Elemente der Funktionalanalysis

Vorlesungsskript
Sommersemester 2014

von Markus Kunze

Vorwort

Das vorliegende Skriptum fasst die Vorlesung "Elemente der Funktionalanalysis" zusammen, die ich im Sommersemester 2014 an der Universität Ulm gehalten habe. Die Vorlesung richtet sich an Bachelorstudenten im 4. oder 6. Semester und baut auf den Vorlesungen Analysis 1 und 2, lineare Algebra 1 und 2 und der Vorlesung Maßtheorie auf.

Teilweise wird auch auf Resultate der Vorlesung "Topologie" zurückgegriffen. Im Falle des Satzes von Arzelà–Ascoