z\|^2 = \|\underbrace{x - Px}_{\in M^\perp} + \underbrace{Px - z}_{\in M}\|^2 = \|x - Px\|^2 + \|Px - z\|^2,$$

d.h. $\|x - z\|$ wird für $z \in M$ genau dann am kleinsten, wenn $z = Px$. Also ist $P = P_M$. ∎

Wir beenden diesen Abschnitt mit einem Ergebnis, das eine Beziehung herstellt zwischen den Kernen bzw. Bildräumen eine Operators und seiner Adjungierten.

Satz 2.44 *Sei H ein Hilbertraum und $A \in L(H)$. Dann gilt*

$$\operatorname{Ker} A^* = (\operatorname{Im} A)^\perp \quad \textit{und} \quad \overline{\operatorname{Im} A^*} = (\operatorname{Ker} A)^\perp.$$

Beweis Wir überlegen uns die erste Beziehung. Für die zweite $\nearrow$ Übung. Für jedes $x \in H$ gilt

$$\begin{aligned} x \in \operatorname{Ker} A^* \quad &\Leftrightarrow \quad A^*x = 0 \\ &\Leftrightarrow \quad \langle A^*x, y \rangle = 0 \quad \forall y \in H \\ &\Leftrightarrow \quad \langle x, Ay \rangle = 0 \quad \forall y \in H \\ &\Leftrightarrow \quad x \perp \operatorname{Im} A \\ &\Leftrightarrow \quad x \in (\operatorname{Im} A)^\perp. \end{aligned}$$

■

2.6 Kompakte Operatoren

Die kompakten linearen Operatoren bilden eine Teilmenge der linearen beschränkten Operatoren, die in ihren Eigenschaften den aus der linearen Algebra bekannten Abbildungen zwischen endlich-dimensionalen Räumen ($\widehat{=}$ Matrizen) recht nahe kommen. Dies werden wir jedoch erst später deutlich sehen; hier beschränken wir uns auf die Definition, einige grundlegende Eigenschaften sowie Beispiele.

Definition 2.45 *Seien X, Y normierte Räume. Ein linearer Operator $K : X \to Y$ heißt* kompakt, *wenn er beschränkte Mengen (in X) in relativ kompakte Mengen (in Y) abbildet.*

Da eine Teilmenge Z von X genau dann beschränkt ist, wenn sie in einem Vielfachen der Einheitskugel $B = \{x \in X : \|x\| \leq 1\}$ liegt, lassen sich kompakte Operatoren auch so charakterisieren: Ein linearer Operator $K : X \to Y$ ist kompakt genau dann, wenn er die abgeschlossene Einheitskugel von X in eine relativ kompakte Menge (in Y) überführt. Wir bezeichnen die Menge der kompakten Operatoren $K : X \to Y$ mit $K(X, Y)$ und schreiben $K(X)$ für $K(X, X)$. Da relativ kompakte Mengen beschränkt sind, folgt aus der Definition sofort

$$K(X, Y) \subseteq L(X, Y).$$

Im folgenden Satz sind wichtige Eigenschaften kompakter Operatoren zusammengefasst.

Satz 2.46 *Seien X, Y, Z Banachräume. Dann gilt*

(a) *$K(X, Y)$ ist ein abgeschlossener linearer Teilraum von $L(X, Y)$.*

(b) *Ist $K \in K(X, Y)$ und $A \in L(Z, X)$ bzw. $A \in L(Y, Z)$, so ist*

$$KA \in K(Z, Y) \quad \textit{bzw.} \quad AK \in K(X, Z).$$

Das Produkt eines kompakten Operators mit einem beschränkten Operator ist also ein kompakter Operator.

Beweis: Linearität von $K(X, Y)$. Seien $K_1, K_2 \in K(X, Y)$, $\alpha_1, \alpha_2 \in \mathbb{K}$ und $K := \alpha_1 K_1 + \alpha_2 K_2$. Wir müssen zeigen, dass K kompakt ist, d.h. dass sich aus $(Kx_n)_{n\in\mathbb{N}}$ eine konvergente Teilfolge auswählen lässt, wenn die $x_n \in X$ beliebige Elemente mit $\|x_n\| \leq 1$ sind.

Da K_1 kompakt ist, gibt es eine Teilfolge $(x_n)_{n\in\mathbb{N}'}$ von $(x_n)_{n\in\mathbb{N}}$ (mit einer unendlichen Teilmenge $\mathbb{N}' \subseteq \mathbb{N}$) so, dass die Folge $(K_1 x_n)_{n\in\mathbb{N}'}$ konvergiert. Da K_2 kompakt ist, findet man weiter eine Teilfolge $(x_n)_{n\in\mathbb{N}''}$ von $(x_n)_{n\in\mathbb{N}'}$ so, dass auch $(K_2 x_n)_{n\in\mathbb{N}''}$ konvergiert. Dann konvergiert aber auch die Folge $(Kx_n)_{n\in\mathbb{N}''} = \big((\alpha_1 K_1 + \alpha_2 K_2)x_n\big)_{n\in\mathbb{N}''}$.

Abgeschlossenheit von $K(X, Y)$. Sei $(K_n) \subseteq K(X, Y)$ eine Folge von kompakten Operatoren, die gegen einen Operator $K \in L(X, Y)$ konvergiert. Wir haben zu zeigen, dass $K \in K(X, Y)$. Sei $(x_n) \subseteq X$ eine Folge mit $\|x_n\| \leq 1$ für alle $n \in \mathbb{N}$. Da K_1 kompakt ist, finden wir eine Teilfolge $(x_{1n})_{n\in\mathbb{N}}$ von (x_n) so, dass $(K_1 x_{1n})$ konvergiert. Weiter gibt es wegen der Kompaktheit von K_2 eine Teilfolge $(x_{2n})_{n\in\mathbb{N}}$ von (x_{1n}) so, dass $(K_2 x_{2n})$ konvergiert. Wir fahren so fort und erhalten für jedes $k \geq 2$ eine Teilfolge (x_{kn}) von $(x_{k-1,n})$ so, dass $(K_k x_{kn})$ konvergiert. Aus

$$\begin{array}{ccccc}
x_{11} & x_{12} & x_{13} & x_{14} & \cdots \\
x_{21} & x_{22} & x_{23} & x_{24} & \cdots \\
x_{31} & x_{32} & x_{33} & x_{34} & \cdots \\
\vdots & \vdots & \vdots & \vdots &
\end{array}$$

wählen wir die „Diagonalfolge“ (z_n) mit $z_n := x_{nn}$ aus. Da $(z_n)_{n\geq k}$ eine Teilfolge von $(x_{kn})_{n\geq 1}$ ist, konvergiert die Folge $(K_k z_n)_{n\geq 1}$ für jedes k. Dann muss aber auch die Folge (Kz_n) konvergieren: Für jedes $\varepsilon > 0$ gibt es nämlich ein k so, dass $\|K - K_k\| \leq \varepsilon/3$, und es gibt weiter ein N so, dass

$$\|K_k z_n - K_k z_m\| \leq \varepsilon/3 \quad \text{für alle } m, n \geq N.$$

Für alle $n, m \geq N$ gilt daher

$$\begin{aligned}
\|Kz_n - Kz_m\| &\leq \|Kz_n - K_k z_n\| + \|K_k z_n - K_k z_m\| + \|K_k z_m - Kz_m\| \\
&\leq \|K - K_k\|\,\|z_n\| + \|K_k z_n - K_k z_m\| + \|K_k - K\|\,\|z_m\| \leq \varepsilon.
\end{aligned}$$

Also ist (Kz_n) eine Cauchyfolge in Y und mithin konvergent. Es gibt also eine Teilfolge (z_n) von (x_n), so dass (Kz_n) konvergiert.

(b) Seien z.B. $K \in K(X, Y)$ und $A \in L(Y, Z)$, und sei $(x_n) \subseteq X$ eine Folge mit $\|x_n\| \leq 1$ für alle n. Wegen der Kompaktheit von K existiert eine Teilfolge $(x'_n)_{n\in\mathbb{N}}$ von (x_n) so, dass (Kx'_n) konvergiert. Wegen der Stetigkeit von A konvergiert dann auch die Folge (AKx'_n). Also ist AK kompakt. Die Kompaktheit von KA zeigen wir in der Übung. ■

Wir sehen uns nun einige Beispiele für kompakte Operatoren an.

Beispiel 1: Operatoren von endlichem Rang. Wie bei Matrizen bezeichnet man die Dimension des Bildraumes $AX = \operatorname{Im} A$ eines Operators $A : X \to Y$ auch als *Rang* des Operators. Ein Operator $A \in L(X, Y)$ heißt *von endlichem Rang*, wenn $\dim \operatorname{Im} A < \infty$. Die Menge aller Operatoren $A \in L(X, Y)$ von endlichem Rang bezeichnet man oft mit $F(X, Y)$ bzw. $F(X)$ falls $X = Y$.

Lemma 2.47 *Für beliebige Banachräume gilt $F(X, Y) \subseteq K(X, Y)$.*

Beweis Sei $A \in F(X, Y)$. Wegen der Beschränktheit bildet A jede beschränkte Menge $M \subseteq X$ auf eine beschränkte Menge $AM \subseteq \operatorname{Im} A$ ab. Da $\operatorname{Im} A$ endlichdimensional ist, ist die beschränkte Menge AM relativ kompakt. ∎

Einige Banachräume sind dadurch ausgezeichnet, dass sich *jeder* kompakte Operator durch Operatoren von endlichem Rang approximieren läßt. Man sagt auch, dass diese Räume die *Approximationseigenschaft* besitzen. Insbesondere gilt

Satz 2.48 *Ist H ein separabler Hilbertraum, so ist* $\operatorname{clos} F(H) = K(H)$.

Beweis Aus Lemma 2.47 und Satz 2.46 (a) wissen wir, dass $\operatorname{clos} F(H) \subseteq K(H)$. Wir müssen noch zeigen, dass sich jeder kompakte Operator $K \in K(H)$ durch Operatoren von endlichem Rang approximieren läßt. Sei B die Einheitskugel von H. Weiter sei $(e_i)_{i=1}^{\infty}$ eine Hilbertbasis von H und P_n der Orthoprojektor auf $\operatorname{span}\{e_1, \ldots, e_n\}$. Dann ist $K_n := P_n K \in F(H)$, und für jedes $x \in H$ gilt

$$\|Kx - K_n x\| = \|Kx - P_n Kx\| \to 0 \quad \text{für } n \to \infty \tag{2.31}$$

(vgl. Satz 2.34). Weiter: für jedes $\varepsilon > 0$ gibt es ein endliches $\varepsilon/3$-Netz $\{Kx_1, \ldots, Kx_r\}$ in KB, und wegen (2.31) finden wir ein $N = N(\varepsilon)$ so, dass

$$\|Kx_i - K_n x_i\| \le \varepsilon/3 \quad \text{für alle } i \in \{1, \ldots, r\} \text{ und für } n \ge N.$$

Ist nun $x \in B$ beliebig und x_i so gewählt, dass $\|Kx - Kx_i\| \le \varepsilon/3$, so ist für alle $n \ge N$

$$\begin{aligned} \|Kx - K_n x\| &\le \|Kx - Kx_i\| + \|Kx_i - K_n x_i\| + \|K_n x_i - K_n x\| \\ &\le \varepsilon/3 + \varepsilon/3 + \|P_n(Kx_i - Kx)\| \le \varepsilon. \end{aligned}$$

Also ist $\|K - K_n\| \le \varepsilon$ für alle $n \ge N$, woraus die Behauptung folgt. ∎

Man kann sich sogar leicht überlegen, dass *jeder* Hilbertraum die Approximationseigenschaft besitzt: Ist B die Einheitskugel in H, so ist $\overline{K(B)}$ kompakt und daher auch separabel. Dann ist auch $\overline{\operatorname{Im} K}$ ein separabler Hilbertraum, und man kann weiter schließen wie oben.

Auch die uns bereits bekannten Banachräume ℓ^p und L^p mit $1 < p < \infty$ sowie $C(X)$ haben die Approximationseigenschaft (für $C(X)$ siehe z.B. Folgerung 4.1.6 in [Schröder]). Allerdings hat P. Enflo 1973 einen separablen Banachraum X konstruiert, für den $F(X)$ nicht dicht in $K(X)$ liegt. ■

Beispiel 2: Sei $H = \ell^2$ und $(a_n) \in \ell^\infty$. Durch $K(x_n) := (a_n x_n)$ wird ein linearer (klar) und beschränkter Operator $K : \ell^2 \to \ell^2$ definiert:

$$\|K(x_n)\| = \left(\sum_{i=1}^{\infty} |a_i x_i|^2\right)^{1/2} \leq \sup_{i\in\mathbb{N}} |a_i| \left(\sum_{i=1}^{\infty} |x_i|^2\right)^{1/2} = \|(a_n)\|_\infty \, \|(x_n)\|_2.$$

Im Falle $(a_n) \in c_0$ ist dieser Operator kompakt. Mit dem Operator

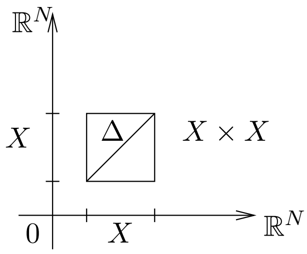
$$P_n : \ell^2 \to \ell^2, \quad (x_i) \mapsto (x_1, x_2, \ldots, x_n, 0, 0, \ldots),$$

ist nämlich

$$\|K(x_i) - P_n K(x_i)\| = \left(\sum_{i\geq n+1} |a_i x_i|^2\right)^{1/2} \leq \sup_{i\geq n+1} |a_i| \cdot \|(x_n)\|_2,$$

und aus $\sup_{i\geq n+1} |a_i| \to 0$ für $n \to \infty$ (da $a_i \to 0$) folgt $\|K - P_n K\| \to 0$. ■

Beispiel 3: Integraloperatoren auf $C(X)$.
Sei $X \subseteq \mathbb{R}^N$ eine kompakte Menge mit nichtleerem Inneren und Jordan-meßbar, und $R > 0$ sei so gewählt, dass $X \subseteq B_R(0) = \{x \in \mathbb{R}^N : \|x\| \leq R\}$. Weiter sei $\Delta = \{(x,x) : x \in X\}$ die „Diagonale“ von $X \times X$. Schließlich sei $k : X \times X \to \mathbb{R}$ eine *schwach singuläre* Kernfunktion, d.h. k sei stetig auf $(X \times X)\backslash\Delta$, und es gebe Konstanten $m \geq 0$ und $\alpha \in [0, N)$ so, dass



$$|k(t,s)| \leq m|t-s|^{-\alpha} \quad \text{für alle } (t,s) \in (X \times X)\backslash\Delta.$$

Wir betrachten den durch

$$(Kf)(t) := \int_X k(t,s) f(s)\, ds\,, \quad t \in X,$$

definierten *schwach singulären Integraloperator* K und wollen dessen Kompaktheit als Operator von $C(X)$ nach $C(X)$ zeigen.
Sei zuerst k stetig auf ganz $X \times X$. In diesem Fall wissen wir bereits, dass $K \in L\big(C(X)\big)$ ist (Beispiel 3 aus 2.1). Also bildet K beschränkte Mengen auf beschränkte Mengen ab. Insbesondere ist das Bild der Einheitskugel beschränkt.

Wir zeigen noch die gleichgradige Stetigkeit dieser Menge. Die Kernfunktion k ist gleichmäßig stetig auf $X \times X$. Für jedes $\varepsilon > 0$ finden wir also ein δ so, dass $|k(t_1, s) - k(t_2, s)| < \varepsilon$ für alle t_1, t_2 mit $\|t_1 - t_2\| < \delta$ und für alle s. Ist nun $\|t_1 - t_2\| < \delta$ und $f \in C(X)$ eine beliebige Funktion mit $\|f\|_\infty \leq 1$, so wird

$$|(Kf)(t_1) - (Kf)(t_2)| \leq \int_X |k(t_1, s) - k(t_2, s)|\, |f(s)|\, ds \leq \varepsilon \cdot \mathrm{Vol}\,(X),$$

d.h. die Einheitskugel von $C(X)$ wird durch K auf eine gleichgradig stetige Menge abgebildet. Nach Arzelà-Ascoli ist K in diesem Fall also kompakt.
Im allgemeinen Fall überlegen wir uns zuerst, dass Kf für jedes $f \in C(X)$ eine beschränkte Funktion ist:

$$\begin{aligned}
\|Kf\|_\infty &= \sup_{t \in X} \left| \int_X k(t, s) f(s)\, ds \right| \leq \sup_{t \in X} \int_X |k(t, s)|\, |f(s)|\, ds \\
&\leq \|f\|_\infty \sup_{t \in X} \int_X |k(t, s)|\, ds \leq m \|f\|_\infty \sup_{t \in X} \int_{B_R(0)} |t - s|^{-\alpha} ds \\
&\leq m \|f\|_\infty \int_{B_{2R}(0)} |s|^{-\alpha} ds.
\end{aligned}$$

Wie wir aus der Analysis wissen, ist dieses Integral endlich.

Nun wählen wir ein $\beta \in (\alpha, N)$ und betrachten die Kernfunktionen

$$k_n(t, s) := \begin{cases} k(t, s) & \text{für } |t - s| > 1/n, \\ n^\beta |t - s|^\beta k(t, s) & \text{für } |t - s| \leq 1/n. \end{cases}$$

Wegen $\beta > \alpha$ sind diese Funktionen auf $X \times X$ stetig. Den durch die Kernfunktion k_n definierten Integraloperator bezeichnen wir mit K_n. Für beliebige Funktionen $f \in C(X)$ hat man dann

$$\begin{aligned}
\|Kf - K_n f\|_\infty &= \sup_{t \in X} \left| \int_X \big(k(t, s) - k_n(t, s)\big) f(s)\, ds \right| \\
&\leq \|f\|_\infty \sup_{t \in X} \int_X |k(t, s) - k_n(t, s)|\, ds \\
&\leq \|f\|_\infty \sup_{t \in X} \int_{s \in X, |t - s| \leq 1/n} m |t - s|^{-\alpha} (1 - n^\beta |t - s|^\beta)\, ds \\
&\leq m \|f\|_\infty \int_{B_{1/n}(0)} |s|^{-\alpha} \big(1 - n^\beta |s|^\beta\big)\, ds.
\end{aligned}$$

Integrale über radialsymmetrische Funktionen lassen sich durch Benutzung von Kugelkoordinaten auswerten. Allgemein gilt

$$\int_{B_R(0)} f(|s|)\, ds = 2\, \frac{\pi^{N/2}}{\Gamma(\frac{N}{2})} \int_0^R r^{N-1} f(r)\, dr$$

(vgl. Fichtenholz III, Nr. 676, Beispiel 12), und wir erhalten damit weiter

$$\begin{aligned}
\|Kf - K_n f\|_\infty &\leq C\|f\|_\infty \int_0^{1/n} r^{N-1-\alpha}(1 - n^\beta r^\beta)\, dr \\
&= C\|f\|_\infty \int_0^{1/n} (r^{N-1-\alpha} - n^\beta r^{N-1-\alpha+\beta})\, dr \\
&= C\|f\|_\infty \left(\frac{r^{N-\alpha}}{N-\alpha} - \frac{n^\beta r^{N-\alpha+\beta}}{N-\alpha+\beta} \right)\Bigg|_0^{1/n} \\
&= C\|f\|_\infty \, n^{\alpha-N} \left(\frac{1}{N-\alpha} - \frac{1}{N-\alpha+\beta} \right). \qquad (2.32)
\end{aligned}$$

Für $n \to \infty$ geht dies wegen $\alpha - N < 0$ gegen 0. Hieraus folgt erstens, dass Kf als gleichmäßiger Grenzwert stetiger Funktionen wieder stetig ist, d.h. K bildet $C(X)$ in $C(X)$ ab, und zweitens, dass

$$\|K - K_n\|_{L(C(X))} \leq C\, n^{\alpha-N} \left(\frac{1}{N-\alpha} - \frac{1}{N-\alpha+\beta} \right) \to 0.$$

Da alle Operatoren K_n kompakt sind, ist K nach Satz 2.46 (a) kompakt. ■

Beispiel 4: Hilbert-Schmidt-Operatoren. Sei X wie im vorigen Beispiel und k eine quadratisch integrierbare Kernfunktion, d.h. $k \in L^2(X \times X)$. Der Integraloperator

$$(Kf)\,(t) := \int_X k(t,s) f(s)\, ds$$

heißt dann auch ein *Hilbert-Schmidt-Operator*. Wir zeigen, dass $K \in L\big(L^2(X)\big)$. Für jede Funktion $f \in L^2(X)$ ist mit der Hölder-Ungleichung

$$\begin{aligned}
|(Kf)(t)| &\leq \int_X |k(t,s)|\, |f(s)|\, ds \\
&\leq \left(\int_X |k(t,s)|^2\, ds \right)^{1/2} \left(\int_X |f(s)|^2\, ds \right)^{1/2},
\end{aligned}$$

also

$$\|Kf\|_{L^2(X)} = \left(\int_X |(Kf)(t)|^2\, dt \right)^{1/2} \leq \left(\int_X \int_X |k(t,s)|^2\, ds\, dt \right)^{1/2} \left(\int_X |f(s)|^2 ds \right)^{1/2}$$

bzw. $\|Kf\|_{L^2(X)} \leq \|k\|_{L^2(X \times X)} \|f\|_{L^2(X)}$. Insbesondere ist K beschränkt und

$$\|K\| \leq \|k\|_{L^2(X \times X)}.$$

Hilbert-Schmidt-Operatoren erweisen sich sogar als kompakt. Um dies zu sehen, wählen wir eine Hilbertbasis $(e_n)_{n \in \mathbb{N}}$ von $L^2(X)$. Weiter sei P_n der Orthoprojektor von $L^2(X)$ auf span $\{e_1, \dots, e_n\}$. Die Operatoren $K_n := P_nK + KP_n - P_nKP_n$ sind dann von endlichem Rang, und die Kompaktheit von K folgt aus Satz 2.46, sobald wir gezeigt haben, dass $\|K - K_n\|_{L(L^2(X))} \to 0$ für $n \to \infty$. Dazu sei $x \in L^2(X)$ eine beliebige Funktion mit Fourierentwicklung $x = \sum_{j\in\mathbb{N}} x_j e_j$ und Norm $\|x\|_{L^2(X)} = 1$. Mit dieser Funktion gilt wegen der Parsevalschen Gleichung

$$\begin{aligned}
\|(K-K_n)x\|^2 &= \sum_{i=1}^{\infty} |\langle (K-K_n)x, e_i\rangle|^2 = \sum_{i=1}^{\infty} \left| \sum_{j=1}^{\infty} x_j \langle (K-K_n)e_j, e_i\rangle \right|^2 \\
&\le \sum_{i=1}^{\infty} \Big(\underbrace{\sum_{j=1}^{\infty} |x_j|^2}_{=1} \sum_{j=1}^{\infty} |\langle (K-K_n)e_j, e_i\rangle|^2 \Big) \quad \text{(Hölder)} \\
&= \sum_{i,j=1}^{\infty} |\langle Ke_j, e_i\rangle - \langle KP_ne_j, e_i\rangle - \langle P_nKe_j, e_i\rangle + \langle P_nKP_ne_j, e_i\rangle|^2.
\end{aligned}$$

Man überlegt sich leicht, dass diese Summe gleich

$$\sum_{i,j\ge n+1} |\langle Ke_j, e_i\rangle|^2$$

ist. Z.B. ist für $i \le n$ und $j \ge n+1$

$$\begin{aligned}
&\langle Ke_j, e_i\rangle - \langle KP_ne_j, e_i\rangle - \langle P_nKe_j, e_i\rangle + \langle P_nKP_ne_j, e_i\rangle = \\
&\quad = \langle Ke_j, e_i\rangle - 0 - \langle Ke_j, P_ne_i\rangle + 0 = 0.
\end{aligned}$$

Um zu zeigen, dass

$$\sum_{i,j\ge n+1} |\langle Ke_j, e_i\rangle|^2 \to 0 \quad \text{für } n \to \infty,$$

müssen wir nachweisen, dass

$$\sum_{i,j\ge 1} |\langle Ke_j, e_i\rangle|^2 < \infty.$$

Nun ist

$$\begin{aligned}
\langle Ke_j, e_i\rangle_{L^2(X)} &= \int_X (Ke_j)(t)\, \overline{e_i(t)}\, dt = \int_X \int_X k(t,s) e_j(s) \overline{e_i(t)}\, ds\, dt \\
&= \iint_{X\times X} k(t,s) f_{ji}(s,t)\, d(s,t)
\end{aligned}$$

mit

$$f_{ji}(s,t) := e_j(s)\, \overline{e_i(t)} \quad \text{für } (s,t) \in X \times X,$$

d.h.

$$\langle Ke_j, e_i\rangle_{L^2(X)} = \langle k, f_{ji}\rangle_{L^2(X\times X)}.$$

Man sieht leicht, dass die Funktionen f_{ji} mit $i, j \geq 1$ ein ONS in $L^2(X \times X)$ bilden. Aus der Besselschen Ungleichung folgt daher

$$\sum_{i,j\geq 1} |\langle Ke_j, e_i\rangle_{L^2(X)}|^2 = \sum_{i,j\geq 1} |\langle k, f_{ji}\rangle_{L^2(X\times X)}|^2 \leq \|k\|^2_{L^2(X\times X)} < \infty$$

und damit die Behauptung. ∎

Wir beenden diesen Abschnitt mit einem Resultat über kompakte Operatoren auf Hilberträumen.

Satz 2.49 *Sei H ein separabler Hilbertraum. Ein Operator $K \in L(H)$ ist genau dann kompakt, wenn der adjungierte Operator K^* kompakt ist.*

Beweis Sei (e_i) eine ONB von H und P_n der Orthoprojektor von H auf $\operatorname{span}\{e_1, \ldots, e_n\}$. Im Beweis von Satz 2.48 haben wir gesehen, dass für kompakte Operatoren K gilt $\|K - P_nK\| \to 0$. Dann gilt aber auch $\|(K - P_nK)^*\| = \|K^* - K^*P_n\| \to 0$, und da die Operatoren K^*P_n von endlichem Rang sind, folgt mit Satz 2.46 die Kompaktheit von K^*. ∎

Dieses Resultat gilt sogar für beliebige Hilberträume.

3 Grundprinzipien der Funktionalanalysis

3.1 Der Satz von Baire und seine Folgerungen

3.1.1 Der Satz von Baire

Alle funktionalanalytischen Resultate dieses Abschnittes beruhen auf folgendem Satz.

Satz 3.1 (Baire) *Sei X ein vollständiger metrischer Raum, und die $A_j \subseteq X$ mit $j \in \mathbb{N}$ seien abgeschlossene Teilmengen von X mit $X = \cup_{j=1}^{\infty} A_j$. Dann hat wenigstens eine der Mengen A_j ein nichtleeres Inneres.*

Beweis Angenommen, keine der Mengen A_j besitzt innere Punkte. Wir wählen $x_0 \in X$ und $\varepsilon_0 > 0$ beliebig. Mit $B_r(x)$ bezeichnen wir die offene Kugel um $x \in X$ mit Radius $r > 0$.
Die Menge $B_{\varepsilon_0/2}(x_0) \cap (X \setminus A_1)$ ist nicht leer (sonst würde ja $B_{\varepsilon_0/2}(x_0)$ vollständig in A_1 liegen, d.h. A_1 hätte innere Punkte) und offen (da $B_{\varepsilon_0/2}(x_0)$ und $X \setminus A_1$ offen sind). Es gibt daher ein $x_1 \in X$ und ein $\varepsilon_1 > 0$ (welches wir außerdem kleiner als 1 wählen können) so, dass

$$B_{\varepsilon_1}(x_1) \subset B_{\varepsilon_0/2}(x_0) \cap (X \setminus A_1).$$

Weiter existiert ebenso ein $x_2 \in X$ und ein $\varepsilon_2 \in (0, 1/2)$ mit

$$B_{\varepsilon_2}(x_2) \subset B_{\varepsilon_1/2}(x_1) \cap (X \setminus A_2).$$

Wir fahren so fort und konstruieren Folgen $(x_j) \subset X$ und $(\varepsilon_j) \subseteq \mathbb{R}$ mit $\varepsilon_j \in (0, 1/j)$ so, dass

$$B_{\varepsilon_j}(x_j) \subset B_{\varepsilon_{j-1}/2}(x_{j-1}) \cap (X \setminus A_j).$$

Die Folge (x_j) ist eine Cauchyfolge. Nach Konstruktion ist nämlich für $k \geq j \geq 1$

$$d(x_k, x_j) < \frac{\varepsilon_j}{2} < \frac{1}{2j}\,.$$

Da X vollständig ist, konvergiert die Folge (x_j), und für den Grenzwert $x \in X$ dieser Folge gilt

$$\lim_{k\to\infty} d(x_k, x_j) = d(x, x_j) \leq \frac{\varepsilon_j}{2} \quad \forall j \in \mathbb{N}.$$

Es ist also

$$x \in \overline{B_{\varepsilon_j/2}(x_j)} \subseteq B_{\varepsilon_j}(x_j) \subset B_{\varepsilon_{j-1}/2}(x_{j-1}) \cap (X \setminus A_j)$$

und daher $x \notin A_j$ für alle j. Dies widerspricht der Voraussetzung $\bigcup_{j=1}^{\infty} A_j = X$. ■

Eine Teilmenge A eines metrischen Raumes heißt

- *nirgends dicht*, wenn $\operatorname{clos} A$ keine inneren Punkte enthält.
- *von 1. Kategorie*, wenn sie eine Vereinigung höchstens abzählbar vieler nirgends dichter Mengen ist.
- *von 2. Kategorie*, wenn sie nicht von erster Kategorie ist.

Mit diesen Begriffen läßt sich Satz 3.1 wie folgt formulieren.

Satz 3.2 (Baire'scher Kategoriensatz) *Jeder vollständige metrische Raum ist von 2. Kategorie.*

Der Baire'sche Kategoriensatz hat bemerkenswerte Konsequenzen. Z.B. ist die Menge aller algebraischen Zahlen (d.h. die Menge aller Nullstellen von Polynomen mit ganzzahligen Koeffizienten) abzählbar und folglich eine Menge von erster Kategorie. Da $\mathbb{R}$ vollständig ist, folgt aus dem Baire'schen Kategoriensatz die Existenz transzendenter Zahlen. Auch die Menge der auf dem Intervall $[0,1]$ stetigen und in wenigstens einem Punkt differenzierbaren Funktionen erweist sich als von erster Kategorie in $C[0,1]$ (diese Menge ist also „sehr klein"). Da $C[0,1]$ vollständig ist, folgt mit Satz 3.2 die Existenz von nirgends differenzierbaren stetigen Funktionen (und es gibt „sehr viele" solcher Funktionen). Weiter ist aus der Analysis bekannt, dass der *punktweise* Grenzwert einer Folge stetiger Funktionen im Allgemeinen nicht stetig ist. Mit Satz 3.2 kann man aber zeigen, dass die Menge der Unstetigkeitsstellen der Grenzfunktion einer punktweise konvergenten Folge in $C[0,1]$ von erster Kategorie ist. Sie hat also „nur wenige" Unstetigkeitsstellen, vgl. [Yosida, S. 14]. Weitere Anwendungen diskutieren wir in den folgenden Abschnitten.

3.1.2 Das Prinzip der gleichmäßigen Beschränktheit

Satz 3.3 (Prinzip der gleichmäßigen Beschränktheit) *Sei X ein Banachraum, Y ein normierter Raum, und $\mathcal{B} \subset L(X,Y)$ sei eine nichtleere Familie von Operatoren mit der Eigenschaft, dass $\sup_{B\in\mathcal{B}} \|Bx\| < \infty$ für jedes $x \in X$. Dann ist $\sup_{B\in\mathcal{B}} \|B\| < \infty$.*

Beweis Für jedes $n \in \mathbb{N}$ sei $A_n := \{x \in X : \sup_{B\in\mathcal{B}} \|Bx\| \le n\}$. Diese Mengen sind abgeschlossen. Es ist nämlich

$$A_n = \bigcap_{B\in\mathcal{B}} \{x \in X : \|Bx\| \le n\},$$

und wegen der Stetigkeit von B ist jede der Mengen $\{x \in X : \|Bx\| \le n\}$ abgeschlossen. Außerdem ist offenbar $\bigcup_{n=1}^{\infty} A_n = X$. Nach Satz 3.1 gibt es ein n so, dass A_n eine komplette abgeschlossene Kugel $\overline{U_\varepsilon(x_0)}$ mit $\varepsilon > 0$ enthält.

Für alle $x \in X$ mit $\|x\| \leq 1$ und alle $B \in \mathcal{B}$ ist dann

$$\|Bx\| = \frac{1}{\varepsilon}\|B(\varepsilon x)\| \leq \frac{1}{\varepsilon}\|B \underbrace{(\varepsilon x + x_0)}_{\in \overline{U_\varepsilon(x_0)} \subset A_n}\| + \frac{1}{\varepsilon}\|B \underbrace{x_0}_{\in A_n}\| \leq \frac{2n}{\varepsilon}.$$

Hieraus folgt $\|B\| \leq 2n/\varepsilon$ für alle $B \in \mathcal{B}$. ■

Das Prinzip der gleichmäßigen Beschränktheit spielt z.B. eine Rolle, wenn man andere Konvergenzarten von Operatoren als die Normkonvergenz betrachten möchte. Stellen wir uns vor, wir möchten eine Operatorgleichung

$$Ax = y, \quad A \in L(X), \quad y \in X \tag{3.1}$$

näherungsweise lösen. In der Regel wird man versuchen, y durch einen endlichen Vektor anzunähern und A durch eine $n \times n$–Matrix zu „approximieren“ und statt (3.1) das lineare Gleichungssystem

$$A_n x_n = y_n, \quad A \in \mathbb{C}^{n \times n}, \quad y \in \mathbb{C}^n$$

betrachten. In welchem Sinne soll die Matrix A_n den Operator A „approximieren“? Die Forderung $\|A_n - A\| \to 0$ ist auf den ersten Blick verlockend; sie engt jedoch die Klasse der möglichen Gleichungen (3.1) zu stark ein. Wir wissen ja bereits, dass die Matrizen A_n Operatoren von endlichem Rang sind, und die einzigen Operatoren, die sich bezüglich der Normkonvergenz durch Operatoren von endlichem Rang approximieren lassen, sind die *kompakten.* Die Forderung $\|A - A_n\| \to 0$ würde also bedeuten, dass wir nur Gleichungen $Ax = y$ mit kompakten Operatoren A näherungsweise lösen könnten. Aus diesem Grund betrachtet man allgemeinere Konvergenzbegriffe.

Definition 3.4 *Seien X, Y normierte Räume und $A_n, A \in L(X,Y)$. Die Folge (A_n) konvergiert* stark *oder* punktweise *gegen A, wenn*

$$\|A_n x - Ax\|_Y \to 0 \quad \textit{für alle } x \in X.$$

Wegen $\|A_n x - Ax\| \leq \|A_n - A\| \, \|x\|$ ist jede norm-konvergente Folge (A_n) auch stark konvergent. Die Umkehrung gilt jedoch nicht.

Beispiel 1: Sei H ein separabler Hilberraum mit ONB $(e_i)_{i=1}^\infty$, und sei P_n der Orthoprojektor von H auf $\operatorname{span}\{e_1, \ldots, e_n\}$. Die Aussagen der Sätze 2.34 und 2.35 können wir dann auch so formulieren:

$$P_n \to I \quad \text{stark für } n \to \infty.$$

Nach Satz 2.35 ist nämlich $P_n x = \sum_{j=1}^n \langle x, e_j \rangle e_j$, und aus Satz 2.34 wissen wir, dass $x = \sum_{j=1}^\infty \langle x, e_j \rangle e_j$.

Die Projektoren P_n konvergieren aber **nicht** in der Norm gegen den identischen Operator I, falls die Dimension von H unendlich ist. Würden sie nämlich in der Norm gegen I konvergieren, so wäre die identische Abbildung I kompakt, da jeder Operator P_n von endlichem Rang ist. Dann wäre aber die Einheitskugel von H relativ kompakt. Wir haben jedoch bereits bemerkt, dass sich aus der Menge $(e_i)_{i\geq 1}$ wegen $\|e_i - e_j\| = \sqrt{2}$ für $i \neq j$ keine normkonvergente Teilfolge auswählen läßt. ∎

Aus Satz 3.3 erhält man leicht, dass man auf die Forderung $A \in L(X,Y)$ in Definition 3.4 verzichten kann, wenn X ein Banachraum ist:

Satz 3.5 *Sei X ein Banachraum, Y ein normierter Raum, und (A_n) sei eine Folge in $L(X,Y)$ mit der Eigenschaft, dass der Grenzwert $Ax := \lim_{n\to\infty} A_n x$ für jedes $x \in X$ existiert. Dann ist $A \in L(X,Y)$, und es gilt $\|A\| \leq \liminf \|A_n\| \leq \sup \|A_n\| < \infty$.*

Beweis Die Linearität von A ist sofort klar. Weiter folgt aus der Konvergenz $A_n x \to Ax$, dass $\sup_n \|A_n x\| < \infty$ für jedes $x \in X$ und somit nach Satz 3.3 $\sup \|A_n\| < \infty$. Damit gilt für jedes $x \in X$

$$\|Ax\| = \lim_{n\to\infty} \|A_n x\| = \liminf_{n\to\infty} \|A_n x\| \leq \liminf_{n\to\infty} \|A_n\|\,\|x\| \leq \sup_{n\to\infty} \|A_n\|\,\|x\|. \qquad ∎$$

Satz 3.6 (Banach/Steinhaus) *Seien X, Y Banachräume, und (A_n) sei eine Folge in $L(X,Y)$. Dann sind folgende Aussagen äquivalent:*

(a) *Es gibt ein $A \in L(X,Y)$ mit $Ax = \lim A_n x$ für alle $x \in X$.*

(b) *Es gilt $\sup \|A_n\| < \infty$, und der Grenzwert $\lim_{n\to\infty} A_n x$ existiert für alle x aus einer dichten Teilmenge X_0 von X.*

Beweis Die Implikation (a) $\Rightarrow$ (b) folgt aus dem Prinzip der gleichmäßigen Beschränktheit. Wir zeigen nun (b) $\Rightarrow$ (a). Sei $M := \sup \|A_n\|$. Zu jedem $x \in X$ und $\varepsilon > 0$ findet man ein $z \in X_0$ so, dass $\|x - z\| < \varepsilon$, und man findet weiter ein N so, dass $\|A_n z - A_m z\| < \varepsilon$ für alle $m, n \geq N$. Für alle $m, n \geq N$ ist dann

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

Also ist für jedes $x \in X$ die Folge $(A_n x)$ eine Cauchyfolge in Y und somit konvergent. Die Behauptung folgt nun sofort aus Satz 3.5. ∎

Beispiel 2: Konvergenz von Quadraturformeln. Für jedes $n \in \mathbb{N}$ seien eine Unterteilung $a \leq t_0^{(n)} < t_1^{(n)} < \ldots < t_{m_n}^{(n)} \leq b$ des Intervalls $[a,b]$ sowie Zahlen $\alpha_k^{(n)} \in \mathbb{R}$, $k = 0, \ldots, m_n$, gegeben. Wir betrachten die Quadraturformel

$$Q_n f := \sum_{k=0}^{m_n} \alpha_k^{(n)} f(t_k^{(n)}),$$

die jeder stetigen Funktion f auf $[a,b]$ eine Zahl Q_nf zuordnet, die wir als Näherung für das Integral $\int_a^b f(x)dx$ betrachten möchten. Definieren wir Q durch $Qf := \int_a^b f(x)dx$, so lautet die Frage, wann $Q_nf \to Qf$ für alle $f \in C[a,b]$ gilt bzw. wann die Funktionale Q_n stark gegen Q konvergieren.

Satz 3.7 *Es gilt genau dann $Q_nf \to Qf = \int_a^b f(x)\,dx$ für alle $f \in C[a,b]$, wenn*

(a) *es ein $M > 0$ gibt mit $\displaystyle\sum_{k=0}^{m_n} |\alpha_k^{(n)}| \le M$ für alle $n \in \mathbb{N}$ und*

(b) *$Q_np \to Qp$ für alle Polynome p auf $[a,b]$.*

Beweis Es ist für jedes $f \in C[a,b]$

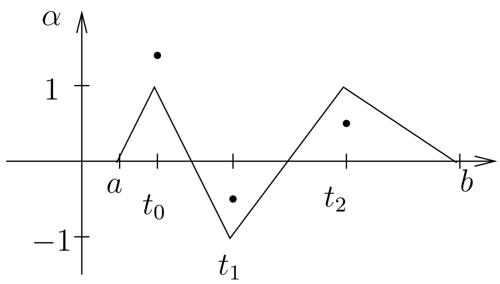
$$|Q_nf| = \left| \sum_{k=0}^{m_n} \alpha_k^{(n)} f(t_k^{(n)}) \right| \le \|f\|_\infty \sum_{k=0}^{m_n} |\alpha_k^{(n)}|$$

und somit $\|Q_n\| \le \sum_{k=0}^{m_n} |\alpha_k^{(n)}|$. In dieser Abschätzung gilt sogar die Gleichheit. Um dies zu sehen, wählen wir für f eine stückweise lineare Funktion mit $\|f\|_\infty = 1$ und

$$f(t_k^{(n)}) = \operatorname{sgn} \alpha_k^{(n)} \quad \forall\, k.$$

Für diese Funktion ist

$$Qf = \sum_{k=0}^{m_n} |\alpha_k^{(n)}|.$$



Die Bedingung in (a) bedeutet also gerade, dass $\sup_n \|Q_n\| \le M < \infty$ ist.
Da die Polynome in $C[a,b]$ dicht liegen (Weierstraß), folgt die Behauptung sofort aus Satz 3.6. ■

3.1.3 Das Prinzip der offenen Abbildung

Als zweite Anwendung des Baireschen Kategoriensatzes erhält man das folgende Resultat. Der dargestellte Beweis stammt von Schauder; vgl. auch [Schröder], S. 71–72.

Satz 3.8 (Prinzip der offenen Abbildung) *Seien X und Y Banachräume und $A \in L(X,Y)$ sei surjektiv. Dann ist für jede in X offene Menge G das Bild AG offen in Y.*

Beweis Für $A, B \subseteq Y$ und $\alpha \in \mathbb{K}$ setzen wir

$$\alpha A := \{\alpha a \in Y : a \in A\} \quad \text{und} \quad A + B := \{a + b \in Y : a \in A,\, b \in B\}.$$

Dann gilt offensichtlich

$$\overline{\alpha A} = \alpha \overline{A} \quad \text{und} \quad \overline{A} + \overline{B} \subseteq \overline{A + B}.$$

Wir setzen zur Abkürzung $U_\varepsilon := B_\varepsilon(0) \subseteq X$ und $V_\varepsilon := B_\varepsilon(0) \subseteq Y$ und zeigen nacheinander:

(a) Zu jedem $\varepsilon > 0$ existiert ein $\delta > 0$ mit $V_\delta \subset \overline{AU_{2\varepsilon}}$.

(b) Zu jedem $\varepsilon > 0$ existiert ein $\delta_1 > 0$ mit $V_{\delta_1} \subseteq AU_\varepsilon$.

(c) AG ist offen in Y für jede offene Menge $G \subseteq X$.

Zu (a). Sei $\varepsilon > 0$. Da $X = \bigcup_{n \geq 1} nU_\varepsilon$ und A surjektiv ist, gilt

$$Y = AX = \bigcup_{n \geq 1} A(nU_\varepsilon) = \bigcup_{n \geq 1} nAU_\varepsilon \subseteq \bigcup_{n \geq 1} \overline{nAU_\varepsilon}.$$

Nach dem Baire'schen Kategoriensatz existieren $n \geq 1$, $y' \in Y$ und $\delta' > 0$ mit

$$y' + V_{\delta'} \subseteq \overline{nAU_\varepsilon} = n\overline{AU_\varepsilon}.$$

Für $y := \frac{1}{n} y'$ und $\delta := \frac{1}{n}\delta'$ ist also

$$y + V_\delta \subseteq \overline{AU_\varepsilon}.$$

Hieraus folgt

$$\begin{aligned} V_\delta \subseteq -y + \overline{AU_\varepsilon} \quad &\subseteq \quad \overline{AU_\varepsilon} + \overline{AU_\varepsilon} \subseteq \overline{AU_\varepsilon + AU_\varepsilon} \\ &= \quad \overline{A(U_\varepsilon + U_\varepsilon)} \subseteq \overline{AU_{2\varepsilon}}. \end{aligned}$$

Zu (b). Sei $\varepsilon > 0$. Wegen (a) kann man $\varepsilon_n > 0$ mit $\sum_{n \geq 1} \varepsilon_n < \varepsilon$ und $\delta_n > 0$ mit $V_{\delta_n} \subset \overline{AU_{\varepsilon_n}}$ mit $\lim_{n \to \infty} \delta_n = 0$ wählen. Wir zeigen, dass δ_1 bereits die verlangten Eigenschaften besitzt. Zu jedem $y \in V_{\delta_1} \subset \overline{AU_{\varepsilon_1}}$ existiert ein $x_1 \in U_{\varepsilon_1}$ mit $\|y - Ax_1\| < \delta_2$, d.h. $y - Ax_1 \in V_{\delta_2} \subset \overline{AU_{\varepsilon_2}}$. Sukzessive finden wir $x_n \in U_{\varepsilon_n}$ mit

$$y - A(x_1 + \ldots + x_n) \in V_{\delta_{n+1}} \subset \overline{AU_{\varepsilon_{n+1}}}.$$

Nun existiert aber der Grenzwert $x := \sum_{n \geq 1} x_n \in X$, da ja

$$\sum_{n \geq 1} \|x_n\| \leq \sum_{n \geq 1} \varepsilon_n < \varepsilon,$$

und es gilt $\|x\| \leq \sum_{n \geq 1} \|x_n\| < \varepsilon$. Außerdem ist $Ax = y$, also $y \in AU_\varepsilon$.

Zu (c). Sei $G \subseteq X$ offen und $y \in AG$. Wir schreiben y als Ax mit $x \in G$ und wählen $\varepsilon > 0$ so, dass $B_\varepsilon(x) \subseteq G$. Nach (b) existiert $\delta > 0$ mit $V_\delta \subseteq AU_\varepsilon$. Damit ist

$$\begin{aligned} B_\delta(y) &= y + V_\delta \subseteq y + AU_\varepsilon = Ax + AU_\varepsilon \\ &= A(x + U_\varepsilon) = A\bigl(B_\varepsilon(x)\bigr) \subseteq AG, \end{aligned}$$

d.h. AG ist offen. ■

Als Folgerung erhält man den bereits angesprochenen Satz von Banach über den inversen Operator.

Satz 3.9 (Banach) *Seien X und Y Banachräume und $A \in L(X, Y)$ ein bijektiver (d.h. invertierbarer) Operator. Dann ist $A^{-1} \in L(Y, X)$, d.h. A ist* stetig *invertierbar.*

Zum Beweis genügt es zu bemerken, dass eine Abbildung $f : M \to N$ zwischen metrischen Räumen genau dann stetig auf ganz M ist, wenn das Urbild jeder in N offenen Menge in M offen ist (↗ Übung). Wegen Satz 3.8 trifft dies gerade auf die Abbildung $A^{-1} : Y \to X$ zu. ■

Folgerung 3.10 *Seien X und Y Banachräume. Die folgenden Bedingungen sind äquivalent für einen Operator $A \in L(X, Y)$:*

(a) *A ist nach unten beschränkt, d.h. es gibt ein $C > 0$ so, dass $\|Ax\| \geq C\|x\|$ für alle $x \in X$.*

(b) $\operatorname{Ker} A = \{0\}$ *und* $\operatorname{Im} A$ *ist abgeschlossen.*

Beweis Die Implikation (a) $\Rightarrow$ (b) haben wir bereits im Beweis von Satz 2.8 gezeigt. Für die umgekehrte Implikation bemerken wir, dass $\operatorname{Im} A$ vollständig ist und dass demzufolge der Operator $A : X \to \operatorname{Im} A$ nach Satz 3.9 stetig invertierbar ist. Sei $B : \operatorname{Im} A \to X$ seine Inverse. Für alle $x \in X$ ist dann

$$\|x\| = \|BAx\| \leq \|B\|\|Ax\|,$$

also $\|Ax\| \geq C\|x\|$ mit $C := \|B\|^{-1}$. ■

Im nächsten Satz erhalten wir eine weitere Charakterisierung der Stetigkeit. Dazu benötigen wir einige Bezeichnungen. Sind X, Y normierte Räume über $\mathbb{K}$, so kann man ihr *Produkt* $X \times Y = \{(x, y) : x \in X, y \in Y\}$ zu einem normierten Raum machen, indem man etwa definiert

$$\|(x, y)\| := \sqrt{\|x\|_X^2 + \|y\|_Y^2} \quad \forall\, (x, y) \in X \times Y. \tag{3.2}$$

Eine Folge in $X \times Y$ konvergiert genau dann bez. dieser Norm, wenn sie komponentenweise konvergiert. Insbesondere gilt also: Sind X, Y Banachräume, so ist auch $X \times Y$ Banachraum. Andere Definitionen wie

$$\|(x, y)\| := \max\{\|x\|, \|y\|\} \quad \text{oder} \quad \|(x, y)\| := \|x\| + \|y\|$$

führen zu äquivalenten Normen auf $X \times Y$. Die Norm (3.2) auf $X \times Y$ ist besonders bequem, wenn X und Y Hilberträume sind. In diesem Fall definiert nämlich

$$\langle (x_1, y_1), (x_2, y_2) \rangle := \langle x_1, x_2 \rangle_X + \langle y_1, y_2 \rangle_Y$$

ein Skalarprodukt auf $X \times Y$, das genau die Norm (3.2) erzeugt.

Satz 3.11 (Satz vom abgeschlossenen Graphen) *Seien X, Y Banachräume und $A : X \to Y$ eine lineare Abbildung. Dann sind folgende Aussagen äquivalent:*

(a) *A ist stetig auf X.*

(b) *Ist $(x_n) \subseteq X$ eine Folge, für die (x_n, Ax_n) gegen $(x, y) \in X \times Y$ konvergiert, so ist $y = Ax$.*

(c) *Der Graph von A, d.h. die Menge $\{(x, Ax) \in X \times Y : x \in X\}$, ist abgeschlossen in $X \times Y$.*

Beweis Die Implikationen (a) $\Rightarrow$ (b) $\Rightarrow$ (c) sind klar. Wir zeigen noch (c) $\Rightarrow$ (a). Sei $G(A) := \{(x, Ax) \in X \times Y : x \in X\}$ ein abgeschlossener linearer Teilraum von $X \times Y$. Dann ist $G(A)$ selbst ein Banachraum. Die Abbildung

$$\pi_1 : G(A) \to X, \quad (x, Ax) \mapsto x$$

ist offenbar linear, bijektiv und stetig. Wegen Satz 3.9 ist π_1^{-1} ebenfalls stetig. Weiter ist die Abbildung

$$\pi_2 : G(A) \to Y, \quad (x, Ax) \mapsto Ax$$

linear und stetig. Dann ist auch die Hintereinanderausführung $\pi_2 \circ \pi_1^{-1}$ stetig. Die Abbildung $\pi_2 \circ \pi_1^{-1}$ ist aber gerade der Operator A:

$$X \ni x \quad \overset{\pi_1^{-1}}{\leadsto} \quad (x, Ax) \quad \overset{\pi_2}{\leadsto} \quad Ax. \qquad \blacksquare$$

Als Anwendung erhalten wir das folgende klassische Resultat.

Satz 3.12 (Hellinger/Toeplitz) *Sei H ein Hilbertraum, $A : H \to H$ ein linearer Operator, und es gelte $\langle Ax, y \rangle = \langle x, Ay \rangle$ für alle $x, y \in H$. Dann ist A stetig.*

Beweis Wir benutzen die Implikation (b) $\Rightarrow$ (a) aus dem Satz vom abgeschlossenen Graphen. Sei $(x_n) \subseteq H$ eine Folge mit $x_n \to x$ und $Ax_n \to y$ in H. Dann gilt für alle $z \in H$

$$\langle z, y \rangle = \lim_{n\to\infty} \langle z, Ax_n \rangle = \lim_{n\to\infty} \langle Az, x_n \rangle = \langle Az, x \rangle = \langle z, Ax \rangle,$$

also $Ax = y$. $\blacksquare$

3.2 Duale Räume und der Satz von Hahn-Banach

3.2.1 Der Fortsetzungssatz von Hahn-Banach

In diesem Abschnitt untersuchen wir den dualen Raum $X' = L(X, \mathbb{K})$ eines Banachraumes X genauer. Ist X ist Hilbertraum, so wissen wir aus dem Darstellungssatz von Riesz, dass es auf X sehr viele lineare beschränkte Funktionale gibt. Im Fall eines beliebigen Banachraumes ist dagegen im Moment nicht einmal klar, ob es überhaupt lineare beschränkte Funktionale (abgesehen von der Abbildung $x \mapsto 0$) gibt. Dies wird durch den Fortsetzungssatz von Hahn/Banach gesichert.

Satz 3.13 (Hahn/Banach) *Sei X ein normierter Raum über dem Körper $\mathbb{K}$ ($= \mathbb{R}$ oder $\mathbb{C}$), $M \subseteq X$ sei ein linearer (nicht notwendig abgeschlossener) Teilraum von X, und $f : M \to \mathbb{K}$ sei ein lineares und beschränktes Funktional auf M, d.h. es ist $|f(x)| \leq C\|x\|$ für alle $x \in M$. Dann gibt es ein lineares beschränktes Funktional $F : X \to \mathbb{K}$ mit*

(i) $F(x) = f(x)$ *für alle* $x \in M$

(ii) $\|F\|_{L(X,\mathbb{K})} = \|f\|_{L(M,\mathbb{K})}$.

Beweisidee Wir wollen den Beweis nur für den Fall erläutern, wo $\mathbb{K} = \mathbb{R}$ und der Raum X *separabel* ist, d.h. wenn es eine abzählbare Teilmenge von X gibt, welche dicht in X ist. Diese Einschränkungen erlauben es, den Beweis im wesentlichen mittels vollständiger Induktion zu führen. Im allgemeinen Fall nichtseparabler Räume benötigt man für den Beweis nichtkonstruktive Hilfsmittel wie das *Zorn'sche Lemma* oder (dazu äquivalent) die Methode der *transfiniten Induktion*. Einen kompletten Beweis findet man z.B. in Schröder, Funktionalanalysis, S. 77–78.

In einem ersten Schritt überlegen wir uns, wie man f von M auf einen Raum fortsetzen kann, der eine „Dimension mehr“ als M besitzt und natürlich so, dass (i) und (ii) entsprechend gelten. Genauer: Sei $x \in X \backslash M$ und

$$M_1 = \operatorname{span}\{M, x\} = \{m + tx : \quad m \in M,\, t \in \mathbb{R}\}.$$

Es ist einfach zu sehen, dass jedes Element von M_1 *eindeutig* in der Form $m + tx$ mit $m \in M$ und $t \in \mathbb{R}$ geschrieben werden kann. Für jede Fortsetzung f_1 von f auf M_1 gilt

$$f_1(m + tx) = f_1(m) + tf_1(x) = f(m) + tc$$

mit einem $c \in \mathbb{R}$, über das wir im Moment nichts wissen. Die Frage ist, ob wir $c = f_1(x)$ so wählen können, dass $\|f_1\|_{L(M_1,\mathbb{R})} = \|f\|_{L(M,\mathbb{R})}$, d.h. dass

$$|f_1(m + tx)| = |f(m) + tc| \leq \|f\| \, \|m + tx\| \quad \forall\, m \in M, t \in \mathbb{R} \tag{3.3}$$

ist. Wenn wir c so wählen können, dass (3.3) gilt, ist $\|f_1\| \leq \|f\|$, und die umgekehrte Ungleichung ist ohnehin klar.

Weiter: es genügt zu zeigen, dass es ein $c \in \mathbb{R}$ so gibt, dass

$$f(m) + tc \leq \|f\| \, \|m + tx\| \quad \forall m \in M, \, t \in \mathbb{R}. \tag{3.4}$$

Aus (3.4) folgt nämlich auch

$$-\big(f(m) + tc\big) = f(-m) + (-t)c \overset{(3.4)}{\leq} \|f\| \, \|-m - tx\| = \|f\| \, \|m + tx\|$$

und damit (3.3) (hier haben wir ausgenutzt, dass $\mathbb{K} = \mathbb{R}$ ist). Schreiben wir (3.4) für $t = 1$ und $t = -1$ auf, so folgt

$$c \leq \|f\| \, \|m + x\| - f(m) \qquad \forall m \in M, \tag{3.5}$$
$$f(m) - \|f\| \, \|m - x\| \leq c \qquad \forall m \in M. \tag{3.6}$$

Es genügt nun zu zeigen, dass man c so wählen kann, dass (3.5) und (3.6) gelten. Aus (3.5) folgt nämlich für alle $t > 0$

$$f(m) + tc = t\Big(f\big(\frac{m}{t}\big) + c\Big) \overset{(3.5)}{\leq} t\|f\| \, \|\frac{m}{t} + x\| = \|f\| \, \|m + tx\|,$$

und ähnlich erhält man aus (3.6) für alle $t < 0$

$$f(m) + tc = -t\Big(f\big(\frac{m}{-t}\big) - c\Big) \overset{(3.6)}{\leq} -t\|f\| \, \|\frac{m}{-t} - x\| = \|f\| \, \|m + tx\|.$$

Damit folgt (3.4) für alle $t \neq 0$, und für $t = 0$ ist (3.4) ohnehin für jedes c richtig.

Es ist nun klar, dass man genau dann eine Zahl c (unabhängig von m) so wählen kann, dass (3.5) und (3.6) gelten, wenn

$$f(m_1) - \|f\| \, \|m_1 - x\| \leq \|f\| \, \|m_2 + x\| - f(m_2) \quad \forall m_1, m_2 \in M.$$

Diese Ungleichung ist äquivalent zur folgenden:

$$f(m_1) + f(m_2) \leq \|f\| \, \|m_1 - x\| + \|f\| \, \|m_2 + x\| \quad \forall m_1, m_2 \in M.$$

Dies ist offenbar richtig, denn für alle $m_1, m_2 \in M$, $x \in X$ gilt

$$f(m_1) + f(m_2) = f(m_1 + m_2) \leq \|f\| \, \|m_1 + m_2\| \leq \|f\| \, (\|m_1 - x\| + \|m_2 + x\|).$$

Also kann man tatsächlich ein $c \in \mathbb{R}$ so wählen, dass

$$f_1(m + tx) = f(m) + tc \quad \text{und} \quad \|f_1\| = \|f\|.$$

Sei nun $\Omega = \{y_1, y_2, \ldots\}$ eine abzählbare dichte Teilmenge von X. Als x_1 wählen wir das y_i mit dem kleinsten i, welches nicht in M liegt. Wenn es kein solches x_1 gibt, so liegt M dicht in X, und wir können sofort zum letzten Beweisschritt übergehen. Wie oben beschrieben setzen wir f von M zu einem Funktional f_1 auf $M_1 := \operatorname{span}\{M, x_1\}$ so fort, dass $\|f_1\| = \|f\|$. Nun wählen wir als x_2 das y_i mit dem kleinsten i, welches nicht in M_1 liegt. Dann können wir f_1 von M_1 auf $M_2 := \operatorname{span}\{M_1, x_2\}$ so fortsetzen, dass $\|f_2\| = \|f_1\| = \|f\|$. Wir fahren auf diese Weise fort und erhalten schließlich ein Funktional F auf $\bigcup_n M_n$, welches auf M_n mit f_n (und damit auf M mit f) übereinstimmt und für das gilt $\|F\| = \|f_n\| = \|f\|$. Die Vereinigung $\bigcup_n M_n$ kann sich dabei über endlich oder abzählbar viele Mengen erstrecken.

Nun ist auf Grund unserer Konstruktion klar, dass $\bigcup_n M_n$ eine *dichte* Teilmenge von X ist. Mit Satz 2.6 können wir F von dieser dichten Teilmenge zu einem Funktional auf ganz $\operatorname{clos}(\bigcup_n M_n) = X$ stetig fortsetzen, wobei die Norm erhalten bleibt. ∎

Folgerung 3.14 *Sei X ein normierter Raum und $x_0 \in X$ mit $x_0 \neq 0$. Dann gibt es ein Funktional $f \in X'$ so, dass $\|f\| = 1$ und $f(x_0) = \|x_0\|$.*

Beweis Sei $M := \operatorname{span}\{x_0\}$. Jedes Element $x \in M$ läßt sich eindeutig schreiben als $x = \alpha x_0$ mit $\alpha \in \mathbb{K}$. Wir definieren f auf M durch $f(x) = \alpha\|x_0\|$. Dann ist $f(x_0) = \|x_0\|$, und für jedes $x = \alpha x_0 \in M$ gilt

$$|f(x)| = |\alpha|\,\|x_0\| = \|\alpha x_0\| = \|x\|,$$

d.h. $\|f\|_{M'} = 1$. Nach Hahn-Banach läßt sich f zu einem Funktional auf ganz X fortsetzen, welches auf M mit f übereinstimmt und gleiche Norm wie f hat. ∎

Insbesondere gilt also für jedes $x \in X$: ist $f(x) = 0$ für alle $f \in X'$, so ist $x = 0$. In diesem Sinne besitzt X genügend viele Elemente.

Folgerung 3.15 *Sei X ein normierter Raum, M ein linearer Teilraum, und sei $y_0 \in X$ so, dass $d := \inf_{x \in M} \|y_0 - x\| > 0$. Dann gibt es ein $f \in X'$ so, dass*

(i) $f(x) = 0$ *für alle $x \in M$,*

(ii) $f(y_0) = 1$,

(iii) $\|f\|_{X'} = 1/d$.

Beweis Sei $M_0 := \operatorname{span}\{M, y_0\}$. Da y_0 nicht in M liegt, läßt sich jedes Element x aus M_0 eindeutig schreiben als $x = m + \alpha y_0$ mit $m \in M$, $\alpha \in \mathbb{K}$. Wir definieren

$$f : M_0 \to \mathbb{K}, \quad x = m + \alpha y_0 \mapsto \alpha.$$

Offenbar ist f ein lineares Funktional, welches (i) und (ii) erfüllt. Weiterhin ist $f : M_0 \to \mathbb{K}$ beschränkt, denn für $x = m + \alpha y_0$ ist im Falle $\alpha \neq 0$

$$|f(x)| = |\alpha| = \frac{|\alpha|\,\|x\|}{\|m + \alpha y_0\|} = \frac{\|x\|}{\|y_0 + \frac{m}{\alpha}\|} \leq \frac{1}{d}\|x\|,$$

und im Falle $\alpha = 0$ gilt die Abschätzung $|f(x)| \leq \frac{1}{d}\|x\|$ offenbar auch. Insbesondere ist also $\|f\|_{M_0'} \leq 1/d$. Wir überlegen uns noch, dass $\|f\|_{M_0'} = 1/d$. Dazu wählen wir eine Folge $(m_n) \subseteq M$ so, dass $\|y_0 - m_n\| \to d$ für $n \to \infty$. Wegen

$$|f(y_0 - m_n)| = |f(y_0) - f(m_n)| = 1$$

gilt dann

$$1 \leq \|f\|\,\|y_0 - m_n\| \leq \frac{1}{d}\,\|y_0 - m_n\| \to 1 \quad \text{für } n \to \infty.$$

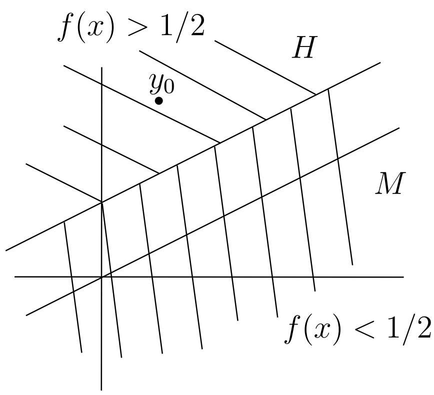
Also ist $\lim_{n\to\infty} \|f\|\,\|y_0 - m_n\| = 1$, und andererseits ist natürlich

$$\lim_{n\to\infty} \|f\|\,\|y_0 - m_n\| = \|f\| \lim_{n\to\infty} \|y_0 - m_n\| = d\|f\|.$$

Hieraus folgt $\|f\| = 1/d$. ■

Als bemerkenswerte Konsequenz erhält man hieraus beispielsweise, dass es (in den Bezeichnungen von Folgerung 3.15) eine Hyperebene H gibt, welche M und y_0 trennt, d.h. so, dass y_0 und M auf verschiedenen Seiten von H liegen. Man kann z.B. wählen

$$H = \{x \in X : \ f(x) = 1/2\}.$$



Es gibt auch weitaus verblüffendere Konsequenzen des Hahn-Banach-Satzes. So ist ja die Abbildung

$$\lim : \quad c \to \mathbb{R}, \quad (x_n) \mapsto \lim x_n$$

ein lineares Funktional auf c, welches darüber hinaus stetig ist:

$$|\lim x_n| \leq \sup |x_n| = \|(x_n)\|_\infty.$$

Nach Hahn-Banach gibt es eine Fortsetzung dieses Funktionals vom Raum c auf den Raum ℓ^∞ aller beschränkten Folgen. Es existiert also eine stetige lineare Abbildung Lim auf ℓ^∞, welche *jeder* beschränkten Folge $(x_n) \in \ell^\infty$ eine reelle Zahl $\operatorname{Lim} x_n$ zuordnet, wobei $\operatorname{Lim} x_n$ der gewöhnliche Limes ist, falls $(x_n) \in c$. Man kann Lim sogar so wählen, dass

$$\liminf x_n \leq \operatorname{Lim} x_n \leq \limsup x_n$$

für alle $(x_n) \in \ell^\infty$ ist. Eine solche Abbildung Lim heißt *Banach-Limes*.

3.2.2 Duale Räume einiger konkreter Banachräume

Für einige konkrete Banachräume lassen sich die dualen Räume explizit beschreiben. Wir sehen uns einige Beispiele an.

Beispiel 1: Hilberträume. Ist H ein Hilbertraum, so wissen wir aus dem Riesz'schen Darstellungssatz, dass es für jedes Funktional $f \in H'$ ein eindeutig bestimmtes $y \in H$ gibt, so dass $f(x) = \langle x, y\rangle$ und $\|f\| = \|y\|$. Umgekehrt erzeugt natürlich jedes Element $y \in H$ vermöge $x \mapsto \langle x, y\rangle$ ein lineares und stetiges Funktional. Wir können also H und H' miteinander in einem gewissen Sinn identifizieren. Genauer: die Abbildung $J : H \to H'$, die jedem $y \in H$ das Funktional $x \mapsto \langle x, y\rangle$ zuordnet, ist eine Isometrie von H auf H'. Diese Isometrie hat aber im Unterschied zu früher betrachteten Isometrien eine besondere Eigenschaft: Sie ist für $\mathbb{K} = \mathbb{C}$ NICHT linear. Wenn nämlich den Elementen y_1, y_2 die Funktionale f_1, f_2 mit

$$f_1(x) = \langle x, y_1\rangle \quad \text{bzw.} \quad f_2(x) = \langle x, y_2\rangle$$

entsprechen, so entspricht der Linearkombination $\alpha_1 y_1 + \alpha_2 y_2$ $(\alpha_1, \alpha_2 \in \mathbb{K})$ das Funktional

$$x \mapsto \langle x, \alpha_1 y_1 + \alpha_2 y_2\rangle = \overline{\alpha}_1\langle x, y_1\rangle + \overline{\alpha}_2\langle x, y_2\rangle = (\overline{\alpha}_1 f_1 + \overline{\alpha}_2 f_2)(x),$$

d.h. es ist

$$J(\alpha_1 y_1 + \alpha_2 y_2) = \overline{\alpha}_1 J y_1 + \overline{\alpha}_2 J y_2.$$

Solche Abbildungen heißen *antilinear*.

Um eine *lineare* Isometrie von H auf H' zu erhalten, kann man wie folgt vorgehen: Wenn H separabel ist, kann man eine ON-Basis $(e_i)_{i=1}^\infty$ von H wählen und jedes $x \in H$ ist eindeutig darstellbar als $x = \sum_{i=1}^\infty x_i e_i$. Dann konvergiert aber auch die Reihe $\sum_{i=1}^\infty \overline{x}_i e_i$ (vgl. Satz 2.30), und wir können einen Operator

$$K : H \to H, \quad x \mapsto \sum_{i=1}^\infty \overline{x}_i e_i$$

definieren. Man macht sich schnell klar, dass dieser Operator ebenfalls eine antilineare Isometrie ist, und zwar von H auf H. Führt man zwei antilineare Operatoren nacheinander aus, erhält man aber einen linearen Operator:

$$(JK)(\alpha_1 x_1 + \alpha_2 x_2) = J(\overline{\alpha}_1 K x_1 + \overline{\alpha}_2 K x_2) = \alpha_1 (JK) x_1 + \alpha_2 (JK) x_2.$$

Daher ist JK eine *lineare* Isometrie von H auf H'. Eine ähnliche Konstruktion gibt es auch für nichtseparable Räume. ■

Beispiel 2: ℓ^p-Räume $(1 < p < \infty)$. Sei $p \in (1, \infty)$, und $q \in (1, \infty)$ sei so gewählt, dass $\frac{1}{p} + \frac{1}{q} = 1$. Wegen der Hölderungleichung

$$\sum_{n=1}^{\infty} |x_n y_n| \leq \|(x_n)\|_p \, \|(y_n)\|_q \tag{3.7}$$

konvergiert für beliebige Folgen $x = (x_n) \in \ell^p$ und $y = (y_n) \in \ell^q$ die Reihe $\sum_{n=1}^{\infty} x_n y_n$. Jede Folge $y \in \ell^q$ definiert also ein lineares Funktional $\Phi(y)$ auf ℓ^p durch

$$\Phi(y)(x) := \sum_{n=1}^{\infty} x_n y_n \ , \quad x = (x_n) \in \ell^p.$$

Aus (3.7) folgt, dass $\Phi(y)$ stetig ist und dass für alle $y \in \ell^q$ gilt:

$$\|\Phi(y)\|_{(\ell^p)'} \leq \|y\|_q. \tag{3.8}$$

Wir zeigen, dass in (3.8) sogar stets die Gleichheit gilt, d.h. dass $\Phi : \ell^q \to (\ell^p)'$ eine Isometrie ist. Zu einem gegebenen $y = (y_k) \in \ell^q$ mit $y_k = |y_k| e^{i\alpha_k}$ definieren wir dazu eine Folge $x = (x_k)$ durch $x_k := |y_k|^{q-1} e^{-i\alpha_k}$. Diese Folge gehört zu ℓ^p:

$$\sum_{k=1}^{\infty} |x_k|^p = \sum_{k=1}^{\infty} |y_k|^{p(q-1)} = \sum_{k=1}^{\infty} |y_k|^q < \infty, \tag{3.9}$$

und es gilt für dieses spezielle x

$$\begin{aligned} \Phi(y)(x) &= \sum_{k=1}^{\infty} x_k y_k = \sum_{k=1}^{\infty} |y_k|^{q-1} |y_k| = \sum_{k=1}^{\infty} |y_k|^q = \|y\|_q^q \\ &= \|y\|_q \, \|y\|_q^{q-1} \overset{(3.9)}{=} \|y\|_q \, \|x\|_p^{p(q-1)/q} = \|y\|_q \, \|x\|_p. \end{aligned}$$

Also ist tatsächlich $\|\Phi(y)\|_{(\ell^p)'} = \|y\|_q$.

Nun überlegen wir uns noch, dass Φ den Raum ℓ^q *auf* $(\ell^p)'$ abbildet. Sei $\varphi \in (\ell^p)'$. Mit e_k bezeichnen wir die spezielle Folge $(0, 0, \ldots, 0, 1, 0, \ldots) \in \ell^p$ mit der 1 an der k-ten Stelle, und wir definieren $y = (y_k)$ durch $y_k := \varphi(e_k)$. Wir zeigen, dass $y \in \ell^q$ und dass $\varphi(x) = F(y)(x)$ für alle $x \in \ell^p$.

Für die erste dieser Aussagen wählen wir zu $y = (y_k)$ die *spezielle* Folge $x = (x_k)$ wie oben und schreiben $x^{(n)}$ für $(x_1, \ldots, x_n, 0, 0, \ldots)$. Für jedes n ist dann

$$\begin{aligned} \|x^{(n)}\|_p^p &= \sum_{k=1}^{n} x_k y_k = \sum_{k=1}^{n} x_k \varphi(e_k) \\ &= \varphi\left(\sum_{k=1}^{n} x_k e_k \right) = \varphi(x^{(n)}) \leq \|\varphi\| \, \|x^{(n)}\|_p \end{aligned}$$

woraus folgt $\|x^{(n)}\|_p^{p-1} \leq \|\varphi\|$ bzw. $\|x^{(n)}\|_p \leq \|\varphi\|^{1/(p-1)}$. Durch Grenzübergang $n \to \infty$ erhalten wir $\|x\|_p \leq \|\varphi\|^{1/(p-1)}$, d.h. es ist $x \in \ell^p$. Wie wir oben gesehen haben, folgt dann auch $y \in \ell^q$.

Sei nun $x = (x_n) \in \ell^p$ *beliebig*. Wegen der Stetigkeit von φ gilt dann

$$\varphi(x) = \varphi\left(\lim_{n\to\infty} \sum_{k=1}^{n} x_k e_k\right) = \lim_{n\to\infty} \sum_{k=1}^{n} x_k \varphi(e_k) = \sum_{k=1}^{\infty} x_k y_k.$$

Fazit: Φ ist lineare Isometrie von ℓ^q auf $(\ell^p)'$. In diesem Sinne identifiziert man die Räume $(\ell^p)'$ und ℓ^q und schreibt einfach $(\ell^p)' = \ell^q$. ■

Beispiel 3: L^p-Räume $(1 < p < \infty)$. Sei wieder q durch $1/p + 1/q = 1$ definiert. Ähnlich wie in Beispiel 2 gilt dann: Für jede Funktion $g \in L^q(X)$ wird durch

$$\Phi(g)(f) := \int_X f(x)g(x)\,dx$$

ein lineares und beschränktes Funktional $\Phi(g)$ auf $L^p(X)$ definiert, und die hierdurch erklärte Abbildung Φ ist eine lineare Isometrie von L^q auf $(L^p)'$. In diesem Sinne ist also $(L^p)' = L^q$. Der Beweis der Surjektivität von Φ bereitet hier aber wesentlich mehr Mühe als in Beispiel 2. ■

Beispiel 4: $C[a,b]$. Die Beschreibung des Dualraumes zu $C(X)$ ist - selbst im Fall $X = [a,b]$ - aufwendiger als in den übrigen Beispielen. Wir geben hier nur die nötigen Definitionen an, um das Hauptresultat - den Rieszschen Darstellungssatz - verstehen zu können und verweisen auf die Literatur (z.B. Schröder, Abschnitt 2.4) für ausführlichere Herleitungen.

Sei $f : [a,b] \to \mathbb{R}$ und $Z = \{t_i\}_{i=0}^n$ mit $a = t_0 < t_1 < \ldots < t_n = b$ eine *Zerlegung* von $[a,b]$. Die *Variation* von f bezüglich Z ist die Zahl

$$V(f,Z) := \sum_{i=1}^{n} |f(t_i) - f(t_{i-1})|.$$

Ist die Zahl

$$V(f) := \sup\,\{V(f,Z) : Z \text{ ist Zerlegung von } [a,b]\}$$

endlich, so heißt f eine *Funktion von beschränkter Variation*, und die Zahl $V(f)$ heißt die *Totalvariation* von f. Die Menge aller Funktionen von beschränkter Variation auf $[a,b]$ bezeichnen wir mit $BV[a,b]$. Zu $BV[a,b]$ gehören z.B. alle monotonen und alle Lipschitzstetigen Funktionen, jedoch nicht alle stetigen Funktionen. Die Menge $BV[a,b]$ wird ein Banachraum, wenn man sie mit folgender Norm versieht:

$$\|f\|_{BV} := |f(a)| + V(f).$$

Insbesondere ist die Menge aller Funktionen $f \in BV[a,b]$ mit $f(a) = 0$ ein linearer abgeschlossener Teilraum von $BV[a,b]$, den wir mit $BV_0[a,b]$ bezeichnen. Für Funktionen $f \in BV_0[a,b]$ gilt also $\|f\|_{BV} = V(f)$.
Ist f eine *stetige* Funktion auf $[a,b]$, $Z = \{t_i\}_{i=0}^n$ eine Zerlegung von $[a,b]$ und $\zeta = (\zeta_1, \ldots, \zeta_n)$ ein dazu gehörender Zwischenvektor (d.h. $t_{i-1} \leq \zeta_i \leq t_i$), so heißt

$$R(f, Z, \zeta) := \sum_{i=1}^n f(\zeta_i)(t_i - t_{i-1})$$

eine *Riemann-Summe* für f. Bekanntlich gilt für jede Folge (Z_n) von Zerlegungen, deren „Maschenweite" $|Z_n| := \max |t_i - t_{i-1}|$ gegen 0 geht, und für jede Folge zugehöriger Zwischenvektoren ζ_n, dass die Folge $\big(R(f, Z_n, \zeta_n)\big)_{n \geq 1}$ konvergiert und dass alle diese Folgen den gleichen Grenzwert besitzen. Dieser Grenzwert ist das Riemann-Integral

$$\int_a^b f(s)\,ds.$$

In Verallgemeinerung dieses Begriffes definiert man für jedes $f \in C[a,b]$ und jedes $g \in BV[a,b]$ sowie für jede Zerlegung Z mit Zwischenvektor ζ die *Riemann-Stieltjes-Summe*

$$R(f, g, Z, \zeta) := \sum_{i=1}^n f(\zeta_i)\big(g(t_i) - g(t_{i-1})\big)$$

(für $g(t) = t$ erhalten wir gerade die übliche Riemann-Summe). Wie oben gilt: Für jede Folge (Z_n) von Zerlegungen mit $|Z_n| \to 0$ und jede Folge (ζ_n) zugehöriger Zwischenvektoren konvergiert die Folge der RS-Summen $\big(R(f, g, Z_n, \zeta_n)\big)$. Weiter besitzen alle diese Folgen den gleichen Grenzwert. Dieser heißt das *Riemann-Stieltjes-Integral von f bezüglich g* und wird mit $\int_a^b f(s)d\,g(s)$ bezeichnet.
Schließlich bezeichnen wir noch mit $BV_*[a,b]$ den Teilraum von $BV_0[a,b]$, der aus allen Funktionen besteht, die in jedem Punkt aus (a,b) von rechts stetig sind. Dann gilt

Satz 3.16 (Riesz'scher Darstellungssatz) *Jede Funktion $g \in BV_*[a,b]$ erzeugt durch*

$$\Phi(g)(f) := \int_a^b f(s)d\,g(s)$$

ein lineares stetiges Funktional $\Phi(g)$ auf $C[a,b]$. Die Abbildung Φ ist eine lineare Isometrie von $BV_*[a,b]$ auf $C[a,b]'$. In diesem Sinne ist also

$$C[a,b]' = BV_*[a,b].$$

Die Voraussetzung der Stetigkeit von rechts haben wir benötigt, um eine *eindeutige* Zuordnung zwischen Funktionen von beschränkter Variation und Funktionalen auf C zu garantieren. Man beachte, dass z.B. die Funktionen $g(s) = 0$ für alle $s \in [a,b]$ und $h(s) = 0$ für $s \in [a,b] \setminus \{\frac{a+b}{2}\}$, $h(\frac{a+b}{2}) = 1$, beide das Nullfunktional auf C erzeugen.

3.2.3 Bidual, reflexive Räume

Als *Bidual* eines normierten Raumes X bezeichnet man den normierten Raum $X'' = (X')'$. Man kann jeden normierten Raum X als einen Unterraum seines bidualen Raumes X'' auffassen. Dazu sei $x \in X$ und $f \in X'$. Bisher haben wir in $f(x) \in \mathbb{K}$ immer das Funktional f als fest und x als variabel betrachtet, d.h. wir haben uns für die Abbildung

$$X \to \mathbb{K}, \quad x \mapsto f(x)$$

interessiert. Nun wollen wir $x \in X$ fixieren und das Funktional $f \in X'$ in $f(x)$ als variabel betrachten. Für jedes feste x betrachten wir also die Abbildung

$$X' \to \mathbb{K}, \quad f \mapsto f(x),$$

die wir mit $F(x)$ bezeichnen. Es ist also $F(x)(f) = f(x)$. Wegen

$$\begin{aligned} F(x)(\alpha_1 f_1 + \alpha_2 f_2) &= (\alpha_1 f_1 + \alpha_2 f_2)(x) = \alpha_1 f_1(x) + \alpha_2 f_2(x) \\ &= \alpha_1 F(x)(f_1) + \alpha_2 F(x)(f_2) \end{aligned}$$

ist $F(x)$ ein lineares Funktional auf X', und aus

$$|F(x)(f)| = |f(x)| \leq \|f\|_{X'} \, \|x\|_X$$

folgt die Stetigkeit dieses Funktionals sowie die Normabschätzung $\|F(x)\|_{X''} \leq \|x\|_X$. Es gilt sogar die Gleichheit in dieser Abschätzung. Wie wir aus Folgerung 3.14 aus dem Satz von Hahn-Banach wissen, gibt es nämlich zu jedem $x \in X$ ein $f_x \in X'$, so dass $\|f_x\| = 1$ und $f_x(x) = \|x\|$. Für dieses konkrete Funktional ist

$$|F(x)(f_x)| = |f_x(x)| = \|x\|$$

und daher $\|F(x)\|_{X''} = \sup_{\|f\|=1} |F(x)(f)| \geq \|x\|$. Also ist $\|F(x)\| = \|x\|$. Die Abbildung $F : X \to X''$, die jedem $x \in X$ das Funktional $F(x)$ auf X' zuordnet, ist also eine Isometrie. Schließlich überlegen wir uns, dass F eine lineare Abbildung ist. Seien $x_1, x_2 \in X$, $\alpha_1, \alpha_2 \in \mathbb{K}$ und $f \in X'$ beliebig. Dann ist

$$\begin{aligned} F(\alpha_1 x_1 + \alpha_2 x_2)(f) &= f(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 f(x_1) + \alpha_2 f(x_2) \\ &= \alpha_1 F(x_1)(f) + \alpha_2 F(x_2)(f). \end{aligned}$$

Da dies für beliebiges $f \in X'$ gilt, folgt

$$F(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 F(x_1) + \alpha_2 F(x_2).$$

Wir fassen zusammen.

Satz 3.17 *Für jeden normierten Raum X ist die durch*

$$F(x) : X' \to \mathbb{K}, \quad f \mapsto f(x)$$

definierte Abbildung $F : X \to X''$ eine lineare Isometrie von X in X''.

Mit anderen Worten: Jeder normierte Raum ist auf natürliche Weise in seinen Bidual eingebettet. Da X'' als dualer Raum von X' vollständig ist, erhalten wir mit Satz 3.17 auch einen neuen Beweis der Aussage, dass sich normierte Räume vervollständigen lassen. Die Vervollständigung von X ist einfach der Abschluss von $F(X)$ in X''.

Definition 3.18 *Ein normierter Raum X heißt* reflexiv*, wenn die in Satz 3.17 beschriebene Abbildung F surjektiv, d.h. eine lineare Isometrie von X* auf *X'' ist.*

Man beachte, dass wir nicht nur schlechthin verlangen, dass es eine lineare Isometrie von X auf X'' *gibt*, sondern wir verlangen, dass eine *ganz bestimmte* Abbildung (nämlich F) eine lineare und surjektive Isometrie ist. In der Tat gibt es Beispiele von Banachräumen, die linear isometrisch zu ihrem Bidual, aber NICHT reflexiv sind.

Beispiele für reflexive und nicht reflexive Räume. Von den uns bekannten Räumen sind die ℓ^p– und $L^p(X)$–Räume für $1 < p < \infty$ reflexiv. Für $p = 1$ oder $p = \infty$ sind diese Räume nicht reflexiv. Ebenso sind $C[a, b]$ und auch $BV_*[a, b]$ nicht reflexiv. Etwas genauer überlegen wir uns, dass Hilberträume reflexiv sind.

Satz 3.19 *Hilberträume sind reflexiv.*

Beweis Wir erinnern daran, dass für jeden Hilbertraum H die Abbildung $J : H \to H'$, die jedem $y \in H$ das Funktional $x \mapsto \langle x, y\rangle$ zuordnet, eine antilineare Isometrie von H auf H' ist.

Als erstes überlegen wir uns, dass man H' zu einem Hilbertraum machen kann, indem ein Skalarprodukt $\langle\cdot,\cdot\rangle_{H'}$ definiert wird durch

$$\langle f, g\rangle_{H'} := \langle J^{-1}g, J^{-1}f\rangle_H.$$

Wir überprüfen nicht alle Eigenschaften eines Skalarproduktes sondern zeigen nur die Linearität der Abbildung $f \mapsto \langle f, g\rangle_{H'}$. Für beliebige $f_1, f_2, g \in H'$, $\alpha_1, \alpha_2 \in \mathbb{K}$ ist

$$\begin{aligned}\langle \alpha_1 f_1 + \alpha_2 f_2, g\rangle_{H'} &= \langle J^{-1}g, J^{-1}(\alpha_1 f_1 + \alpha_2 f_2)\rangle_H = \langle J^{-1}g, \overline{\alpha}_1 J^{-1}f_1 + \overline{\alpha}_2 J^{-1}f_2\rangle_H \\ &= \alpha_1\langle J^{-1}g, J^{-1}f_1\rangle_H + \alpha_2\langle J^{-1}g, J^{-1}f_2\rangle_H = \alpha_1\langle f_1, g\rangle_{H'} + \alpha_2\langle f_2, g\rangle_{H'},\end{aligned}$$

da mit J auch J^{-1} antilinear ist. Sei nun $\varphi \in H'' = (H')'$. Da H' ein Hilbertraum ist, gibt es nach dem Rieszschen Darstellungssatz ein $g \in H'$, so dass

$$\varphi(f) = \langle f, g\rangle_{H'} \quad \text{für alle } f \in H'.$$

Setzen wir $J^{-1}g = x \in H$, so erhalten wir für alle $f \in H'$

$$\langle f, g\rangle_{H'} = \langle J^{-1}g, J^{-1}f\rangle_H = \langle x, J^{-1}f\rangle_H = J(J^{-1}f)(x) = f(x) = F(x)(f).$$

Es ist also $\varphi(f) = f(x)$, d.h. φ ist von der Gestalt $F(x)$. Also ist F surjektiv. ∎

3.2.4 Adjungierte Operatoren

Seien X und Y normierte Räume und $A \in L(X,Y)$. Es ist klar, dass für jedes Funktional $f \in Y'$ durch $x \mapsto f(Ax)$ ein lineares und stetiges Funktional auf X definiert wird. Dieses bezeichnen wir mit $A'f$. Der so erklärte Operator $A' : Y' \to X'$ heißt die *Adjungierte* von A. Es ist also $(A'f)(x) = f(Ax)$.

Satz 3.20 (a) *Seien X und Y normierte Räume und $A \in L(X,Y)$. Dann ist $A' \in L(Y', X')$, und es gilt*

$$\|A\|_{L(X,Y)} = \|A'\|_{L(Y',X')}.$$

(b) *Es ist $(\alpha_1 A_1 + \alpha_2 A_2)' = \alpha_1 A_1' + \alpha_2 A_2'$ für alle $A_1, A_2 \in L(X,Y)$ und $\alpha_1, \alpha_2 \in \mathbb{K}$ sowie $(AB)' = B'A'$ für alle $B \in L(X,Y)$ und $A \in L(Y,Z)$. Ist $A \in L(X,Y)$ invertierbar, so ist auch $A' \in L(Y',X')$ invertierbar, und es ist $(A')^{-1} = (A^{-1})'$.*

Beweis: *Linearität von A'.* Seien $f_1, f_2 \in Y'$ und $\alpha_1, \alpha_2 \in \mathbb{K}$. Dann gilt für alle $x \in X$

$$\begin{aligned} A'(\alpha_1 f_1 + \alpha_2 f_2)(x) &= (\alpha_1 f_1 + \alpha_2 f_2)(Ax) = \alpha_1 f_1(Ax) + \alpha_2 f_2(Ax) \\ &= \alpha_1 (A'f_1)(x) + \alpha_2 (A'f_2)(x), \end{aligned}$$

d.h.

$$A'(\alpha_1 f_1 + \alpha_2 f_2) = \alpha_1 A'f_1 + \alpha_2 A'f_2.$$

Beschränktheit von A'. Für alle $x \in X$ und $f \in Y'$ ist

$$|(A'f)(x)| = |f(Ax)| \le \|f\|_{Y'} \|Ax\|_Y \le \|f\|_{Y'} \|A\|_{L(X,Y)} \|x\|_X.$$

Hieraus folgt zunächst

$$\|A'f\|_{X'} \le \|f\|_{Y'} \|A\|_{L(X,Y)}$$

und weiter

$$\|A'\|_{L(Y',X')} \le \|A\|_{L(X,Y)}.$$

Normgleichheit. Wir zeigen nun noch $\|A\|_{L(X,Y)} \le \|A'\|_{L(Y',X')}$ bzw. $\|Ax\| \le \|A'\| \, \|x\|$ für alle $x \in X$. Ist $Ax = 0$, so gilt dies offenbar. Sei also $Ax \neq 0$. Nach Hahn-Banach (Folgerung 3.14) gibt es ein $f \in Y'$ so, dass $\|f\| = 1$ und $f(Ax) = \|Ax\|$. Für dieses f ist

$$\|Ax\| = f(Ax) = |f(Ax)| = |(A'f)(x)| \le \|f\| \, \|A'\| \, \|x\| = \|A'\| \, \|x\|.$$

Damit ist (a) bewiesen, und (b) ist HA. ■

Satz 3.21 *Seien X und Y normierte Räume und $A \in L(X,Y)$. Für den Operator $A'' := (A')' : X'' \to Y''$ gilt dann $A''\big|_X = A$.*

Hier haben wir implizit benutzt, dass X in X'' und Y in Y'' isometrisch eingebettet sind. Genauer müsste man die Aussage des Satzes so formulieren:

$$A''\big(F_X(x)\big) = F_Y(Ax) \quad \text{für alle } x \in X, \tag{3.10}$$

wobei $F_X : X \to X''$ und $F_Y : Y \to Y''$ die Einbettungen gemäß Satz 3.17 sind.

Beweis Für beliebige $x \in X$ und $f \in Y'$ ist

$$\begin{aligned} \big(A'' F_X(x)\big)(f) &= F_X(x)(A'f) = (A'f)(x) \\ &= f(Ax) = F_Y(Ax)(f). \end{aligned}$$

Da dies für alle $f \in Y'$ gilt, folgt die Behauptung (3.10). ■

Beispiel 1: Ist H Hilbertraum und $A \in L(H)$, so haben wir zwei verschiedene Arten von adjungierten Operatoren für A definiert:

- die Hilbertraum-Adjungierte $A^* : H \to H$ durch $\langle Ax, y\rangle = \langle x, A^*y\rangle$ $\forall x, y \in H$,
- die "Banachraum-Adjungierte" $A' : H' \to H'$ durch $(A'f)(x) = f(Ax)$ $\forall x \in H$, $f \in H'$.

Der Zusammenhang zwischen A^* und A' wird wieder über die antilineare Isometrie $J : H \to H'$ hergestellt. Es ist nämlich für alle $x \in H$ und $f \in H'$

$$(A'f)(x) = f(Ax) = \langle Ax, J^{-1}f\rangle = \langle x, A^*J^{-1}f\rangle = (JA^*J^{-1}f)(x),$$

d.h.

$$A'f = JA^*J^{-1}f \quad \forall f \quad \text{bzw.} \quad A' = JA^*J^{-1}.$$

Natürlich ist dann auch $A^* = J^{-1}A'J$. ■

Beispiel 2: Für $1 < p < \infty$ bestimmen wir die Adjungierte des Verschiebungsoperators

$$V_1 : \ell^p \to \ell^p, \quad (x_1, x_2, x_3, \ldots) \mapsto (0, x_1, x_2, x_3, \ldots).$$

Erklären wir q durch $1/p + 1/q = 1$, so können wir, wie wir früher gesehen haben, die Räume $(\ell^p)'$ und ℓ^q miteinander identifizieren. Genauer: die Abbildung $\Phi : \ell^q \to (\ell^p)'$, die der Folge $(y_n) \in \ell^q$ das Funktional

$$(x_n) \mapsto \sum_{n=1}^{\infty} x_n y_n$$

zuordnet, ist eine lineare Isometrie von ℓ^q auf $(\ell^p)'$. Erklären wir für $f \in (\ell^p)'$ die Folge $y = (y_n) \in \ell^q$ durch $y = \Phi^{-1}f$, so erhalten wir für alle $x = (x_n) \in \ell^p$:

$$\begin{aligned} (V_1'f)(x) &= f(V_1x) = (\Phi y)(V_1x) = \sum_{n=1}^{\infty} (V_1x)_n\, y_n = \sum_{n=2}^{\infty} x_{n-1}y_n \\ &= \sum_{n=1}^{\infty} x_n y_{n+1} = \sum_{n=1}^{\infty} x_n (V_{-1}y)_n, \end{aligned} \tag{3.11}$$

wobei $V_{-1} : \ell^q \to \ell^q$ für den Operator der "Rückwärtsverschiebung" steht:

$$V_{-1}(y_1, y_2, y_3, \ldots) = (y_2, y_3, y_4, \ldots).$$

Aus (3.11) erhalten wir weiter

$$(V_1' f)(x) = (\Phi V_{-1} y)(x) = (\Phi V_{-1} \Phi^{-1} f)(x),$$

d.h. es ist $V_1' f = \Phi V_{-1} \Phi^{-1} f$ bzw. $V_1' = \Phi V_{-1} \Phi^{-1}$. Im gleichen Sinn, wie man $(\ell^p)'$ und ℓ^q identifiziert, identifiziert man auch V_1' mit V_{-1} und schreibt einfach $V_1' = V_{-1}$. ■

Beispiel 3: Für $k \in C(X \times X)$ ist der durch

$$(Af)(t) = \int_X k(t,s) f(s) ds, \quad t \in X$$

definierte Operator beschränkt als Operator von $L^p(X)$ nach $L^p(X)$ $(1 < p < \infty)$. Identifiziert man $L^p(X)'$ mit $L^q(X)$ $(1/p + 1/q = 1)$, so kann man den adjungierten Operator zu A identifizieren mit

$$(A'g)(t) = \int_X k'(t,s) g(s) ds,$$

wobei $k'(t,s) = k(s,t)$. ■

Abschließend überlegen wir uns, dass sich die Kompaktheit eines Operators auf dessen Adjungierte überträgt.

Satz 3.22 *Aus der Kompaktheit von $A \in L(X,Y)$ folgt die von $A' \in L(Y', X')$.*

Beweis Für jeden normierten Raum Z bezeichnen wir mit B_Z seine abgeschlossene Einheitskugel. Wir müssen zeigen, dass $A'(B_{Y'})$ relativ kompakt ist. Dazu zeigen wir, dass für jedes $\varepsilon > 0$ ein endliches ε-Netz in dieser Menge existiert.

Sei also $\varepsilon > 0$ beliebig vorgegeben. Da $A : X \to Y$ kompakt ist, finden wir ein endliches $\varepsilon/3$-Netz $\{Ax_1, Ax_2, \ldots, Ax_n\}$ in $A(B_X)$. Wir betrachten $\mathbb{K}^n$ mit der Maximumnorm. Durch

$$S : Y' \to \mathbb{K}^n, \quad f \mapsto \big(f(Ax_1), \ldots, f(Ax_n)\big)$$

wird eine (offenbar) lineare und (wegen $\|Sf\|_\infty \le \|f\| \, \|A\|$) beschränkte Abbildung definiert. Wegen der Endlichdimensionalität von $\mathbb{K}^n$ ist S kompakt, und man findet daher ein endliches $\varepsilon/3$–Netz $\{Sf_1, \ldots, Sf_m\}$ in $S(B_{Y'})$. Wir zeigen, dass $\{A'f_1, \ldots, A'f_m\}$ ein ε–Netz in $A'(B_{Y'})$ bildet.

Sei $f \in B_{Y'}$ beliebig. Dann findet man ein k so, dass

$$|f(Ax_j) - f_k(Ax_j)| = |(Sf)_j - (Sf_k)_j| \le \|Sf - Sf_k\|_\infty < \varepsilon/3$$

für alle $j = 1, \dots, n$. Weiter: Für jedes $x \in B_X$ finden wir ein x_j derart, dass $\|Ax - Ax_j\| < \frac{\varepsilon}{3}$. Mit diesem x_j ist

$$\begin{aligned}|(A'f)(x) - (A'f_k)(x)| &= |f(Ax) - f_k(Ax)| \\ &\leq |f(Ax) - f(Ax_j)| + |f(Ax_j) - f_k(Ax_j)| + |f_k(Ax_j) - f_k(Ax)| \\ &\leq \|Ax - Ax_j\| + |f(Ax_j) - f_k(Ax_j)| + \|Ax_j - Ax\| < \varepsilon.\end{aligned}$$

Da dies für jedes $x \in B_X$ gilt, ist $\|A'f - A'f_k\| \leq \varepsilon$. ■

Es gilt auch die umgekehrte Aussage: Aus der Kompaktheit von $A' \in L(Y', X')$ folgt die von $A \in L(X, Y)$. Dies folgt aus Satz 3.22 und aus $A''|_X = A$.

3.2.5 Normal auflösbare Operatoren

Zur Motivation sei zunächst $A : \mathbb{R}^n \to \mathbb{R}^m$ eine Matrix. Aus der linearen Algebra wissen wir, dass das lineare Gleichungssystem $Ax = y$ genau dann lösbar ist (d.h., dass $y \in \operatorname{Im} A$), wenn y senkrecht steht zu allen Lösungen f der homogenen Gleichung $A^* f = 0$. Für Operatoren auf unendlich-dimensionalen Räumen gilt dies i.Allg. nicht mehr. Wir definieren daher:

Definition 3.23 *Seien X und Y normierte Räume. Ein Operator $A \in L(X, Y)$ heißt* normal auflösbar, *wenn gilt: die Gleichung $Ax = y$ ist genau dann lösbar, wenn $f(y) = 0$ für alle $f \in Y'$ mit $A'f = 0$.*

Satz 3.24 *Seien X und Y normierte Räume und $A \in L(X, Y)$. Dann gilt*

(a)
$$\overline{\operatorname{Im} A} = \{y \in Y : f(y) = 0 \quad \text{für alle } f \in \operatorname{Ker} A'\}. \tag{3.12}$$

(b) *A ist genau dann normal auflösbar, wenn $\operatorname{Im} A$ abgeschlossen ist.*

Beweis (a) Wir bezeichnen die Menge auf der rechten Seite von (3.12) mit Z. Sei $y \in \operatorname{Im} A$, d.h. $y = Ax$ mit einem $x \in X$. Für alle $f \in \operatorname{Ker} A'$ ist dann

$$f(y) = f(Ax) = (A'f)(x) = 0, \quad \text{d.h. } y \in Z.$$

Folglich ist $\operatorname{Im} A \subseteq Z$. Außerdem ist klar, dass Z abgeschlossen ist (Stetigkeit von f beachten). Daher gilt auch $\overline{\operatorname{Im} A} \subseteq Z$.
Angenommen, $\overline{\operatorname{Im} A}$ ist echt enthalten in Z. Dann gibt es ein $y_0 \in Z$, welches nicht zu $\overline{\operatorname{Im} A}$ gehört. Aus Hahn-Banach (Folgerung 3.15) folgt die Existenz eines Funktionals $f \in Y'$ mit $f(y_0) \neq 0$ und $f(\operatorname{Im} A) = 0$. Dieses Funktional liegt wegen $(A'f)(x) = f(Ax) = 0$ für alle x in $\operatorname{Ker} A'$. Dann ist aber $f(Z) = 0$ und insbesondere $f(y_0) = 0$, was der Bedingung $f(y_0) \neq 0$ widerspricht.

(b) Die Bedingung der normalen Auflösbarkeit für A läßt sich schreiben als

$$\operatorname{Im} A = \{y \in Y : f(y) = 0 \quad \text{für alle } f \in \operatorname{Ker} A'\}.$$

Ein Vergleich mit (3.12) liefert sofort die Behauptung. ■

In Verallgemeinerung des Begriffes der Orthogonalität in Hilberträumen trifft man für allgemeine normierte Räume X die folgenden Definitionen: Für jede Teilmenge M von X ist

$$M^{\perp} := \{f \in X' : \quad f(x) = 0 \text{ für alle } x \in M\},$$

und für jede Teilmenge N von X' ist

$$^{\perp}N := \{x \in X : \quad f(x) = 0 \text{ für alle } f \in N\}.$$

Mit diesen Bezeichnungen kann (3.12) auch wie folgt geschrieben werden:

$$\overline{\operatorname{Im} A} \;=\; {}^{\perp}(\operatorname{Ker} A'). \tag{3.13}$$

Man kann sich auch leicht klarmachen, dass für jeden Operator $A \in L(X, Y)$ gilt

$$^{\perp}(\operatorname{Im} A') = \operatorname{Ker} A, \quad (\operatorname{Im} A)^{\perp} = \operatorname{Ker} A'. \tag{3.14}$$

Folgerung 3.25 *Seien X und Y Banachräume, $A \in L(X, Y)$, und A und A' seien nach unten beschränkt (d.h. $\|Ax\| \geq M\|x\|$, $\|A'f\| \geq M'\|f\|$ für alle $x \in X$ und $f \in Y'$ mit positiven Konstanten M und M'). Dann ist A stetig invertierbar.*

Beweis Wir wissen bereits, dass aus $\|Ax\| \geq M\|x\|$ folgt, dass $\operatorname{Im} A$ abgeschlossen und $\operatorname{Ker} A = \{0\}$ ist. Wegen (3.12) erhalten wir weiter

$$\operatorname{Im} A = \{y \in Y : \quad f(y) = 0 \text{ für alle } f \in \operatorname{Ker} A'\}. \tag{3.15}$$

Aus $\|A'f\| \geq M'\|f\|$ folgt aber, dass auch $\operatorname{Ker} A' = \{0\}$. Folglich ist die rechte Seite von (3.15) gleich Y, d.h. $\operatorname{Im} A = Y$. Also ist A invertierbar, und nach dem Satz von Banach ist A^{-1} stetig. ■

Offenbar gilt auch die Umkehrung von Folgerung 3.25: Ist A stetig invertierbar, so sind A und A' nach unten beschränkt.

Es stellt sich die Frage, ob sich Klassen von Operatoren angeben lassen, die normal auflösbar sind. Dies wird eines der Probleme in den nächsten Abschnitten sein. Hier vermerken wir - ohne Beweis - nur folgendes:

Satz 3.26 *Ein kompakter Operator $K \in L(X, Y)$ ist genau dann normal auflösbar, wenn er von endlichem Rang ist (d.h. wenn $\operatorname{Im} K$ ein endlichdimensionaler Raum ist).*

Unter unseren üblichen Voraussetzungen sind Fredholmsche oder Volterrasche Integralgleichungen *1. Art* gerade von der Form $Kx = y$ mit einem kompakten Operator K, der in der Regel *nicht* von endlichem Rang ist. Satz 3.26 deutet einen Grund an, warum die Behandlung solcher Gleichungen ziemliche Schwierigkeiten bereitet (man spricht auch von *schlecht gestellten Aufgaben*). Wir wenden uns nun der genaueren Untersuchung von Gleichungen der 2. Art zu.

4 Spektraltheorie

4.1 Das Spektrum eines linearen beschränkten Operators

Eine Zahl λ heißt bekanntlich *Eigenwert* einer $n \times n$-Matrix A, wenn es einen Vektor $x \neq 0$ (einen sog. Eigenvektor von A) gibt, für den $(A - \lambda I)x = 0$. Insbesondere ist $A - \lambda I$ dann nicht invertierbar. Für Operatoren auf unendlichdimensionalen Räumen gibt es noch andere Ursachen, warum ein Operator $A - \lambda I$ nicht invertierbar ist.
Im weiteren sei X ein Banachraum, $A \in L(X)$, und I sei der identische Operator auf X. Außerdem wollen wir von nun an $\mathbb{K} = \mathbb{C}$ annehmen.

Definition 4.1 *Das* Spektrum $\sigma(A)$ *von* $A \in L(X)$ *ist die Menge aller* $\lambda \in \mathbb{C}$, *für die der Operator* $A - \lambda I$ *nicht invertierbar ist. Die Differenz* $\mathbb{C} \backslash \sigma(A)$ *heißt die* Resolventenmenge *von* A. *Bezeichnung:* $\rho(A)$.

Satz 4.2 *Für jeden Operator* $A \in L(X)$ *ist* $\sigma(A)$ *eine kompakte (d.h. beschränkte und abgeschlossene) und nichtleere Teilmenge von* $\mathbb{C}$.

Beweis Ist $\lambda \in \mathbb{C}$ eine Zahl mit $|\lambda| > \|A\|$, so ist $\|\frac{1}{\lambda}A\| < 1$, und der Satz über die Neumann-Reihe garantiert, dass $I - \frac{1}{\lambda}A$ invertierbar ist. Dann ist aber auch $(-\lambda)\big(I - \frac{1}{\lambda}A\big) = A - \lambda I$ invertierbar. Das Spektrum $\sigma(A)$ ist also in der Kreisscheibe $\{z \in \mathbb{C} : |z| \leq \|A\|\}$ enthalten und folglich beschränkt.
Die Abgeschlossenheit von $\sigma(A)$ zeigen wir, indem wir nachweisen, dass das Komplement $\rho(A) = \mathbb{C} \backslash \sigma(A)$ offen ist. Sei $\lambda \in \rho(A)$, d.h. $A - \lambda I$ ist invertierbar. Für alle μ mit

$$|\mu| < \frac{1}{\|(A - \lambda I)^{-1}\|}$$

ist $\|\mu(A - \lambda I)^{-1}\| < 1$, d.h. der Operator $I - \mu(A - \lambda I)^{-1}$ ist invertierbar (Neumann-Reihe). Dann ist aber auch der Operator

$$(A - \lambda I)\big(I - \mu(A - \lambda I)^{-1}\big) = A - \lambda I - \mu I = A - (\lambda + \mu)I$$

invertierbar. Folglich ist $\rho(A)$ offen und $\sigma(A)$ abgeschlossen.
Der Beweis, dass $\sigma(A)$ niemals leer ist, wird meist mit Mitteln der komplexen Funktionentheorie geführt. Die Grundidee ist wie folgt. Angenommen, $\sigma(A)$ ist leer. Dann ist die Funktion $\lambda \mapsto (A - \lambda I)^{-1}$ auf der gesamten komplexen Ebene definiert. Man zeigt nun, dass diese Funktion holomorph ist, dass sie überall beschränkt ist, und dass $\lim_{\lambda \to \infty} \|(A - \lambda I)^{-1}\| = 0$. Die Holomorphie kann man wie folgt einsehen: Setzt man in der Identität $X^{-1} - Y^{-1} = X^{-1}(Y - X)Y^{-1}$ für $X = \lambda I - A$ und für $Y = \mu I - A$ mit $\lambda, \mu \in \rho(A)$, so gelangt man zur so genannten *Resolventengleichung*

$$(\lambda I - A)^{-1} - (\mu I - A)^{-1} = (\mu - \lambda)(\lambda I - A)^{-1}(\mu I - A)^{-1}.$$

Aus dieser folgt, dass die Funktion $\lambda \mapsto (\lambda I - A)^{-1}$ auf der gesamten Resolventenmenge $\rho(A)$ komplex differenzierbar und damit holomorph ist.

Aus der Holomorphie und der Beschränktheit folgt mit dem Satz von Liouville, dass die Funktion $\lambda \mapsto (A - \lambda I)^{-1}$ nur eine Konstante sein kann, und aus der Grenzwerteigenschaft folgt, dass diese Konstante gleich 0 ist. Also wäre $(A - \lambda I)^{-1} = 0$ für alle λ, was natürlich unmöglich ist. ∎

Anmerkung Die Eigenschaft $\sigma(A) \neq \emptyset$ kann *nicht* mehr gezeigt werden, wenn man Operatoren auf reellen Banachräumen betrachtet und $\sigma(A)$ entsprechend als Menge aller $\lambda \in \mathbb{R}$ definiert, für die $A - \lambda I$ nicht invertierbar ist. Dies ist einer der Gründe dafür, warum wir in diesem Abschnitt $\mathbb{K} = \mathbb{C}$ wählen. Für $A = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ ist beispielsweise $\sigma(A) = \emptyset$ für $\mathbb{K} = \mathbb{R}$ und $\sigma(A) = \{\pm i\}$ für $\mathbb{K} = \mathbb{C}$. ∎

Die kleinste Zahl $r \geq 0$ mit der Eigenschaft, dass $\sigma(A)$ im Kreis $\{z \in \mathbb{C} : |z| \leq r\}$ liegt, heißt *Spektralradius* von A. Wir bezeichnen den Spektralradius mit $r(A)$. Im Beweis von Satz 4.2 haben wir gesehen, dass

$$r(A) \leq \|A\|. \tag{4.1}$$

Man kann den Spektralradius wie folgt berechnen.

Satz 4.3 *Sei X ein Banachraum und $A \in L(X)$. Dann ist*

$$r(A) = \lim_{n\to\infty} \sqrt[n]{\|A^n\|} \tag{4.2}$$

(die Existenz des Grenzwertes ist Teil der Behauptung).

Bemerkenswert an (4.2) ist, dass algebraische (Spektralradius) mit metrischen (Norm von A^n) Größen verknüpft werden.

Beweisidee Aus dem Beweis von Satz 4.2 wissen wir, dass die Funktion $\lambda \mapsto (\lambda I - A)^{-1}$ auf der gesamten Resolventenmenge holomorph ist. Insbesondere ist diese Funktion also auf dem Kreisring $\{\lambda \in \mathbb{C} : |\lambda| > r(A)\}$ holomorph. Holomorphe Funktionen auf Kreisringen lassen sich in Laurentreihen entwickeln, und die Reihenentwicklung von $(\lambda I - A)^{-1} = \lambda^{-1}(I - \lambda^{-1}A)^{-1}$ führt gerade auf die Reihe $\sum_{n=0}^{\infty} \frac{1}{\lambda^{n+1}} A^n$. Diese Reihe konvergiert also für alle λ mit $|\lambda| > r(A)$.

Nach dem Wurzelkriterium (was auch für Reihen in $L(X)$ gilt) konvergiert die Potenzreihe $z \mapsto \sum_{n=0}^{\infty} z^n A^n$ für $|z| < (\limsup \sqrt[n]{\|A^n\|})^{-1}$, und sie divergiert für $|z| > (\limsup \sqrt[n]{\|A^n\|})^{-1}$. Vergleichen wir dies mit dem vorher gewonnenen Resultat, folgt

$$\limsup \sqrt[n]{\|A^n\|} \leq r(A). \tag{4.3}$$

Weiter: Für jedes $\lambda \in \mathbb{C}$ und $n \geq 1$ gilt offenbar

$$\lambda^n I - A^n = (A^{n-1} + \lambda A^{n-2} + \ldots + \lambda^{n-1} I)(\lambda I - A).$$

Hieraus folgt: Ist $\lambda \in \sigma(A)$, so ist $\lambda^n \in \sigma(A^n)$. Wegen (4.1) bedeutet das

$$|\lambda^n| \leq \|A^n\| \quad \text{bzw.} \quad |\lambda| \leq \sqrt[n]{\|A^n\|}$$

für alle $\lambda \in \sigma(A)$ und $n \in \mathbb{N}$. Dann ist aber auch $|\lambda| \leq \liminf \sqrt[n]{\|A^n\|}$, und da dies für alle $\lambda \in \sigma(A)$ gilt, folgt

$$r(A) \leq \liminf \sqrt[n]{\|A^n\|}. \tag{4.4}$$

Zusammen mit (4.3) liefert dies die Behauptung. ■

Die Punkte des Spektrums eines Operators werden weiter klassifiziert. Ist $\lambda \in \sigma(A)$, und gibt es einen Vektor $x \in X\backslash\{0\}$, so dass $(A - \lambda I)x = 0$ (d.h. ist $\mathrm{Ker}\,(A - \lambda I) \neq \{0\}$), so heißt λ ein *Eigenwert* und x ein *Eigenvektor* von A. Die Menge aller Eigenwerte von A heißt das *Punktspektrum* von A und wird mit $\sigma_p(A)$ bezeichnet. Alle $\lambda \in \sigma(A)$ mit

$$\mathrm{Ker}\,(A - \lambda I) = \{0\} \quad \text{und} \quad \overline{\mathrm{Im}\,(A - \lambda I)} = X$$

faßt man zum *kontinuierlichen Spektrum* $\sigma_c(A)$ zusammen, und die übrigen Punkte aus $\sigma(A)$, für die dann also gilt

$$\mathrm{Ker}\,(A - \lambda I) = \{0\} \quad \text{und} \quad \overline{\mathrm{Im}\,(A - \lambda I)} \neq X$$

zum *Restspektrum* $\sigma_r(A)$. Offenbar ist $\sigma(A) = \sigma_p(A) \cup \sigma_c(A) \cup \sigma_r(A)$. Ist $A : \mathbb{C}^n \to \mathbb{C}^n$ (der aus der linearen Algebra bekannte Fall), so ist $\sigma(A) = \sigma_p(A)$, d.h. alle Punkte im Spektrum sind Eigenwerte.

Ist $X = H$ ein komplexer Hilbertraum, so folgt aus den bekannten Beziehungen $(AB)^* = B^*A^*$ (Hilbertraumadjungierte) und $(\alpha I)^* = \overline{\alpha} I$ sofort, dass

$$\sigma(A^*) = \overline{\sigma(A)}.$$

Unter Beachtung von (vgl. Satz 2.44 aus 2.5)

$$\mathrm{Ker}\,(A^* - \overline{\lambda} I) = \big(\mathrm{Im}\,(A - \lambda I)\big)^\perp, \quad \mathrm{clos}\,\mathrm{Im}\,(A^* - \overline{\lambda} I) = \big(\mathrm{Ker}\,(A - \lambda I)\big)^\perp$$

lässt sich darüber hinaus leicht zeigen, dass für $A \in L(H)$ gilt

$$\begin{aligned} \lambda \in \sigma_p(A) \quad &\Rightarrow \quad \overline{\lambda} \in \sigma_p(A^*) \cup \sigma_r(A^*)\,, \\ \lambda \in \sigma_r(A) \quad &\Rightarrow \quad \overline{\lambda} \in \sigma_p(A^*)\,, \\ \lambda \in \sigma_c(A) \quad &\Leftrightarrow \quad \overline{\lambda} \in \sigma_c(A^*). \end{aligned}$$

($\nearrow$ Übung)

Für Operatoren mit speziellen Eigenschaften lassen sich genauere Aussagen über Lage und Klassifikation des Spektrums machen. Wir fassen einige solcher Resultate im folgenden Satz zusammen.

Satz 4.4 *Sei H Hilbertraum und $A \in L(H)$.*

(a) *Ist A normal (d.h. $AA^* = A^*A$), so gilt für alle $\lambda \in \mathbb{C}$:*

$$\operatorname{Ker}(A - \lambda I) = \operatorname{Ker}(A^* - \overline{\lambda} I) = \big(\operatorname{Im}(A - \lambda I)\big)^{\perp}. \tag{4.5}$$

Damit gilt insbesondere

(i) $\operatorname{Ker} A = \operatorname{Ker} A^*$.

(ii) $\overline{\operatorname{Im} A} = H \Leftrightarrow \operatorname{Ker} A = \{0\}$. *Somit ist stets $\sigma_r(A) = \emptyset$.*

(iii) *A invertierbar $\Leftrightarrow \exists C > 0$ so, dass $\|Ax\| \geq C\|x\|$ für alle $x \in H$.*

(iv) $\lambda \in \sigma_p(A) \Leftrightarrow \overline{\lambda} \in \sigma_p(A^*)$.

(v) $\lambda, \mu \in \sigma_p(A)$, $\lambda \neq \mu \Rightarrow \operatorname{Ker}(A - \lambda I) \perp \operatorname{Ker}(A - \mu I)$.

(b) *Ist A selbstadjungiert (d.h. $A^* = A$), so ist $\sigma(A) \subseteq \mathbb{R}$. Außerdem ist $\|A\| = r(A)$.*

(c) *Ist A unitär (d.h. A invertierbar und $A^{-1} = A^*$), so ist $\sigma(A) \subseteq \mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}$.*

(d) *Ist A positiv (d.h. selbstadjungiert und $\langle Ax, x\rangle \geq 0$ für alle $x \in H$), so ist $\sigma(A) \subseteq \mathbb{R}^+ := \{z \in \mathbb{R} : z \geq 0\}$.*

(e) *Ist A ein Orthoprojektor (d.h. $A = A^* = A^2$), so ist $\sigma(A) \subseteq \{0, 1\}$.*

Beweis Wir zeigen zuerst (4.5). Wie wir bereits wissen, gilt für *jeden* Operator $B \in L(H)$ die Beziehung $\operatorname{Ker} B^* = (\operatorname{Im} B)^{\perp}$ (Satz 2.44 in 2.5). Ersetzen wir hierin B durch $A - \lambda I$, folgt sofort die zweite Identität aus (4.5). Für die erste Identität in (4.5) überlegen wir uns, dass für jeden *normalen Operator* $B \in L(H)$ gilt

$$\|Bx\| = \|B^*x\| \quad \forall x \in H. \tag{4.6}$$

Es ist nämlich

$$\begin{aligned} \|Bx\|^2 - \|B^*x\|^2 &= \langle Bx, Bx\rangle - \langle B^*x, B^*x\rangle = \langle B^*Bx, x\rangle - \langle BB^*x, x\rangle \\ &= \langle (B^*B - BB^*)x, x\rangle = 0, \end{aligned}$$

woraus (4.6) folgt. Beachten wir weiter, dass für jeden normalen Operator A auch $A - \lambda I$ wieder normal ist, so können wir in (4.6) B durch $A - \lambda I$ ersetzen und bekommen $\|(A - \lambda I)x\| = \|(A^* - \overline{\lambda} I)x\|$ $\forall x \in H$. Insbesondere ist also $(A - \lambda I)x = 0$ genau dann, wenn $(A^* - \overline{\lambda} I)x = 0$, d.h. es gilt auch die erste Identität in (4.5).

Damit sind die Aussagen (i), (ii) und (iv) sofort klar. Zu Aussage (iii): Wir wissen bereits, dass für jeden Operator $A \in L(H)$ gilt:

$$A \text{ invertierbar } \Rightarrow \exists\, C > 0 : \|Ax\| \geq C\|x\| \quad \forall x \in H$$

und dass umgekehrt aus der Beschränktheit nach unten folgt: $\operatorname{Ker} A = \{0\}$ und $\operatorname{Im} A$ ist abgeschlossen (Satz 2.8). Falls A normal ist, folgt mit (ii) hieraus, dass $\overline{\operatorname{Im} A} = \operatorname{Im} A = H$, d.h. A ist invertierbar.

Schließlich zeigen wir noch (v): Sei $(A - \lambda I)x = 0$, $(A - \mu I)y = 0$ und $\lambda \neq \mu$. Dann ist wegen (4.5) auch $(A^* - \overline{\mu} I)y = 0$, und wir erhalten

$$\lambda\langle x, y\rangle = \langle \lambda x, y\rangle = \langle Ax, y\rangle = \langle x, A^*y\rangle = \langle x, \overline{\mu}y\rangle = \mu\langle x, y\rangle,$$

also (da $\lambda \neq \mu$) $\langle x, y\rangle = 0$. Damit ist (a) gezeigt.

(b): Sei $A = A^*$ und $\lambda = \alpha + \beta i \in \mathbb{C}$. Für jedes $x \in H$ ist

$$\begin{aligned}\|(A-\lambda I)x\|^2 &= \|Ax - \alpha x - \beta i x\|^2 = \langle Ax - \alpha x - \beta i x,\ Ax - \alpha x - \beta i x\rangle \\ &= \langle Ax - \alpha x,\ Ax - \alpha x\rangle \underbrace{-\langle Ax - \alpha x, \beta i x\rangle - \langle \beta i x,\ Ax - \alpha x\rangle}_{=\beta i\langle (A-\alpha I)x, x\rangle - \beta i\langle x, (A-\alpha I)x\rangle = 0} + \langle \beta i x, \beta i x\rangle \\ &= \|Ax - \alpha x\|^2 + \|\beta x\|^2,\end{aligned}$$

also erst recht

$$\|(A - \lambda I)x\| \geq |\beta|\,\|x\| \quad \text{für alle } x \in H.$$

Im Falle $\beta \neq 0$ folgt hieraus und aus (a)(iii) die Invertierbarkeit von $A - \lambda I$. Für $\lambda \in \sigma(A)$ ist also $\lambda \in \mathbb{R}$. Aus der Formel für den Spektralradius und der für alle Operatoren $A \in L(H)$ gültigen Beziehung $\|A^*A\| = \|A\|^2$ folgt sofort, dass $r(A) = \|A\|$.

(c): Für unitäres A ist $\|A\| = \|A^*\| = 1$ (vgl. Satz 2.41 (b) aus 2.5) und folglich $\sigma(A) \subseteq \{z \in \mathbb{C} :\ |z| \leq 1\}$. Für $|\lambda| < 1$ ist $\|\lambda A^*\| < 1$. Daher ist in diesem Fall der Operator

$$A - \lambda I = A - \lambda A A^* = A(I - \lambda A^*)$$

invertierbar (Neumann-Reihe) und $\lambda \notin \sigma(A)$.

(d): Für $\lambda \in \mathbb{R}$, $\lambda < 0$, und A positiv ist

$$\begin{aligned}\|(A-\lambda I)x\|^2 &= \langle Ax - \lambda x, Ax - \lambda x\rangle = \underbrace{\|Ax\|^2}_{\geq 0} + \lambda^2\|x\|^2 \underbrace{-2\lambda}_{\geq 0} \underbrace{\langle Ax, x\rangle}_{\geq 0} \\ &\geq \lambda^2\|x\|^2,\end{aligned}$$

d.h. die Invertierbarkeit von $A - \lambda I$ folgt aus (a)(iii).

(e): Sei P Orthoprojektor und $Q := I - P$. Dann ist auch Q Orthoprojektor:

$$Q^* = I - P^* = I - P = Q, \quad Q^2 = (I - P)^2 = I - 2P + P^2 = I - P = Q,$$

und es gilt

$$PQ = P(I - P) = P - P^2 = 0$$

und analog $QP = 0$. Hieraus folgt, dass für $\alpha, \beta \in \mathbb{C}\backslash\{0\}$ der Operator $\alpha P + \beta Q$ invertierbar ist und dass $\alpha^{-1}P + \beta^{-1}Q$ sein Inverser ist:

$$(\alpha P+\beta Q)(\alpha^{-1}P+\beta^{-1}Q) = \alpha\alpha^{-1}P^2+\alpha\beta^{-1}PQ+\beta\alpha^{-1}QP+\beta\beta^{-1}Q^2 = P+Q = I.$$

Sein nun $\lambda \in \sigma(P)$, d.h. sei $P - \lambda I$ nicht invertierbar. Wegen

$$P - \lambda I = P - \lambda(P + Q) = (1 - \lambda)P + \lambda Q$$

ist dann notwendigerweise $\lambda = 1$ oder $\lambda = 0$. ■

4.2 Die Riesz-Theorie in Banachräumen

Die Riesz-Theorie macht Aussagen zum Spektrum kompakter Operatoren und damit zur Lösbarkeit von Fredholmschen Integralgleichungen $(\lambda I - K)u = f$ mit $\lambda \in \mathbb{C}$ und K kompakt. Einige der nachstehenden Resultate gelten für kompakte Operatoren auf normierten Räumen; der Einfachheit halber beschränken wir uns auf den Banachraumfall. Weitergehende Aussagen lassen sich für selbstadjungierte kompakte Operatoren auf Hilberträumen zeigen; diese so genannte Hilbert-Schmidt-Theorie betrachten wir im nächsten Abschnitt.

Satz 4.5 *Sei X ein Banachraum und $K \in K(X)$. Dann ist der Kern des Operators $A := I - K$ endlich-dimensional, und sein Bild ist abgeschlossen.*

Für den Beweis des ersten Teils benötigen wir eine Aussage, die uns in Spezialfällen bereits bekannt ist.

Rieszsches Lemma *Sei X ein normierter Raum. Dann ist die abgeschlossene Einheitskugel $B := \{x \in X : \|x\| \leq 1\}$ genau dann kompakt, wenn X endlich-dimensional ist.*

Beweis Sei X endlich-dimensional. Da B beschränkt und abgeschlossen ist, ist B nach dem Satz von Bolzano-Weierstraß kompakt.

Sei umgekehrt B kompakt, und $x_1, \ldots, x_m$ sei ein (endliches) 1/2-Netz für B.

Wir bezeichnen mit Y die lineare Hülle $\operatorname{span}\{x_1, \ldots, x_m\}$ und wollen zeigen, dass $X = Y$. Angenommen, es gibt ein x in $X\backslash Y$. Dann ist

$$d := \inf\{\|x - y\| : y \in Y\} > 0$$

(vgl. Abschnitt über beste Approximation und beachte, dass Y endlich-dimensional ist). Es gibt also ein $y \in Y$ so, dass $d \leq \|x - y\| \leq \frac{3}{2}\,d$. Wir wählen x_i so, dass $\left\|\frac{x-y}{\|x-y\|} - x_i\right\| < 1/2$. Dann ist

$$\frac{1}{2}\|x - y\| > \|x - y\| \left\|\frac{x - y}{\|x - y\|} - x_i\right\| = \Big\|x - \underbrace{y - \|x - y\|x_i}_{\in Y}\Big\| \geq d,$$

also $\|x-y\| > 2d$ im Widerspruch zu $\|x-y\| \leq \frac{3}{2}\, d$. ■

Beweis von Satz 4.5: Endlichdimensionalität des Kernes. Sei B die abgeschlossene Einheitskugel in $\operatorname{Ker} A = \operatorname{Ker}(I-K)$. Da $Ax=0$ gleichbedeutend mit $Kx=x$ ist, folgt $KB=B$. Als Bild einer beschränkten Menge ist KB und daher auch B relativ kompakt, und da B abgeschlossen ist, sogar kompakt. Nach dem Rieszschen Lemma ist $\operatorname{Ker} A$ endlich-dimensional.

Abgeschlossenheit des Bildes. Sei $(y_n) \subseteq \operatorname{Im} A$ eine konvergente Folge mit $\lim y_n =: y \in X$. Wir wählen $x_n \in X$ mit $y_n = Ax_n$, setzen

$$d_n := \inf\{\|x_n - z\| \,:\, z \in \operatorname{Ker} A\}$$

und wählen $z_n \in \operatorname{Ker} A$ so, dass $\|x_n - z_n\| \leq d_n + 1$. Wir zeigen, dass die Folge $(x_n - z_n)$ beschränkt ist. Angenommen, diese Folge ist unbeschränkt. Dann gibt es eine Teilfolge $(x_n - z_n)_{n\in\mathbb{N}'}$ mit $0 < \|x_n - z_n\| \to \infty$ für $n \to \infty$. Sei $w_n := \frac{x_n - z_n}{\|x_n - z_n\|}$ für $n \in \mathbb{N}'$. Dann ist einerseits

$$Aw_n = \frac{1}{\|x_n - z_n\|}\, A(x_n - z_n) = \frac{y_n}{\|x_n - z_n\|} \to 0, \tag{4.7}$$

andererseits besitzt die Folge $(w_n)_{n\in\mathbb{N}'}$ eine Teilfolge $(w_n)_{n\in\mathbb{N}''}$ so, dass die Folge $(Kw_n)_{n\in\mathbb{N}''}$ gegen ein $x \in X$ konvergiert (Kompaktheit von K). Wegen

$$w_n = Aw_n + Kw_n \to 0 + x = x \quad (n \in \mathbb{N}'')$$

gilt weiter $Aw_n \to Ax \overset{(4.7)}{=} 0$, also $x \in \operatorname{Ker} A$. Für die Folge $(w_n)_{n\in\mathbb{N}''}$ ist daher

$$\begin{aligned} 0 \leftarrow \|w_n - x\| &= \Big\|\frac{x_n - z_n}{\|x_n - z_n\|} - x\Big\| = \frac{1}{\|x_n - z_n\|}\,\big\|x_n - \underbrace{z_n - \|x_n - z_n\|x}_{\in \operatorname{Ker} A}\big\| \\ &\geq \frac{d_n}{\|x_n - z_n\|} \geq \frac{\|x_n - z_n\| - 1}{\|x_n - z_n\|} = 1 - \frac{1}{\|x_n - z_n\|} \to 1. \end{aligned}$$

Der erhaltene Widerspruch zeigt, dass die Folge $(x_n - z_n)$ beschränkt ist.

Nun ist es leicht, den Beweis zu vollenden: Da die Folge $(x_n - z_n)$ beschränkt ist, besitzt die Folge $\big(K(x_n - z_n)\big)$ eine konvergente Teilfolge $\big(K(x_n - z_n)\big)_{n\in\mathbb{N}'}$ mit Grenzelement z. Für $n \in \mathbb{N}'$ ist

$$x_n - z_n = A(x_n - z_n) + K(x_n - z_n) = y_n + K(x_n - z_n) \to y + z$$

und folglich

$$y_n = A(x_n - z_n) \to A(y+z),$$

d.h. $y = A(y+z)$. Also liegt y im Bild von A. ■

Eine Verallgemeinerung dieses Resultates lernen wir in Lemma 4.21 kennen.

Als unmittelbare Folgerung ergibt sich: Ist $\lambda \neq 0$ ein Eigenwert des kompakten Operators K, so ist der zugehörige *Eigenunterraum*

$$N_\lambda(K) := \{x \in X : (\lambda I - K)x = 0\} = \mathrm{Ker}\,(\lambda I - K)$$

endlich-dimensional, und der *Bildraum*

$$R_\lambda(K) := \{(\lambda I - K)x : x \in X\} = \mathrm{Im}\,(\lambda I - K)$$

ist abgeschlossen. Es ist nämlich $\lambda I - K = \lambda(I - \frac{1}{\lambda} K)$ und daher

$$\mathrm{Ker}\,(\lambda I - K) = \mathrm{Ker}\,\big(I - \frac{1}{\lambda}K\big), \quad \mathrm{Im}\,(\lambda I - K) = \mathrm{Im}\,(I - \frac{1}{\lambda}K),$$

und der Operator $\frac{1}{\lambda}K$ ist kompakt.

Will man eine Matrix A in ihre Jordansche Normalform überführen, genügt es nicht, die Eigenunterräume zu den einzelnen Eigenwerten von A zu betrachten. Vielmehr ist es in der Regel erforderlich, neben dem Kern von $\lambda I - A$ auch die Kerne der Matrizen $(\lambda I - A)^2$, $(\lambda I - A)^3$, $\ldots$ zu untersuchen. In der linearen Algebra zeigt man dann, dass sich die Kerne $(\lambda I - A)^k$ und $(\lambda I - A)^{k+1}$ von einer Stelle k_0 an nicht mehr unterscheiden, und man konstruiert die Jordansche Normalform von A mit Hilfe einer Basis in $\mathrm{Ker}\,(\lambda I - A)^{k_0}$. Bis zu einem gewissen Grad übertragen sich diese Resultate auf kompakte Operatoren auf Banachräumen.

Für jeden kompakten Operator $K \in L(X)$, jedes $\lambda \in \mathbb{C}$ und jedes $n \in \mathbb{N}$ definieren wir

$$\begin{aligned} N_\lambda^n(K) &:= \{x \in X : (\lambda I - K)^n x = 0\} = \mathrm{Ker}\,(\lambda I - K)^n, \\ R_\lambda^n(K) &:= \{(\lambda I - K)^n x : x \in X\} = \mathrm{Im}\,(\lambda I - K)^n, \end{aligned}$$

mit $(\lambda I - K)^0 := I$. Offenbar gilt dann:

$$\begin{aligned} \{0\} &= N_\lambda^0(K) \subseteq N_\lambda^1(K) \subseteq N_\lambda^2(K) \subseteq \ldots, \\ X &= R_\lambda^0(K) \supseteq R_\lambda^1(K) \supseteq R_\lambda^2(K) \supseteq \ldots. \end{aligned}$$

Weiter ist für $\lambda \neq 0$

$$(\lambda I - K)^n = \sum_{j=0}^{n} \binom{n}{j} \lambda^j (-K)^{n-j} = \lambda^n I - K_n = \lambda^n (I - \frac{1}{\lambda^n} K_n)$$

mit einem kompakten Operator K_n. Daher sind alle Räume $N_\lambda^n(K)$ endlich-dimensional und alle Räume $R_\lambda^n(K)$ abgeschlossen.

Satz 4.6 *Sei X ein Banachraum, $K \in K(X)$ und $\lambda \neq 0$. Dann gibt es eine Zahl m so, dass*

$$\begin{aligned} \{0\} &= N_\lambda^0(K) \overset{echt}{\subset} N_\lambda^1(K) \overset{echt}{\subset} \ldots \overset{echt}{\subset} N_\lambda^m(K) &= N_\lambda^{m+1}(K) = \ldots, \\ X &= R_\lambda^0(K) \overset{echt}{\supset} R_\lambda^1(K) \overset{echt}{\supset} \ldots \overset{echt}{\supset} R_\lambda^m(K) &= R_\lambda^{m+1}(K) = \ldots. \end{aligned}$$

Mit diesem m gilt außerdem

$$N_\lambda^m(K) \cap R_\lambda^m(K) = \{0\}, \quad N_\lambda^m(K) + R_\lambda^m(K) = X. \tag{4.8}$$

Beweis *1. Schritt:* Wir zeigen, dass die Folge $(N_\lambda^m(K))$ stationär wird. Angenommen, es ist $N_\lambda^n(K) \overset{echt}{\subset} N_\lambda^{n+1}(K)$ für alle n. Dann gibt es in X eine Folge (x_n) mit folgenden Eigenschaften: Für alle n ist

$$\begin{aligned} x_{n+1} \in N_\lambda^{n+1}(K)\backslash N_\lambda^n(K), \quad \|x_{n+1}\| = 1, \\ d\big(x_{n+1}, N_\lambda^n(K)\big) := \inf\{\|x_{n+1} - x\| : x \in N_\lambda^n(K)\} \geq 1/2. \end{aligned} \tag{4.9}$$

Dies kann man so einsehen: Da $N_\lambda^n(K)$ endlich-dimensional ist, gibt es ein $y_{n+1} \in N_\lambda^{n+1}(K)\backslash N_\lambda^n(K)$ und ein $z_n \in N_\lambda^n(K)$ so, dass

$$0 < d\big(y_{n+1}, N_\lambda^n(K)\big) \leq \|y_{n+1} - z_n\| \leq 2d\big(y_{n+1}, N_\lambda^n(K)\big).$$

Wir setzen

$$x_{n+1} := \frac{y_{n+1} - z_n}{\|y_{n+1} - z_n\|} \in N_\lambda^{n+1}(K)\backslash N_\lambda^n(K).$$

Dann ist $\|x_{n+1}\| = 1$, und für alle $z \in N_\lambda^n(K)$ gilt tatsächlich

$$\begin{aligned} \|x_{n+1} - z\| &= \frac{\big\|y_{n+1} - z_n - \|y_{n+1} - z_n\|z\big\|}{\|y_{n+1} - z_n\|} \\ &\geq \frac{d\big(y_{n+1}, N_\lambda^n(K)\big)}{2d\big(y_{n+1}, N_\lambda^n(K)\big)} = \frac{1}{2}. \end{aligned}$$

Es sei nun (x_n) eine Folge wie in (4.9). Wir zeigen, dass die Folge (Kx_n) keine konvergente Teilfolge besitzt. Dazu betrachten wir für beliebiges $n, j \in \mathbb{N}$

$$\|Kx_n - Kx_{n+j}\| = \|\lambda x_n - (\lambda I - K)x_n + (\lambda I - K)x_{n+j} - \lambda x_{n+j}\|.$$

Das Element $\lambda x_n - (\lambda I - K)x_n + (\lambda I - K)x_{n+j}$ liegt in $N_\lambda^{n+j-1}(K)$. Es ist nämlich

$$x_n \in N_\lambda^n(K) \subseteq \ldots \subseteq N_\lambda^{n+j-1}(K),$$

und wegen $(\lambda I - K)N_\lambda^{n+1}(K) \subseteq N_\lambda^n(K)$ folgt auch

$$(\lambda I - K)x_n \in N_\lambda^{n-1}(K) \subseteq \ldots \subseteq N_\lambda^{n+j-1}(K), \quad (\lambda I - K)x_{n+j} \in N_\lambda^{n+j-1}(K).$$

Wegen der Wahl von (x_n) ist dann

$$\|Kx_n - Kx_{n+j}\| \geq |\lambda|/2 \quad \forall n, j$$

d.h. die Folge (Kx_n) besitzt tatsächlich keine konvergente Teilfolge, was der Kompaktheit von K widerspricht. Also war die Annahme falsch; es gibt folglich ein k mit

$$N_\lambda^k(K) = N_\lambda^{k+1}(K).$$

Wir überlegen uns nun noch, dass aus $N_\lambda^k(K) = N_\lambda^{k+1}(K)$ folgt $N_\lambda^{k+1}(K) = N_\lambda^{k+2}(K)$. Sei $y \in N_\lambda^{k+2}(K)$, d.h. $(\lambda I - K)^{k+2} y = 0$. Dann ist

$$(\lambda I - K)^{k+1}(\lambda I - K)y = 0,$$

d.h. $(\lambda I - K)y \in N_\lambda^{k+1}(K)$. Nach Annahme folgt $(\lambda I - K)y \in N_\lambda^k(K)$, und dies wiederum bedeutet $(\lambda I - K)^{k+1} y = 0$ bzw. $y \in N_\lambda^{k+1}(K)$. Mit vollständiger Induktion ist nun klar: es gibt ein k so, dass

$$\{0\} = N_\lambda^0(K) \overset{echt}{\subset} N_\lambda^1(K) \overset{echt}{\subset} \ldots \overset{echt}{\subset} N_\lambda^k(K) = N_\lambda^{k+1}(K) = \ldots .$$

2. Schritt: Ähnlich (vgl. Schröder, FA, S. 188) zeigt man, dass es ein ℓ so gibt, dass

$$X = R_\lambda^0(K) \overset{echt}{\supset} R_\lambda^1(K) \overset{echt}{\supset} \ldots \overset{echt}{\supset} R_\lambda^\ell(K) = R_\lambda^{\ell+1}(K) = \ldots .$$

3. Schritt: Sei $m := \max\{k, \ell\}$. Wir zeigen, dass mit diesem m die Beziehungen (4.8) gelten. Ist $x \in N_\lambda^m(K) \cap R_\lambda^m(K)$, so ist $(\lambda I - K)^m x = 0$, und es gibt ein y mit $x = (\lambda I - K)^m y$. Dann ist $(\lambda I - K)^{2m} y = 0$, d.h. $y \in N_\lambda^{2m}(K) = N_\lambda^m(K)$. Es ist also $(\lambda I - K)^m y = 0$, d.h. $x = 0$.
Insbesondere ist die Einschränkung $(\lambda I - K)^m|_{R_\lambda^m(K)}$ injektiv. Diese Einschränkung ist aber auch surjektiv:

$$(\lambda I - K)^m R_\lambda^m(K) = R_\lambda^{2m}(K) = R_\lambda^m(K).$$

Da $R_\lambda^m(k)$ ein Banachraum ist, ist der Operator $(\lambda I - K)^m|_{R_\lambda^m(K)} : R_\lambda^m(K) \to R_\lambda^m(K)$ stetig invertierbar (Satz von Banach). Für beliebiges $x \in X$ sei

$$y := \left((\lambda I - K)^m|_{R_\lambda^m(K)}\right)^{-1} (\lambda I - K)^m x.$$

Dann ist $y \in R_\lambda^m(K)$ und $(\lambda I - K)^m y = (\lambda I - K)^m x$, d.h. $x - y \in N_\lambda^m(K)$. Damit ist auch gezeigt, dass $X = R_\lambda^m(K) + N_\lambda^m(K)$.

4. Schritt: Wir müssen noch $k = \ell$ zeigen. Ist $k = 0$, d.h. ist

$$N_\lambda^0(K) = N_\lambda^1(K) = \ldots = \{0\},$$

so folgt aus $R_\lambda^\ell(K) + N_\lambda^\ell(K) = X$, dass $R_\lambda^\ell(K) = X$ und damit

$$R_\lambda^0(K) = R_\lambda^1(K) = \ldots = X.$$

Sei z.B. noch $\ell \geq k \geq 1$. Dann gilt

$$X = N_\lambda^\ell(K) + R_\lambda^\ell(K) = N_\lambda^k(K) + R_\lambda^\ell(K),$$

und wegen $R_\lambda^\ell(K) \subseteq R_\lambda^k(K)$ ist $X = N_\lambda^k(K) + R_\lambda^k(K)$. Hieraus folgt schließlich

$$R_\lambda^k(K) = (\lambda I - K)^k X = (\lambda I - K)^k \, R_\lambda^k(K) = R_\lambda^{2k}(K),$$

also ist auch $R_\lambda^k(K) = R_\lambda^{k+1}(K)$. ■

Unter den Bedingungen des Satzes zerfällt also der gesamte Raum X in die Summe $X = R_\lambda^m(K) + N_\lambda^m(K)$ zweier abgeschlossener Teilräume mit Durchschnitt $\{0\}$. Jeder dieser Teilräume wird durch den Operator $\lambda I - K$ in sich überführt; solche Räume heißen auch *invariante Teilräume* von $\lambda I - K$. Die Bedeutung derartiger Zerlegungen besteht darin, dass man, um den Operator $\lambda I - K$ auf dem Raum X zu untersuchen, sich darauf beschränken kann, die Einschränkungen des Operators $\lambda I - K$ auf die beiden Teilräume zu untersuchen. In unserer Situation haben wir gesehen, dass der Operator $(\lambda I - K)|_{R_\lambda^m(K)}$ stetig invertierbar ist, während der Operator $(\lambda I - K)|_{N_\lambda^m(K)}$ *nilpotent* ist (d.h. eine gewisse Potenz dieses Operators ist gleich Null; in unserem Fall ist es gerade die m-te Potenz). Natürlich sind invariante Teilräume für beliebige Operatoren $A \in L(X)$ (nicht notwendig von der Gestalt $\lambda I - K$) von Interesse. Es ist lange Zeit ein offenes Problem gewesen, ob jeder beschränkte Operator auf einem Banachraum X nichttriviale (d.h. $\neq X, \neq \{0\}$) invariante Teilräume besitzt. Erst um 1975 wurde von P. Enflo ein kompliziertes Gegenbeispiel angegeben (welches so kompliziert war, dass es erst 1987 für richtig befunden und veröffentlicht wurde). Inzwischen kennt man auch Operatoren auf ℓ^1 ohne nichttriviale invariante Teilräume. Für Hilberträume ist dieses Problem immer noch ungeklärt.

Wir können nun den Hauptsatz über das Spektrum kompakter Operatoren formulieren.

Satz 4.7 *Sei X ein Banachraum und $K \in K(X)$. Dann gilt:*

(a) $\sigma(K)\backslash\{0\} \subseteq \sigma_p(K)$.

(b) $\sigma(K)$ *ist höchstens abzählbar mit* 0 *als einzig möglichem Häufungspunkt.*

(c) *Falls X unendlich-dimensional ist, gehört* 0 *zum Spektrum von K.*

Die Null kann ebenfalls Eigenwert sein, muss aber nicht. Alle übrigen Punkte im Spektrum sind aber mit Sicherheit Eigenwerte von K, und diese sind isolierte Punkte.

Beweis (a) Sei $\lambda \in \sigma(K)\backslash\{0\}$. Wäre λ kein Eigenwert, so wäre $\operatorname{Ker}(\lambda I - K) = \{0\}$, d.h. $\lambda I - K$ ist injektiv. Dann ist $\{0\} = N_\lambda^0(K) = N_\lambda^1(K)$. Nach Satz 4.6 ist dann $X = R_\lambda^0(K) = R_\lambda^1(K)$, also $\operatorname{Im}(\lambda I - K) = X$. Somit ist $\lambda I - K$ invertierbar, was der Annahme $\lambda \in \sigma(K)$ widerspricht.

(b) Wir zeigen zuerst, dass alle Punkte in $\sigma(K)\backslash\{0\}$ isoliert sind. Sei $\lambda \in \sigma(K)\backslash\{0\}$ und sei m so gewählt, dass

$$X = R_\lambda^m(K) + N_\lambda^m(K), \quad R_\lambda^m(K) \cap N_\lambda^m(K) = \{0\}.$$

Wie wir oben vermerkt haben, ist dann $(\lambda I - K)|_{R_\lambda^m(K)}$ stetig invertierbar, d.h. λ gehört zur Resolventenmenge von $K|_{R_\lambda^m(K)}$. Die Resolventenmenge ist offen; es gibt also eine Umgebung U von λ so, dass $(\mu I - K)|_{R_\lambda^m(K)}$ für alle $\mu \in U$ stetig invertierbar ist. Weiter: da $N_\lambda^m(K)$ endlich-dimensional ist, ist das Spektrum von $K|_{N_\lambda^m(K)}$ eine endliche Menge. Der Punkt λ selbst gehört zu dieser Menge; wegen der Endlichkeit von $\sigma(K|_{N_\lambda^m(K)})$ gibt es jedoch eine Umgebung V von λ, die außer λ keine weiteren Punkte von $\sigma(K|_{N_\lambda^m(K)})$ enthält. Für alle $\mu \in (U \cap V)\backslash\{\lambda\}$ ist daher der Operator $\mu I - K$ invertierbar. Folglich ist jeder Punkt von $\sigma(K)\backslash\{0\}$ ein isolierter Punkt.
Nun ist klar, dass es für jedes $n \in \mathbb{N}$ nur endlich viele Spektrumspunkte mit $|\lambda| \geq 1/n$ geben kann. Gäbe es nämlich unendlich viele, so müssten sie sich (da Spektren kompakt sind) in wenigstens einem Punkt x^* häufen. Dieser Punkt x^* gehört wieder zum Spektrum und es ist $|x^*| \geq 1/n$. Also ist x^* Eigenwert und damit isoliert. Widerspruch. Wegen

$$\sigma(K)\backslash\{0\} = \bigcup_n \{\lambda \in \sigma(K) : |\lambda| \geq 1/n\}$$

ist somit $\sigma(K)$ höchstens abzählbar.

(c) Angenommen, 0 liege nicht im Spektrum von K. Dann ist K invertierbar, und aus $K^{-1}K = I$ folgt die Kompaktheit des identischen Operators. Dieser überführt jede beschränkte Menge in sich selbst. Also muss jede beschränkte Menge (z.B. die Einheitskugel) relativ kompakt sein. Aus dem Rieszschen Lemma wissen wir, dass dies in unendlich-dimensionalen Räumen nicht der Fall ist. ■

Was liefert dieser Satz für eine Fredholmsche Integralgleichung

$$\lambda u(t) - \int_T k(t,s)u(s)\,ds = f(t), \quad t \in T \tag{4.10}$$

die wir auf einem geeigneten Banachraum X von Funktionen auf T und unter geeigneten Voraussetzungen an die Kernfunktion k betrachten, so dass der Integraloperator in (4.10) kompakt wird (also etwa $X = C(T)$, k stetig auf $T \times T$ oder schwach singulär oder $X = L^2(T)$, $k \in L^2(T \times T)$)?
Für $\lambda = 0$ sagt der Satz nur aus, dass der Integraloperator nicht invertierbar ist. Wir wissen aber in diesem Fall (Fredholmsche Integralgleichung erster Art), dass die Voraussetzungen zur Lösung der Gleichung $Ku = f$ (wobei K für den Integraloperator in (4.10) steht) denkbar ungünstig sind. Insbesondere ist i.Allg. nicht einmal der Bildraum $\operatorname{Im} K$ abgeschlossen.
Sei also $\lambda \neq 0$. Dann ist entweder λ in der Resolventenmenge von K (d.h. der Operator $\lambda I - K$ ist stetig invertierbar, und (4.10) hat für jede rechte Seite eine eindeutig bestimmte Lösung), oder λ ist Eigenwert von K. Insbesondere gilt:

wenn die Gleichung $(\lambda I - K)x = 0$ nur die Lösung $x = 0$ besitzt, so ist die Gleichung $(\lambda I - K)u = f$ für jede rechte Seite f eindeutig lösbar. Ist dagegen $\mathrm{Ker}\,(\lambda I - K) \neq \{0\}$, d.h. ist λ Eigenwert von K, so weiß man immerhin noch, dass der Eigenraum $\mathrm{Ker}\,(\lambda I - K)$ endlich-dimensional ist und dass $\mathrm{Im}\,(\lambda I - K)$ abgeschlossen ist. Die Gleichung $(\lambda I - K)u = f$ kann dann nicht mehr für jede rechte Seite f lösbar sein. Sonst wäre nämlich $R^1_\lambda(K) = R^0_\lambda(K) = X$, also auch

$$N^1_\lambda(K) = N^0_\lambda(K) = \{0\}$$

nach Satz 4.7. Dies widerspricht der Annahme, dass λ Eigenwert von K ist. Damit haben wir bereits den ersten Teil des folgenden Satzes bewiesen.

Satz 4.8 (Fredholmsche Alternative) *Sei X ein Banachraum, K kompakt und $\lambda \neq 0$. Dann gilt:*

(a) *Entweder $(\lambda I - K)x = 0$ hat nur die triviale Lösung und die Gleichung $(\lambda I - K)u = f$ ist für alle rechten Seiten eindeutig lösbar, oder die homogene Gleichung $(\lambda I - K)x = 0$ hat nichttriviale Lösungen (Eigenvektoren), und die Gleichung $(\lambda I - K)u = f$ besitzt nicht für jede rechte Seite eine Lösung.*

(b) *Die gleiche Aussage gilt für den adjungierten Operator K'; darüber hinaus ist*

$$\dim \mathrm{Ker}\,(\lambda I - K) = \dim \mathrm{Ker}\,(\lambda I - K').$$

Einen Beweis der Aussage (b) finden Sie in Schröder, FA, Satz 4.1.12. Im Fall $\dim \mathrm{Ker}\,(\lambda I - K) = 0$ ist dieser Beweis einfach. Dann ist nämlich nach (3.14)

$$\mathrm{Ker}\,(\lambda I - K') = \big(\mathrm{Im}\,(\lambda I - K)\big)^\perp = X^\perp = \{0\},$$

also $\dim \mathrm{Ker}\,(\lambda I - K') = 0$. Umgekehrt folgt aus $\dim \mathrm{Ker}\,(\lambda I - K') = 0$ ebenfalls aus (3.14)

$$\mathrm{Ker}\,(\lambda I - K) = {}^\perp\big(\mathrm{Im}\,(\lambda I - K')\big) = {}^\perp(X') = \{0\},$$

also $\dim \mathrm{Ker}\,(\lambda I - K) = 0$. Der allgemeine Fall wird durch Reduktion auf diesen Spezialfall bewiesen. ■

4.3 Hilbert-Schmidt-Theorie

Falls $X = H$ ein Hilbertraum und K ein kompakter und selbstadjungierter Operator auf H ist, lassen sich die Aussagen des vorigen Abschnittes noch wesentlich ergänzen. Ein Beispiel für einen solchen Operator ist

$$K : L^2(X) \to L^2(X), \quad (Ku)(t) = \int_X k(t,s)u(s)ds$$

mit $k \in L^2(X \times X)$ und $k(t,s) = \overline{k(s,t)}$. Die Spektraltheorie für kompakte selbstadjungierte Operatoren beruht auf folgendem Resultat.

Satz 4.9 *Sei H ein Hilbertraum und $K \in K(H)$ selbstadjungiert. Dann ist wenigstens eine der Zahlen $\|K\|$ und $-\|K\|$ ein Eigenwert von K.*

Beweis Das Spektrum des selbstadjungierten Operators K liegt nach Satz 4.4 (b) auf der reellen Achse, und außerdem stimmen Norm $\|K\|$ und Spektralradius $r(K)$ überein. Da Spektren kompakt sind, gibt es wenigstens einen Punkt $\lambda \in \sigma(K)$ mit $|\lambda| = r(K)$. Die einzigen reellen Zahlen mit dieser Eigenschaft sind $r(K) = \|K\|$ und $-r(K) = -\|K\|$. Folglich gehört wenigstens eine der Zahlen $\|K\|$ und $-\|K\|$ zum Spektrum von K. Nach Satz 4.7 (a) ist diese Zahl notwendig ein Eigenwert (im Fall $\|K\| = 0$ ist dies wegen $K = 0$ offenbar auch richtig). ∎

Alternativer Beweis Sei $(x_n) \subset H$ eine Folge von Vektoren mit $\|x_n\| = 1$ und so, dass $\|Kx_n\| \to \|K\|$. Wegen

$$\begin{aligned}
0 &\leq \langle K^2x_n - \|Kx_n\|^2x_n,\, K^2x_n - \|Kx_n\|^2x_n\rangle \\
&= \langle K^2x_n, K^2x_n\rangle - \|Kx_n\|^2 \langle K^2x_n, x_n\rangle - \|Kx_n\|^2 \langle x_n, K^2x_n\rangle + \|Kx_n\|^4\langle x_n, x_n\rangle \\
&= \|K^2x_n\|^2 - 2\|Kx_n\|^4 + \|Kx_n\|^4 \\
&\leq \|K\|^4 - 2\|Kx_n\|^4 + \|Kx_n\|^4 \to 0
\end{aligned}$$

folgt die Konvergenz von $\left\|K^2x_n - \|Kx_n\|^2x_n\right\| \to 0$. Da weiter

$$K^2x_n - \|K\|^2x_n = \underbrace{K^2x_n - \|Kx_n\|^2x_n}_{\to 0} + \underbrace{(\|Kx_n\|^2 - \|K\|^2)x_n}_{\to 0} \tag{4.11}$$

ist, gilt auch $K^2x_n - \|K\|^2x_n \to 0$. Wegen der Kompaktheit von K läßt sich aus (x_n) eine Teilfolge auswählen, so dass die Folge $(K^2x_{n_r})_{r\geq 1}$ konvergiert. Dann konvergiert aber auch die Folge $(\|K\|^2x_{n_r})$. Da die Aussage des Satzes für $K = 0$ offenbar richtig ist, können wir $K \neq 0$ annehmen und erhalten die Konvergenz der Folge (x_{n_r}). Ihren Grenzwert nennen wir x. Wegen $\|x_{n_r}\| = 1$ ist auch $\|x\| = 1$ (Stetigkeit der Norm). Nun ist aber

$$K^2x - \|K\|^2x = \lim_{r\to\infty}(K^2x_{n_r} - \|K\|^2x_{n_r}) = 0$$

wegen (4.11), d.h.

$$(K + \|K\|I)(K - \|K\|I)x = 0.$$

Falls $(K - \|K\|I)x = 0$, so ist x Eigenvektor zum Eigenwert $\|K\|$; ist jedoch $(K - \|K\|I)x \neq 0$, so ist $(K - \|K\|I)x$ Eigenvektor zum Eigenwert $-\|K\|$. In jedem Fall ist eine der beiden Zahlen $\pm\|K\|$ ein Eigenwert. ∎

Wir können nun den Spektralsatz für kompakte selbstadjungierte Operatoren beweisen. Er besagt, dass man jeden selbstadjungierten kompakten Operator (genauso wie jede selbstadjungierte Matrix $A \in \mathbb{C}^{n\times n}$) auf Diagonalform bringen kann. Genauer:

Satz 4.10 *Sei H ein separabler Hilbertraum und $K \in K(H)$ selbstadjungiert. Dann ist $\sigma_p(K)$ höchstens abzählbar mit 0 als einzig möglichem Häufungspunkt, und es gilt*

$$\max\{|\lambda| : \lambda \in \sigma_p(K)\} = \|K\|.$$

Sind $\lambda_1, \lambda_2, \ldots$ die verschiedenen Eigenwerte von K und P_n die Orthoprojektoren von H auf $\operatorname{Ker}(K - \lambda_n I)$, so ist $K = \sum_{n\geq 1} \lambda_n P_n$ (Konvergenz im Sinne der Norm). Insbesondere besitzt H eine Hilbertraumbasis S, die aus Eigenvektoren von K besteht, und für jedes $x = \sum_{y\in S}\langle x, y\rangle y \in H$ gilt

$$Kx = \sum_{y\in S} \lambda_y \langle x, y\rangle y, \tag{4.12}$$

wobei λ_y der zu $y \in S$ gehörende Eigenwert ist.

Ist H separabel, so ist die Basis $S = (e_1, e_2, \ldots)$ höchstens abzählbar, und wir können uns die Elemente $x \in H$ als Folgen $(\langle x, e_1\rangle, \langle x, e_2\rangle, \ldots) \in \ell^2$ vorstellen (Satz von Fischer/Riesz). Ist λ_{e_i} der zu e_i gehörende Eigenwert, so zeigt (4.12), dass K wie die Diagonalmatrix $\begin{pmatrix} \lambda_{e_1} & & \\ & \lambda_{e_2} & \\ & & \ddots \end{pmatrix}$ wirkt.

Beweis Für $K = 0$ ist die Aussage klar. Sei $K \neq 0$. Wir wählen $\lambda_1 \in \sigma_p(K)$ so, dass $|\lambda_1| = \|K\|$ und setzen $N_1 := \operatorname{Ker}(K - \lambda_1 I)$ sowie $H_2 := N_1^\perp$. Dann gilt $KN_1 \subseteq N_1$ und $KH_2 \subseteq H_2$: Für $x \in N_1$, d.h. $(K - \lambda_1 I)x = 0$ ist nämlich auch $(K - \lambda_1 I)Kx = K(K - \lambda_1 I)x = 0$, also $Kx \in N_1$, und für $x \in H_2$, also $\langle x, y\rangle = 0$ für alle $y \in N_1$, ist auch $\langle Kx, y\rangle = \langle x, Ky\rangle = 0$ für alle $y \in N_1$, d.h. $Kx \in H_2$.

Falls $H_2 \neq \{0\}$, setzen wir $K_2 = K|_{H_2}$. Wie wir gerade gesehen haben, ist $K_2 \in L(H_2)$. Es gilt aber sogar $K_2 \in K(H_2)$, und K_2 ist wieder selbstadjungiert.

Ist $K_2 \neq 0$, wählen wir wie oben $\lambda_2 \in \sigma_p(K_2)$ mit $|\lambda_2| = \|K_2\|$ und setzen $N_2 = \operatorname{Ker}(K_2 - \lambda_2 I)$. Zunächst ist λ_2 ein Eigenwert von $K_2 \in L(H_2)$. Ist $x \in H_2\backslash\{0\}$ aber ein zugehöriger Eigenvektor, d.h. $(K_2 - \lambda_2 I)x = 0$, so ist offenbar auch $(K - \lambda_2 I)x = 0$, d.h. λ_2 ist auch Eigenwert von K. Man überlegt sich leicht ($H = N_1 \oplus N_1^\perp = N_1 \oplus H_2$ benutzen!), dass $\lambda_2 \neq \lambda_1$ und $\operatorname{Ker}(K - \lambda_2 I) = \operatorname{Ker}(K_2 - \lambda_2 I)$ ist. Schließlich ist $|\lambda_2| \leq |\lambda_1|$, da ja $\|K_2\| \leq \|K\|$.

Wir setzen nun $H_3 := (N_1 \oplus N_2)^\perp$ (die orthogonale Summe $\oplus$ ist gerechtfertigt, da Eigenvektoren zu verschiedenen Eigenwerten aufeinander senkrecht stehen; vgl. Satz 4.4) und fahren so fort. Im Ergebnis erhalten wir eine (möglicherweise abbrechende) Folge $(\lambda_n) \subset \sigma_p(K)$ mit

(i) $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n| \geq \ldots$,

(ii) $N_n = \operatorname{Ker}(K - \lambda_n I)$, $H_{n+1} = (N_1 \oplus \ldots \oplus N_n)^\perp$, $|\lambda_{n+1}| = \|K|_{H_{n+1}}\| = \|K_{n+1}\|$.

Diese Konstruktion bricht ab, wenn $H_{n+1} = \{0\}$ oder $\|K|_{H_{n+1}}\| = 0$. Dabei kann der Fall $H_{n+1} = \{0\}$ nur eintreten, wenn der Ausgangsraum endlich-dimensional ist. Der zweite Fall bedeutet, dass es nur endlich viele Eigenwerte ungleich 0 gibt und dass $H_{n+1} = \operatorname{Ker} K$.

Sei nun P_n der Orthoprojektor auf $N_n = \operatorname{Ker}(K - \lambda_n I)$. Wir betrachten die Differenz $K - \sum_{j=1}^n \lambda_j P_j$. Ist $x \in N_k$ für ein k zwischen 1 und n, so ist

$$\Big(K - \sum_{j=1}^{n} \lambda_j P_j\Big)\, x = Kx - \lambda_k P_k x = (K - \lambda_k I)x = 0$$

(man beachte, dass $P_i P_k = P_k P_i = 0$ für $i \neq k$). Also ist $N_1 + \ldots + N_n \subset \operatorname{Ker}\big(K - \sum_{j=1}^n \lambda_j P_j\big)$. Ist dagegen $x \in (N_1 + \ldots + N_n)^\perp = H_{n+1}$, so ist $P_k x = 0$ für alle $k \le n$, also

$$\Big(K - \sum_{j=1}^{n} \lambda_j P_j\Big)x = Kx = K_{n+1}x,$$

und folglich

$$\Big\|K - \sum_{j=1}^{n} \lambda_j P_j\Big\| = \Big\|\Big(K - \sum_{j=1}^{n} \lambda_j P_j\Big)\Big|_{H_{n+1}}\Big\| = \|K_{n+1}\| = |\lambda_{n+1}|.$$

Wie wir aus dem vorigen Abschnitt (Satz 4.7) wissen, gilt aber $|\lambda_n| \to 0$, woraus die behauptete Normkonvergenz folgt.

Für die letzte Behauptung wählen wir Orthonormalbasen S_n von N_n ($n \ge 1$) sowie S_0 von $N_0 := \operatorname{Ker} K$ und setzen $S = \bigcup_{n \ge 0} S_n$. Wir wollen zeigen, dass S eine Orthonormalbasis von H ist. Aus Satz 2.44 in Abschnitt 2.5 wissen wir, dass

$$H = \operatorname{Ker} K \oplus \overline{\operatorname{Im} K} \quad (\text{beachte: } K = K^*).$$

Da S_0 eine ONB von $\operatorname{Ker} K$ ist, genügt es zu zeigen, dass $\bigcup_{n \ge 1} S_n$ eine ONB von $\overline{\operatorname{Im} K}$ ist. Offenbar ist $\bigcup_{n \ge 1} S_n$ ein Orthonormalsystem in $\overline{\operatorname{Im} K}$ (ein Orthonormalsystem liegt vor, da Eigenvektoren zu verschiedenen Eigenwerten zueinander orthogonal sind, und $\bigcup_{n \ge 1} S_n$ liegt in $\operatorname{Im} K$, da jeder Eigenvektor zu einem Nichtnulleigenwert von K im Bild von K liegt. Es folgt ja aus $(K - \lambda I)x = 0$, dass $x = \frac{1}{\lambda} Kx$). Wir erhalten daher die Behauptung, wenn wir zeigen können, dass die Parsevalsche Gleichung erfüllt ist. Nun ist $Kx = \sum_{j \ge 1} \lambda_j P_j x$ für jedes $x \in H$ und daher

$$\begin{aligned}
\|Kx\|^2 &= \sum_{j \ge 1} |\lambda_j|^2 \, \|P_j x\|^2 &&= \sum_{j \ge 1} |\lambda_j|^2 \sum_{y \in S_j} |\langle x, y\rangle|^2 &&= \sum_{y \in \bigcup_{n \ge 1} S_n} |\lambda_y|^2 \, |\langle x, y\rangle|^2 \\
&= \sum_{y} |\langle x, \lambda_y y\rangle|^2 &&= \sum_{y} |\langle x, Ky\rangle|^2 &&= \sum_{y} |\langle Kx, y\rangle|^2,
\end{aligned}$$

d.h. die Parsevalsche Gleichung ist erfüllt. Damit ist S eine ONB von H, und wir erhalten sofort die Identität (4.12): Für $x = \sum_{y\in S}\langle x,y\rangle y$ ist

$$Kx = K\sum_{y\in S}\langle x,y\rangle y = \sum_{y\in S}\langle x,y\rangle Ky = \sum_{y\in S}\lambda_y\langle x,y\rangle y. \qquad \blacksquare$$

Die Bestimmung des vollständigen Systems aller Eigenwerte und der zugehörigen Eigenvektoren eines kompakten selbstadjungierten Operators K ist i.a. natürlich keine leichte Aufgabe. Sind uns diese jedoch einmal bekannt, so lassen sich damit die Lösungen von Fredholmschen (Integral-)Gleichungen der 2. Art (d.h. $(\lambda I - K)u = f$ mit $\lambda \neq 0$) explizit beschreiben. Wir geben diese Resultate noch an.

Satz 4.11 *Sei H ein separabler Hilbertraum, $K \in K(H)$ selbstadjungiert, und (e_n) sei eine Orthonormalbasis von H, bestehend aus Eigenvektoren e_n von K mit zugehörigen Eigenwerten λ_n. Sei $\lambda \notin \sigma_p(K)\cup\{0\}$. Dann ist die Gleichung*

$$(\lambda I - K)u = f \tag{4.13}$$

für jede rechte Seite $f \in H$ eindeutig lösbar, und es ist

$$u = (\lambda I - K)^{-1} f = \sum_n \frac{\langle f, e_n\rangle}{\lambda - \lambda_n} e_n.$$

Beweis Wir wissen bereits, dass $\sigma(K) \subseteq \sigma_p(K)\cup\{0\}$. Für $\lambda \notin \sigma_p(K)\cup\{0\}$ ist die Gleichung (4.13) also stets eindeutig lösbar. Für die Darstellung der Lösung machen wir den Ansatz $u = \sum u_n e_n$ und entwickeln f in seine Fourierreihe: $f = \sum_n \langle f, e_n\rangle e_n$. Eingesetzt in (4.13) ergibt sich

$$\begin{aligned}(\lambda I - K)\sum_n u_n e_n &= \sum_n u_n(\lambda I - K)e_n \\ &= \sum_n u_n(\lambda - \lambda_n)e_n \overset{!}{=} \sum_n \langle f, e_n\rangle e_n.\end{aligned}$$

Koeffizientenvergleich liefert $u_n(\lambda - \lambda_n) = \langle f, e_n\rangle$ für alle n. ■

Satz 4.12 *Es seien H, K, e_n, λ_n wie in Satz 4.11; jedoch sei λ nun ein Nichtnulleigenwert von K, d.h. $\lambda = \lambda_k \neq 0$. Dann ist die Gleichung* (4.13) *genau dann lösbar, wenn f senkrecht auf* $\operatorname{Ker}(\lambda_k I - K)$*, d.h. senkrecht zu allen Eigenvektoren von K steht, für die $\lambda = \lambda_k$ Eigenwert ist. Ist diese Bedingung für $f \in H$ erfüllt, so werden alle Lösungen der Gleichung* (4.13) *beschrieben durch*

$$u = \sum_{n:\lambda_n\neq\lambda_k} \frac{\langle f, e_n\rangle}{\lambda_k - \lambda_n} e_n + \sum_{n:\lambda_n=\lambda_k} \alpha_n e_n \tag{4.14}$$

mit beliebigen Zahlen $\alpha_n \in \mathbb{C}$

Beweis Wie im Beweis von Satz 4.11 liefert der Ansatz $u = \sum u_n e_n$, dass

$$u_n(\lambda - \lambda_n) = \langle f, e_n \rangle \quad \text{für alle } n.$$

Ist $\lambda = \lambda_k$, so muss notwendigerweise $\langle f, e_n \rangle = 0$ für alle zu λ_k gehörenden Eigenvektoren e_n sein, damit überhaupt Lösungen existieren können. Ist diese Bedingung erfüllt, läßt sich u_n für die entsprechenden n beliebig wählen. Wir gelangen so zur Formel (4.14). Umgekehrt überprüft man durch Einsetzen leicht, dass jedes u wie in (4.14) tatsächlich die Gleichung $(\lambda I - K)u = f$ löst. ■

Bei Kenntnis der Eigenwerte und Eigenvektoren von K lassen sich auch im Fall $\lambda = 0$ noch notwendige und hinreichende Lösbarkeitskriterien sowie eine Darstellungsformel für die Lösung der Gleichung $Ku = f$ angeben. Dies soll in der Übung diskutiert werden.

Im Falle eines Integraloperators mit stetiger oder differenzierbarer Kernfunktion sind die Eigenfunktionen ebenfalls stetig oder differenzierbar (Sätze über Stetigkeit bzw. Differenzierbarkeit von parameterabhängigen Integralen). Man ist in diesem Fall an stärkeren Konvergenzaussagen für die Entwicklung der Lösung nach Eigenfunktionen interessiert als nur an L^2-Konvergenz, insbesondere an der gleichmäßigen Konvergenz. Eine Aussage in diese Richtung macht

Satz 4.13 (Mercer) *Sei $X \subseteq \mathbb{R}^n$ kompakt, k stetig auf $X \times X$ und symmetrisch, d.h. $k(t,s) = \overline{k(s,t)}$, K sei der Integraloperator mit der Kernfunktion k, und wie bisher seien (λ_n) bzw. (e_n) die Folgen der Eigenwerte bzw. Eigenvektoren von K. Ist $\langle Kf, f \rangle \geq 0$ für alle $f \in L^2(X)$, so läßt sich k in die auf $X \times X$ gleichmäßig und absolut konvergente Reihe*

$$k(t,s) = \sum_{n=1}^{\infty} \lambda_n e_n(t) \, \overline{e_n(s)}$$

entwickeln. Insbesondere gilt die "Spurformel"

$$\sum_{n=1}^{\infty} \lambda_n = \int_X k(t,t) \, dt.$$

Beweis ↗ Schröder, FA, Satz 1.4.10. Die Stetigkeit der Eigenfunktionen e_n folgt aus Sätzen über die Stetigkeit von Parameterintegralen. ■

Die Spurformel ist ein Analogon zu der von Matrizen $A = (a_{ij})_{i,j=1}^n \in \mathbb{R}^{n \times n}$ bekannten Identität

$$\sum_{i=1}^{n} (\text{Eigenwerte von } A) = \sum_{i=1}^{n} a_{ii}.$$

4.4 Fredholmoperatoren

Wir betrachten nun eine Klasse von Operatoren, die die Fredholmschen Integraloperatoren 2. Art umfasst, aber wesentlich größer ist. Und zwar wollen wir Operatoren untersuchen, die “fast” invertierbar sind. Wie wir wissen, ist ein linearer und beschränkter Operator $A : X \to Y$ zwischen Banachräumen X, Y genau dann invertierbar, wenn sein Kern gleich $\{0\}$ und sein Bild ganz Y ist. Sind X und Y unendlich-dimensional, so kann man sich auf den Standpunkt stellen, dass alles, was auf *endlich-dimensionalen* Teilräumen passiert, eher unwesentlich ist. So ist es sinnvoll, Operatoren zu betrachten, deren Kern zwar nicht $\{0\}$, aber eben nur ein *endlich*-dimensionaler Teilraum von X ist. Oder man betrachtet Operatoren $A \in L(X, Y)$, deren Bild $\operatorname{Im} A$ zwar nicht ganz Y ist, für die es aber einen *endlich*-dimensionalen Teilraum Y_0 von Y gibt, so dass $\operatorname{Im} A \cap Y_0 = \{0\}$ und $\operatorname{Im} A + Y_0 = Y$. Ein solcher Teilraum heißt auch *Kokern* von A. Der Raum Y_0 ist durch diese beiden Bedingungen noch nicht eindeutig bestimmt. Jedoch liegt seine Dimension durch diese Forderungen fest. Seien X, Y, Z Banachräume über $\mathbb{C}$.

Definition 4.14 *Ein Operator $A \in L(X, Y)$ heißt* Fredholmoperator*, wenn er sowohl einen endlich-dimensionalen Kern als auch einen endlich-dimensionalen Kokern besitzt. Ist dies der Fall, so heißt die Zahl*

$$\operatorname{ind} A := \dim \operatorname{Ker} A - \dim \operatorname{Coker} A$$

der Index *von A.*

Man kann zeigen, dass $\operatorname{Im} A$ automatisch abgeschlossen ist, wenn A einen endlich-dimensionalen Kokern besitzt. Oft nimmt man die Bedingung $\operatorname{Im} A = \overline{\operatorname{Im} A}$ (d.h. A normal auflösbar) in die Definition der Fredholmeigenschaft mit auf. Der Index eines Fredholmoperators erlaubt auch Aussagen über die Lösbarkeit der Gleichung $Au = f$. Ist $\operatorname{ind} A > 0$, so ist diese Gleichung mit Sicherheit nicht eindeutig lösbar. Ist dagegen $\operatorname{ind} A < 0$, so ist $Au = f$ garantiert nicht für alle rechten Seiten lösbar.

Beispiel 0: Sei $A : \mathbb{C}^n \to \mathbb{C}^m$ linear. Dann ist nach dem Dimensionssatz

$$\dim \operatorname{Ker} A + \dim \operatorname{Im} A = n,$$

und nach Definition des Kokerns

$$\dim \operatorname{Im} A + \dim \operatorname{Coker} A = m.$$

Aus diesen beiden Gleichungen folgt

$$\dim \operatorname{Ker} A - \dim \operatorname{Coker} A = n - m,$$

also $\operatorname{ind} A = n - m$. Der Index von A ist in diesem Fall also nur von den Raumdimensionen abhängig. ■

Beispiel 1: Sei X ein Banachraum, $K \in K(X)$ und $A = I + K$. Dann ist $\operatorname{Im} A$ abgeschlossen und $\dim \operatorname{Ker} A < \infty$ (Satz 4.5). Außerdem ist

$$\dim \operatorname{Ker} A = \dim \operatorname{Ker} A'$$

(Satz 4.8), und es gilt

$$(\operatorname{Im} A)^{\perp} = \{f \in X' : f(y) = 0 \text{ für alle } y \in \operatorname{Im} A\} \stackrel{!}{=} \operatorname{Ker} A'$$

(vgl. Abschnitt 3.2.5). Es gibt also genau $\dim \operatorname{Ker} A$ linear unabhängige Funktionale, die auf ganz $\operatorname{Im} A$ verschwinden. Da der duale Raum eines endlich-dimensionalen Raumes die gleiche Dimension hat wie der Raum selbst, folgt hieraus leicht, dass die Kokerndimension von $I + K$ gleich $\dim \operatorname{Ker} A$ ist (beachte, dass A nach Satz 4.6 einen endlichen Kokern besitzt.)

Fazit *Operatoren der Gestalt $I + K$ mit K kompakt sind Fredholmoperatoren, und ihr Index ist gleich* 0. ■

Beispiel 2: Der Differentiationsoperator

$$D : C^1[0,1] \to C[0,1], \quad f \mapsto f'$$

ist ein Fredholmoperator und hat den Index 1. ■

Beispiel 3: Die Verschiebungsoperatoren

$$\begin{aligned} V : \ell^2 \to \ell^2, \quad (x_1, x_2, x_3, \ldots) \quad &\mapsto \quad (0, x_1, x_2, x_3, \ldots), \\ V^* : \ell^2 \to \ell^2, \quad (x_1, x_2, x_3, \ldots) \quad &\mapsto \quad (x_2, x_3, x_4, \ldots) \end{aligned}$$

sind Fredholmsch und haben die Indizes -1 bzw. 1. ■

Auf den ersten Blick scheint es, als wären $\dim \operatorname{Ker} A$ und $\dim \operatorname{Coker} A$ die wesentlichen Größen und $\operatorname{ind} A$ nur eine wenig interessante abgeleitete Größe. Es zeigt sich aber, dass der Index – im Gegensatz etwa zu $\dim \operatorname{Ker} A$ – *invariant bezüglich kleiner Störungen* ist. Diese Eigenschaft ist grundlegend für die gesamte Theorie der Fredholmoperatoren.

Im Weiteren benötigen wir oft folgende Eigenschaft. Ist X ein Banachraum und X_1 ein endlich-dimensionaler Teilraum von X, so gibt es einen abgeschlossenen Teilraum X_2 von X so, dass

$$X_1 + X_2 = X \quad \text{und} \quad X_1 \cap X_2 = \{0\}. \tag{4.15}$$

Außerdem gilt: Sind X_1 und X_2 lineare Teilräume eines Banachraums und ist X_1 endlich-dimensional, so ist X_2 genau dann abgeschlossen, wenn $X_1 + X_2$ abgeschlossen ist. Insbesondere ist der Bildraum eines Fredholmoperators stets abgeschlossen. (Beweis $\nearrow$ Übung). Sind X_1, X_2 zwei abgeschlossene Teilräume von X wie in (4.15), so schreiben wir auch $X = X_1 \oplus X_2$. In diesem Fall besitzt jedes Element $x \in X$ eine Darstellung $x = x_1 + x_2$ mit $x_1 \in X_1$, $x_2 \in X_2$, und diese Darstellung ist eindeutig. Aus $x_1 + x_2 = y_1 + y_2$ mit $x_1, y_1 \in X_1$ und $x_2, y_2 \in X_2$ folgt nämlich

$$X_1 \ni x_1 - y_1 = y_2 - x_2 \in X_2,$$

und aus $X_1 \cap X_2 = \{0\}$ folgt weiter $x_1 - y_1 = 0 = y_2 - x_2$. Durch

$$P : X \to X_1, \quad x \mapsto x_1$$

wird eine lineare Abbildung definiert. Diese heißt *Projektion von X auf X_1 parallel zu X_2* und ist stetig (Satz vom abgeschlossenen Graphen).

Wir betrachten zunächst eine größere Klasse von Operatoren, die *Semi-Fredholmoperatoren*, von denen wir neben der Abgeschlossenheit des Bildes $\operatorname{Im} A$ noch die Endlichkeit *einer* der Zahlen $\dim \operatorname{Ker} A$ oder $\dim \operatorname{Coker} A$ verlangen. Für solche Operatoren definiert man den Index wie oben, muss nun aber auch die Werte $\pm\infty$ zulassen.

Lemma 4.15 *Folgende Eigenschaften sind für einen Operator $A \in L(X, Y)$ äquivalent:*

(a) $\operatorname{Im} A$ *ist abgeschlossen und* $\operatorname{Ker} A$ *ist endlich-dimensional.*

(b) *Es gibt einen Projektor $P \in L(X)$ mit endlich-dimensionalem Bild und ein $C > 0$ so, dass*
$$\|x\| \le C(\|Ax\| + \|Px\|) \quad \forall x \in X. \tag{4.16}$$

Ist (4.16) *erfüllt, so gilt* $\dim \operatorname{Ker} A \le \dim \operatorname{Im} P$.

Im Fall $P = 0$ ist uns dieses Resultat bereits bekannt (Folgerung 3.10).

Beweis (a) $\Rightarrow$ (b): Da $\operatorname{Ker} A$ endlich-dimensional ist, können wir X zerlegen in $X = \operatorname{Ker} A \oplus X_2$ mit einem abgeschlossenen Teilraum X_2 von X. Sei P der Projektor von X auf $\operatorname{Ker} A$ parallel zu X_2. Der Operator $A_2 := A|_{X_2} : X_2 \to \operatorname{Im} A$ ist invertierbar, und nach dem Satz von Banach ist sein Inverser $(A_2)^{-1} : \operatorname{Im} A \to X_2$ beschränkt, etwa durch C_2. Für alle $x_2 \in X_2$ gilt dann

$$\|x_2\| \le \|(A_2)^{-1} A_2 x_2\| \le C_2 \|A_2 x_2\|.$$

Für beliebiges $x \in X$ schreiben wir $x = x_1 + x_2$ mit $x_1 \in \operatorname{Ker} A$ und $x_2 \in X_2$ und erhalten

$$\|x\| \le \|x_1\| + \|x_2\| \le \|Px\| + C_2 \|A_2 x_2\| = \|Px\| + C_2 \|Ax\|$$

wegen $A_2x_2 = Ax_2 = A(x_1 + x_2) = Ax$. Mit $C := \max\{1, C_2\}$ wird schließlich $\|x\| \leq C(\|Ax\| + \|Px\|)$ für alle $x \in X$.

(b) $\Rightarrow$ (a): Für alle $x \in \operatorname{Ker} A$ folgt aus (4.16), dass

$$\|x\| \leq C\|Px\|.$$

Die Einschränkung von P auf $\operatorname{Ker} A$ ist also injektiv. Zusammen mit der Endlichdimensionalität von $\operatorname{Im} P$ zeigt dies, dass

$$\dim \operatorname{Ker} A = \dim \operatorname{Im} (P|_{\operatorname{Ker} A}) \leq \dim \operatorname{Im} P.$$

Um die Abgeschlossenheit des Bildes von A zu zeigen, zerlegen wir X in die Summe $X = \operatorname{Im} P \oplus \operatorname{Ker} P$. Für alle $x \in \operatorname{Ker} P$ ist

$$\|x\| \leq C\|Ax\|,$$

woraus (wie wir bereits wissen) folgt, dass $A(\operatorname{Ker} P)$ abgeschlossen ist. Da $\operatorname{Im} P$ endlich-dimensional ist, unterscheidet sich $\operatorname{Im} A$ vom abgeschlossenen Raum $A(\operatorname{Ker} P)$ nur um einen endlich-dimensionalen Raum (einer Dimension $\leq \dim \operatorname{Im} P$). Dann ist aber $\operatorname{Im} A$ selbst abgeschlossen ($\nearrow$ Übung). ■

Satz 4.16 (Gohberg/Krein) *Sei $A \in L(X, Y)$ ein Operator mit abgeschlossenem Bild und endlich-dimensionalem Kern. Dann gilt: Hat $S \in L(X, Y)$ eine hinreichende kleine Norm, so ist auch* $\operatorname{Im} (A+S)$ *abgeschlossen und* $\dim \operatorname{Ker} (A+S) < \infty$. *Weiter ist in diesem Fall*

$$\dim \operatorname{Ker} (A + S) \leq \dim \operatorname{Ker} A, \quad \operatorname{ind}(A + S) = \operatorname{ind} A.$$

Beweisidee Sei $\operatorname{Im} A$ abgeschlossen und $\dim \operatorname{Ker} A < \infty$. Wie in Beweis (a) $\Rightarrow$ (b) des vorigen Lemmas wählen wir eine Konstante $C > 0$ und einen stetigen Projektor P von X auf $\operatorname{Ker} A$ so, dass

$$\|x\| \leq C(\|Ax\| + \|Px\|) \quad \text{für alle } x \in X.$$

Ist nun $S \in L(X, Y)$ ein Operator mit $\|S\| < 1/C$, so gilt für alle $x \in X$

$$\|(A + S)x\| + \|Px\| \geq \|Ax\| + \|Px\| - \|Sx\| \geq \frac{1}{C}\|x\| - \|S\| \, \|x\|,$$

woraus mit $C_2 := \left(\frac{1}{C} - \|S\|\right)^{-1}$ folgt

$$\|x\| \leq C_2\big(\|(A + S)x\| + \|Px\|\big) \quad \forall x \in X.$$

Mit Lemma 4.15 folgt hieraus, dass $\operatorname{Im} (A+S)$ abgeschlossen, $\operatorname{Ker} (A+S)$ endlich-dimensional und

$$\dim \operatorname{Ker} (A + S) \leq \dim \operatorname{Im} P = \dim \operatorname{Ker} A$$

ist. Wesentlich aufwändiger ist der Beweis der Index-Identität. Diesen finden Sie z.B. in Schröder, Satz 4.2.4. ∎

Wir sehen uns nun einige der bemerkenswerten Konsequenzen von Satz 4.16 und insbesondere der Index-Identität an. Den folgenden Satz kann man aus Satz 4.16 durch Übergang zum adjungierten Operator herleiten:

Satz 4.17 *Sei $A \in L(X,Y)$ ein Operator mit abgeschlossenem Bild und endlich-dimensionalem Kokern. Dann hat für hinreichend kleines $S \in L(X,Y)$ auch $A+S$ ein abgeschlossenes Bild und einen endlich-dimensionalen Kokern. Weiter gilt*

$$\dim \operatorname{Coker}(A+S) \le \dim \operatorname{Coker} A, \quad \operatorname{ind}(A+S) = \operatorname{ind} A.$$

Durch Kombination von Satz 4.16 und 4.17 folgt sofort

Folgerung 4.18 (Dieudonné) *Die Menge der Fredholmoperatoren ist offen in $L(X,Y)$, und die Funktion $A \mapsto \operatorname{ind} A$ ist stetig auf dieser Menge.*

Folgerung 4.19 (Atkinson) *Sind $A \in L(X,Y)$ und $B \in L(Y,Z)$ Fredholmoperatoren, so ist auch $BA \in L(X,Z)$ ein Fredholmoperator, und es gilt*

$$\operatorname{ind} BA = \operatorname{ind} A + \operatorname{ind} B.$$

Beweis Wir betrachten A als Operator auf $\operatorname{Ker} BA$. Dann ist

$$\operatorname{Im}(A|_{\operatorname{Ker} BA}) = A(\operatorname{Ker} BA) \subseteq \operatorname{Ker} B, \quad \operatorname{Ker}\left(A|_{\operatorname{Ker} BA}\right) = \operatorname{Ker} A,$$

und folglich

$$\dim \operatorname{Ker} BA \le \dim \operatorname{Ker} A + \dim \operatorname{Ker} B$$

(man beachte, dass die Identität $\dim \operatorname{Ker} A + \dim \operatorname{Im} A = \dim X$ für jeden linearen Operator $A : X \to Y$ gilt).

Ähnlich zeigt man, dass $\dim \operatorname{Coker} BA \le \dim \operatorname{Coker} A + \dim \operatorname{Coker} B$, so dass BA ein Fredholmoperator ist. Zum Beweis der Indexformel betrachten wir die Operatorfunktion

$$C(t) := \begin{pmatrix} I_Y & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} \cos t \cdot I_Y & -\sin t\, I_Y \\ \sin t \cdot I_Y & \cos t\, I_Y \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & I_Y \end{pmatrix} \quad \text{mit } t \in [0, \frac{\pi}{2}]$$

von Operatoren von $X \times Y$ nach $Y \times Z$. Die Operatoren $\begin{pmatrix} A & 0 \\ 0 & I_Y \end{pmatrix}$ und $\begin{pmatrix} I_Y & 0 \\ 0 & B \end{pmatrix}$ sind Fredholmsch, und der mittlere Operator ist wegen

$$\det \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} = \cos^2 t + \sin^2 t = 1$$

sogar invertierbar. Also sind alle Operatoren $C(t)$ Fredholmsch. Weiter ist $t \mapsto C(t)$ eine stetige Funktion. Wegen Folgerung 4.18 ist dann auch $t \mapsto \operatorname{ind} C(t)$ eine stetige Funktion. Da der Index aber nur ganzzahlige Werte annimmt, muss $\operatorname{ind} C(t)$ konstant sein. Insbesondere ist also $\operatorname{ind} C(0) = \operatorname{ind} C(\pi/2)$. Nun ist aber

$$C(0) = \begin{pmatrix} I & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & I \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

sowie

$$C(\frac{\pi}{2}) = \begin{pmatrix} I & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & I \end{pmatrix} = \begin{pmatrix} 0 & -I \\ BA & 0 \end{pmatrix}$$

und daher

$$\operatorname{ind} A + \operatorname{ind} B = \operatorname{ind} C(0) = \operatorname{ind} C\Big(\frac{\pi}{2}\Big) = \operatorname{ind}(BA). \qquad \blacksquare$$

Folgerung 4.20 (Yood) *Sei $A \in L(X, Y)$ ein Fredholmoperator und $K \in K(X, Y)$ kompakt. Dann ist $A + K$ Fredholmsch und* $\operatorname{ind}(A + K) = \operatorname{ind} A$.

Wir stützen den Beweis auf die folgende Variante von Lemma 4.15:

Lemma 4.21 *Die folgenden Eigenschaften sind für $A \in L(X, Y)$ äquivalent:*

(a) $\operatorname{Im} A$ *ist abgeschlossen und* $\operatorname{Ker} A$ *endlich-dimensional.*

(b) *Es gibt einen Banachraum Z, einen kompakten Operator $L \in K(X, Z)$, und ein $C > 0$ so, dass $\|x\| \le C(\|Ax\| + \|Lx\|)$ für alle $x \in X$.*

Beweis Die Implikation (a) $\Rightarrow$ (b) folgt aus Lemma 4.15. Die Implikation (b) $\Rightarrow$ (a) sieht man wie folgt ein.

Zunächst ist $\|x\| \le C\|Lx\|$ für alle $x \in \operatorname{Ker} A$. Hieraus folgt, dass $L : \operatorname{Ker} A \to Z$ ein abgeschlossenes Bild und einen trivialen Kern hat. Nach dem Satz von Banach über den inversen Operator hat $L : \operatorname{Ker} A \to \operatorname{Im} L$ einen beschränkten Inversen B. Folglich ist der identische Operator $I = BL$ auf $\operatorname{Ker} A$ kompakt. Nach dem Satz von Riesz ist dann aber $\operatorname{Ker} A$ ein endlich-dimensionaler Raum.

Wir zeigen nun die Abgeschlossenheit des Bildes von A. Dazu zerlegen wir X in eine direkte Summe $X_1 \dot{+} \operatorname{Ker} A$ (es ist also $X_1 \cap \operatorname{Ker} A = \{0\}$ und $X_1 \dot{+} \operatorname{Ker} A = X$) mit einem abgeschlossenen Teilraum X_1 von X.

Sei $y_n = Ax_n$ eine Folge in $\operatorname{Im} A$, die gegen $y \in Y$ konvergiert. Offenbar kann $x_n \in X_1$ gewählt werden. Wir zeigen zunächst die Beschränktheit der Folge (x_n). Angenommen, diese Folge ist unbeschränkt. Indem wir erforderlichenfalls zu einer Teilfolge übergehen, können wir annehmen, dass $\|x_n\| \to \infty$. Sei $x'_n := x_n/\|x_n\|$. Dann ist $Ax'_n = Ax_n/\|x_n\| \to 0$. Wegen $\|x'_n\| = 1$ gibt es eine Teilfolge (x'_{n_k}) von (x'_n) so, dass die Folge (Lx'_{n_k}) konvergiert. Aus $\|x\| \le C\big(\|Ax\| + \|Lx\|\big)$ folgt dann, dass (x'_{n_k}) eine Fundamentalfolge in X ist. Sei x ihr Grenzwert. Offenbar ist $\|x\| = 1$. Aus der Abgeschlossenheit von X_1 folgt $x \in X_1$, und aus der Stetigkeit von A folgt $Ax = 0$. Es ist also $x \in X_1 \cap \operatorname{Ker} A = \{0\}$ im Widerspruch zu $\|x\| = 1$. Dieser Widerspruch zeigt, dass die Folge (x_n) beschränkt ist.

Nun schließen wir wie oben. Aus $\|x\| \leq C\big(\|Ax\| + \|Lx\|\big)$ und der Kompaktheit von L folgt die Existenz einer Teilfolge (x_{n_k}) von (x_n), die in X eine Fundamentalfolge und folglich konvergent ist. Für ihren Grenzwert $x \in X$ gilt

$$Ax = \lim Ax_{n_k} = \lim y_{n_k} = y,$$

also ist $y \in \operatorname{Im} A$. ∎

Beweis von Folgerung 4.20. Ist A Fredholmsch, so gibt es nach Lemma 4.15 (bzw. Lemma 4.21) ein $L \in K(X)$ und ein $C > 0$ so, dass

$$\|x\| \leq C(\|Ax\| + \|Lx\|) \quad \forall x \in X.$$

Für jeden Operator $K \in K(X, Y)$ gilt dann

$$\|x\| \leq C(\|Ax\|_Y + \|Lx\|_X) \leq C\big(\|(A+K)x\|_Y + \|Lx\|_X + \|Kx\|_Y\big). \tag{4.17}$$

Wir versehen den Raum $Z := X \times Y$ mit der Norm $\|(x, y)\| := \|x\|_X + \|y\|_Y$ und erklären den Operator $\binom{L}{K} : X \to X \times Y$ durch $x \mapsto (Lx, Kx)$. Dieser Operator ist kompakt, und da (4.17) auch als

$$\|x\| \leq C\big(\|(A+K)x\|_Y + \|\tbinom{L}{K}x\|_Z\big) \quad \forall x \in X$$

geschrieben werden kann, folgt aus Lemma 4.21, dass $A + K$ ein abgeschlossenes Bild und einen endlich-dimensionalen Kern besitzt. Aus Satz 4.16 wissen wir weiter, dass der Index eine stetige Funktion ist. Wir betrachten die stetige Operatorfunktion

$$C(t) = A + tK \quad \text{mit } t = [0, 1].$$

All diese Operatoren sind Semi-Fredholmsch, und der Index aller Operatoren $C(t)$ ist gleich. Insbesondere ist

$$\operatorname{ind}(A + K) = \operatorname{ind} C(1) = \operatorname{ind} C(0) = \operatorname{ind} A. \tag{4.18}$$

Da A Fredholmoperator ist, ist $\operatorname{ind} A < \infty$. Also ist auch $\operatorname{ind}(A + K) < \infty$. Dann muss aber $A + K$ ein Fredholmoperator sein. Außerdem zeigt (4.18) die behauptete Indexgleichheit. ∎

Diese Eigenschaften lassen sich noch vielfach ergänzen. Z.B. gilt auch

Satz 4.22 *Ist $A \in L(X, Y)$ ein Fredholmoperator, so ist auch $A' \in L(Y', X')$ ein Fredholmoperator, und es gilt*

$$\operatorname{ind} A' = -\operatorname{ind} A.$$

Eine analoge Aussage gilt für die Hilbertraumadjungierte.

Satz 4.23 (Atkinson) *Ein Operator $A \in L(X,Y)$ ist genau dann Fredholmsch, wenn es Operatoren $B, C \in L(Y,X)$ und kompakte Operatoren $K \in K(X)$, $L \in K(Y)$ gibt, so dass*

$$BA = I_X + K, \quad AC = I_Y + L.$$

Man sagt auch: Fredholmoperatoren sind invertierbar modulo kompakter Operatoren.

4.5 Singuläre Integraloperatoren vom Cauchy-Typ

Wir betrachten Integralgleichungen der Gestalt

$$(Au)(t) := a(t)u(t) + \frac{b(t)}{\pi i} \int_{\mathbb{T}} \frac{u(s)}{s-t}\, ds = f(t), \quad t \in \mathbb{T}. \tag{4.19}$$

Der Operator A ist zusammengesetzt aus Operatoren der Multiplikation mit Funktionen a, b, welche wir als stetig auf dem Einheitskreis $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$ voraussetzen, und aus dem *Operator der singulären Integration*

$$(Su)(t) = \frac{1}{\pi i} \int_{\mathbb{T}} \frac{u(s)}{s-t}\, ds, \quad t \in \mathbb{T}.$$

Dieses Integral existiert selbst für gutartige (z.B. Hölderstetige) Funktionen u nur im Sinne des Cauchyschen Hauptwertes. Für $s = e^{ix}$, $t = e^{iy}$ hat man also

$$(Su)(e^{iy}) = \frac{1}{\pi i} \int_y^{y+2\pi} \frac{u(e^{ix})}{e^{ix} - e^{iy}}\, ie^{ix} dx = \frac{1}{\pi} \lim_{\varepsilon \searrow 0} \int_{y+\varepsilon}^{y+2\pi-\varepsilon} \frac{u(e^{ix})}{1 - e^{i(y-x)}}\, dx. \tag{4.20}$$

Die Wahl der Konstanten vor dem Integralzeichen wird uns später klar werden.

Um eine Vorstellung vom Wirken von S zu bekommen, berechnen wir Se_n für $e_n(t) = t^n$ mit $n \in \mathbb{Z}$.

Lemma 4.24 *Es ist $Se_n = e_n$ für $n \geq 0$ und $Se_n = -e_n$ für $n < 0$.*

Beweis Für $n = 0$ ist

$$\begin{aligned}
(Se_0)(e^{iy}) &= \frac{1}{\pi} \lim_{\varepsilon \to 0} \int_{y+\varepsilon}^{y+2\pi-\varepsilon} \frac{1}{1 - e^{-(x-y)i}}\, dx \quad \text{(Substitution } x := x - y\text{)} \\
&= \frac{1}{\pi} \lim_{\varepsilon \to 0} \int_{\varepsilon}^{2\pi-\varepsilon} \frac{1}{1 - e^{-ix}}\, dx = \frac{1}{\pi} \lim_{\varepsilon \to 0} \int_{\varepsilon}^{2\pi-\varepsilon} \frac{e^{ix/2}}{e^{ix/2} - e^{-ix/2}}\, dx \\
&= \frac{1}{2\pi i} \lim_{\varepsilon \to 0} \int_{\varepsilon}^{2\pi-\varepsilon} (\cot \frac{x}{2} + i) dx = 1,
\end{aligned}$$

da die Funktion $x \mapsto \cot \frac{x}{2}$ auf $(0, 2\pi)$ bezüglich des Mittelpunktes π dieses Intervalles ungerade ist. Hieraus folgt leicht für $n > 0$

$$\begin{aligned}
(Se_n)(t) &= \frac{1}{\pi i} \int_{\mathbb{T}} \frac{s^n}{s-t}\, ds = \frac{1}{\pi i} \int_{\mathbb{T}} \frac{s^n - t^n}{s-t}\, ds + \frac{1}{\pi i} \int_{\mathbb{T}} \frac{t^n}{s-t}\, ds \\
&= \frac{t^n}{\pi i} \int_{\mathbb{T}} \frac{1}{s-t}\, ds = t^n (Se_0)(t) = t^n.
\end{aligned}$$

Schließlich ist für $n > 0$

$$\begin{aligned}
(Se_{-n})(t) &= \frac{1}{\pi i} \int_{\mathbb{T}} \frac{s^{-n}}{s-t}\, ds = \frac{1}{\pi i} \int_{\mathbb{T}} \frac{s^{-n} - t^{-n}}{s-t}\, ds + \frac{1}{\pi i} \int_{\mathbb{T}} \frac{t^{-n}}{s-t}\, ds \\
&= t^{-n} + \frac{1}{\pi i} \int_{\mathbb{T}} - \sum_{k=0}^{n-1} t^{k-n} s^{-k-1}\, ds = t^{-n} - 2t^{-n} = -t^{-n},
\end{aligned}$$

da

$$\frac{s^{-n} - t^{-n}}{s-t} = \frac{-1}{st} \frac{s^{-n} - t^{-n}}{s^{-1} - t^{-1}} = \frac{-1}{st} \sum_{k=0}^{n-1} s^{-k} t^{-n+1+k} = - \sum_{k=0}^{n-1} s^{-k-1} t^{-n+k}$$

und da bekanntlich

$$\int_{\mathbb{T}} z^n dz = \begin{cases} 0 & n \neq -1 \\ 2\pi i & n = -1. \end{cases}$$

■

Diese einfache Rechnung hat eine Reihe bemerkenswerter Konsequenzen, wenn wir bedenken, dass die Funktionen $(t^n)_{n \in \mathbb{Z}}$ eine Orthonormalbasis von $L^2(\mathbb{T})$ bilden (dies ist im Wesentlichen äquivalent zu der Tatsache, dass die Funktionen 1 und $\sin 2\pi nt$, $\cos 2\pi nt$ mit $n > 0$ eine ONB von $L^2[0, 1]$ bilden).

Folgerung 4.25 *Der Operator S der singulären Integration ist eine Isometrie von $L^2(\mathbb{T})$ auf sich.*

Beweis Der Operator S überführt die Fourierreihe $\sum_{n \in \mathbb{Z}} a_n e_n$ in die Reihe

$$\sum_{n \geq 0} a_n e_n - \sum_{n < 0} a_n e_n. \tag{4.21}$$

Diese Reihe konvergiert und definiert eine Element von $L^2(\mathbb{T})$, welches die gleiche Norm hat wie das Ausgangselement $\sum_{n \in \mathbb{Z}} a_n e_n$. ■

Folgerung 4.26 *Es ist $S^2 = I$ und $S^* = S$.*

Beweis Anwendung von S auf $\sum_{n \in \mathbb{Z}} a_n e_n$ führt auf (4.21). Nochmalige Anwendung von S auf (4.21) liefert wieder die Ausgangsfunktion $\sum_{n \in \mathbb{Z}} a_n e_n$. Genauso einfach zeigt man, dass $S^* = S$ ist. ■

Folgerung 4.27 *$P := \frac{1}{2}(I+S)$ und $Q := \frac{1}{2}(I-S)$ sind Orthoprojektoren (d.h. es ist $P^2 = P$, $P^* = P$).*

Dies folgt sofort aus Folgerung 4.26 durch einfaches Nachrechnen. ■

Offenbar ist

$$aI + bS = a(P+Q) + b(P-Q) = (a+b)P + (a-b)Q.$$

Setzen wir $c := a+b$ und $d := a-b$, so können wir den singulären Integraloperator $A = aI + bS$ auch in der Form $cP + dQ$ schreiben, die für viele Überlegungen zweckmäßiger ist.

Folgerung 4.28 *Seien $c, d \in \mathbb{C}$. Dann ist der Operator $cP + dQ$ genau dann invertierbar, wenn $c \neq 0$ und $d \neq 0$. Ist diese Bedingung erfüllt, so ist der Operator $\frac{1}{c}P + \frac{1}{d}Q$ die Inverse von $cP + dQ$.*

Dies haben wir bereits im Beweis von Satz 4.4 (e) bemerkt. ■

Wir können nun bereits singuläre Integralgleichungen mit konstanten Koeffizienten a, b lösen! Dazu schreiben wir

$$\begin{aligned} A^{-1} &= \frac{1}{c}P + \frac{1}{d}Q = \frac{1}{a+b}\,\frac{I+S}{2} + \frac{1}{a-b}\,\frac{I-S}{2} \\ &= \frac{1}{2}\Big(\frac{1}{a+b} + \frac{1}{a-b}\Big)I + \frac{1}{2}\Big(\frac{1}{a+b} - \frac{1}{a-b}\Big)S = \frac{a}{a^2-b^2}I - \frac{b}{a^2-b^2}\,S. \end{aligned}$$

Also gilt: Für konstante Koeffizienten a, b ist die Gleichung (4.19) genau dann lösbar, wenn $a^2 - b^2 \neq 0$. In diesem Fall ist die Lösung u gegeben durch

$$u(t) = (A^{-1}f)(t) = \frac{a}{a^2-b^2}\,f(t) - \frac{b}{a^2-b^2} \cdot \frac{1}{\pi i}\int_{\mathbb{T}} \frac{f(s)}{s-t}\,ds.$$

Wir wenden uns nun dem Fall zu, wenn a und b stetige Funktionen sind. Dazu ist es zweckmäßig, den Raum $L^2(\mathbb{T})$ mit dem Raum $\ell^2(\mathbb{Z})$ zu identifizieren, der aus allen Folgen $(x_n)_{n\in\mathbb{Z}}$ mit $\sum_{n\in\mathbb{Z}} |x_n|^2 < \infty$ besteht. Die Identifikation geschieht so, dass der Funktion $x = \sum_{n\in\mathbb{Z}} x_n e_n \in L^2(\mathbb{T})$ die Folge $(x_n)_{n\in\mathbb{Z}}$ ihrer Fourierkoeffizienten zugeordnet wird (Satz von Fischer/Riesz). Jeder Operator $A \in L\big(L^2(\mathbb{T})\big)$ wird dabei übersetzt in einen Operator auf $\ell^2(\mathbb{Z})$, den man sich als unendliche Matrix $(a_{jk})_{j,k\in\mathbb{Z}}$ mit den Einträgen $a_{jk} = \langle Ae_k, e_j\rangle$ vorstellen kann. Das Skalarprodukt auf $L^2(\mathbb{T})$ ist dabei definiert als

$$\langle f, g\rangle = \frac{1}{2\pi}\int_0^{2\pi} f(e^{it})\,\overline{g(e^{it})}\,dt.$$

Die Matrixdarstellung des Operators S haben wir praktisch bereits in Lemma 4.24 bestimmt: es ist die Diagonalmatrix

$$S \longleftrightarrow \left(\begin{array}{ccc|ccc}
\ddots & & & & & \\
& -1 & & & & \\
& & -1 & & & \\
\hline
& & & 1 & & \\
& & & & 1 & \\
& & & & & \ddots
\end{array}\right).$$

Da I in die Einheitsmatrix überführt wird, lauten die Matrixdarstellungen der Projektoren P und Q entsprechend

$$P \longleftrightarrow \left(\begin{array}{ccc|ccc}
\ddots & & & & & \\
& 0 & & & & \\
& & 0 & & & \\
\hline
& & & 1 & & \\
& & & & 1 & \\
& & & & & \ddots
\end{array}\right), \quad Q \longleftrightarrow \left(\begin{array}{ccc|ccc}
\ddots & & & & & \\
& 1 & & & & \\
& & 1 & & & \\
\hline
& & & 0 & & \\
& & & & 0 & \\
& & & & & \ddots
\end{array}\right).$$

Wir benötigen noch die Matrixdarstellung von Operatoren der Multiplikation. Für $c \in C(\mathbb{T})$ ist

$$\begin{aligned}
\langle ce_k, e_j \rangle &= \frac{1}{2\pi} \int_0^{2\pi} c(e^{it}) e^{ikt} e^{-ijt}\, dt \\
&= \frac{1}{2\pi} \int_0^{2\pi} c(e^{it}) e^{it(k-j)} dt =: c_{j-k}.
\end{aligned}$$

Der Eintrag $\langle ce_k, e_j \rangle$ hängt also nur von der Differenz der Argumente k und j ab, so dass die Matrixdarstellung von cI lautet

$$cI \longleftrightarrow \left(\begin{array}{ccc|cccc}
 & \ddots & \ddots & \ddots & \ddots & & \\
\ddots & c_0 & c_{-1} & c_{-2} & c_{-3} & \ddots & \\
\ddots & c_1 & c_0 & c_{-1} & c_{-2} & c_{-3} & \ddots \\
\hline
\ddots & c_2 & c_1 & c_0 & c_{-1} & c_{-2} & \ddots \\
\ddots & c_3 & c_2 & c_1 & c_0 & c_{-1} & \ddots \\
 & & c_3 & c_2 & c_1 & c_0 & \\
 & & \ddots & \ddots & \ddots & \ddots &
\end{array}\right).$$

Insbesondere ist die Matrixdarstellung von $c(t) = t$ wegen

$$\frac{1}{2\pi} \int_0^{2\pi} e^{it} e^{it(k-j)} dt = \frac{1}{2\pi} \int_0^{2\pi} e^{it(1+k-j)} dt = \begin{cases} 0 & \text{für } 1+k-j \neq 0 \\ 1 & \text{für } 1+k-j = 0 \end{cases}$$

gleich

$$
tI \longleftrightarrow \left(\begin{array}{ccccc|ccc}
\ddots & \ddots & \ddots & \ddots & & & & \\
 & 0 & 1 & 0 & 0 & & & \\
 & & 0 & 1 & 0 & 0 & & \\
\hline
 & & & 0 & 1 & 0 & 0 & \\
 & & & & 0 & 1 & 0 & \ddots \\
 & & & & & 0 & 1 & \ddots \\
 & & & & & & 0 & \ddots \\
 & & & & & & & \ddots
\end{array}\right).
$$

Dem singulären Integraloperator $cP + dQ$ entspricht daher die Matrix

$$
\left(\begin{array}{cccc|cccc}
\ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\
\ddots & d_0 & d_{-1} & d_{-2} & c_{-3} & c_{-4} & c_{-5} & \ddots \\
\ddots & d_1 & d_0 & d_{-1} & c_{-2} & c_{-3} & c_{-4} & \ddots \\
\ddots & d_2 & d_1 & d_0 & c_{-1} & c_{-2} & c_{-3} & \ddots \\
\hline
\ddots & d_3 & d_2 & d_1 & c_0 & c_{-1} & c_{-2} & \ddots \\
\ddots & d_4 & d_3 & d_2 & c_1 & c_0 & c_{-1} & \ddots \\
\ddots & d_5 & d_4 & d_3 & c_2 & c_1 & c_0 & \ddots \\
 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots
\end{array}\right). \tag{4.22}
$$

Wir interessieren uns nun zunächst für die Fredholmtheorie von Operatoren der Gestalt (4.22). Jede stetige Funktion c auf $\mathbb{T}$ kann man beliebig genau durch sogenannte trigonometrische Polynome $\sum_{k=-n}^{n} a_k t^k$ approximieren (Version des Satzes von Weierstraß). Für Polynome c und d stehen in der rechten oberen bzw. linken unteren Ecke der Matrix (4.22) nur endlich viele Nichtnull-Einträge. Also haben die entsprechenden Operatoren ein endlich-dimensionales Bild und sind insbesondere kompakt. Dann sind diese Operatoren aber für beliebige stetige Funktionen kompakt, da sie sich durch Operatoren von endlichem Rang approximieren lassen. Die Matrix (4.22) ist also genau dann ein Fredholmoperator, wenn der Operator

$$\left(\begin{array}{cccc|cccc}
\ddots & \ddots & \ddots & & & & & \\
\ddots & d_0 & d_{-1} & d_{-2} & & & & \\
\ddots & d_1 & d_0 & d_{-1} & & \multicolumn{2}{c}{\text{O}} & \\
 & d_2 & d_1 & d_0 & & & & \\
\hline
 & & & & c_0 & c_{-1} & c_{-2} & \\
 & \multicolumn{2}{c}{\text{O}} & & c_1 & c_0 & c_{-1} & \ddots \\
 & & & & c_2 & c_1 & c_0 & \ddots \\
 & & & & & \ddots & \ddots & \ddots
\end{array}\right) \tag{4.23}$$

ein Fredholmoperator ist. Außerdem sind die Indizes der Operatoren (4.22) und (4.23) gleich. Die Untersuchung des Operators (4.23) kann man offenbar auf die Untersuchung der Operatoren in seiner rechten unteren bzw. linken oberen Ecke zurückführen. Ein Operator der Gestalt

$$T(c) = \begin{pmatrix} c_0 & c_{-1} & c_{-2} & \ddots \\ c_1 & c_0 & c_{-1} & \ddots \\ c_2 & c_1 & c_0 & \\ & \ddots & \ddots & \ddots \end{pmatrix} \in L(\ell^2)$$

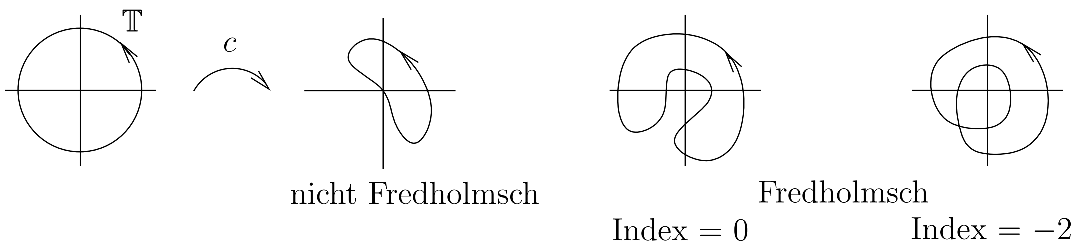
heißt der durch c erzeugte *Toeplitzoperator*. Einen solchen Operator kennen wir bereits: Für $c(t) = t$ ist $T(c)$ gerade der Verschiebungsoperator

$$T(c) : (x_0, x_1, \ldots) \mapsto (0, x_0, x_1, \ldots),$$

und ganz analog ist für $c(t) = t^{-1}$

$$T(t^{-1}) : (x_0, x_1, \ldots) \mapsto (x_1, x_2, \ldots).$$

Satz 4.29 *Der Operator $T(c)$ mit $c \in C(\mathbb{T})$ ist genau dann Fredholmsch, wenn $c(t) \neq 0$ für alle $t \in \mathbb{T}$. Ist dies erfüllt, so ist der Index von $T(c)$ gleich der negativen Windungszahl der Funktion c bezüglich des Nullpunktes.*



Beweis $\rightarrow$ Schröder, Abschnitt 4.4

Wir haben damit ein sehr schönes geometrisches (topologisches) Kriterium dafür gefunden, wann ein Toeplitzoperator Fredholmsch ist, und können außerdem leicht seinen Index bestimmen. In diesem speziellen Fall haben wir damit bereits ein Invertierbarkeitskriterium! Es gilt nämlich

Satz 4.30 (Coburn, Gohberg/Krein) *Ein Toeplitzoperator ist genau dann invertierbar, wenn er Fredholmsch ist und den Index* 0 *besitzt.*

Mit Hilfe der Resultate des vorigen Abschnittes ist es nicht schwer, aus Satz 4.29 und 4.30 das folgende Fredholm- und Invertierbarkeitskriterium für singuläre Integraloperatoren abzuleiten.

Satz 4.31 (Noether) *Der singuläre Integraloperator $A = cP + dQ$ mit $c, d \in C(\mathbb{T})$ ist genau dann ein Fredholmoperator, wenn $c(t) \neq 0$ und $d(t) \neq 0$ für alle $t \in \mathbb{T}$. Ist dies erfüllt, so ist $\operatorname{ind} A$ gleich der negativen Windungszahl der Funktion c/d bezüglich des Nullpunktes. Falls $\operatorname{ind} A = 0$, so ist A invertierbar.*

Beweis $\nearrow$ Schröder, Spezialliteratur.

Anmerkungen

1. Die explizite Bestimmung von A^{-1} ist i.Allg. schwierig. Ein Algorithmus zur Bestimmung von A^{-1} existiert für rationale Funktionen c, d.

2. David Hilbert hat vermutet, dass (genau wie bei Fredholmschen Integralgleichungen 2. Art) auch der Index singulärer Integraloperatoren stets 0 ist. Noethers Satz widerlegt diese Vermutung Hilberts.

3. Die Noethersche Indexformel ist der Vorläufer aller späteren Indexsätze wie dem von Atiyah-Singer. In diesen Sätzen geht es darum, den analytischen Index $\operatorname{ind} A$ eines Operators mit topologischen Größen (wie die Windungszahl) in Beziehung zu setzen.

4. Im Wintersemester gibt es eine Vorlesung, in der ich über Numerische Analysis für Integralgleichungen (u.a.) sprechen möchte. Der Schwerpunkt dieser Vorlesung wird auf dem Einsatz funktionalanalytischer Methoden (lokale Prinzipien in Banachalgebren) zum Studium der Konvergenz von Näherungsverfahren liegen.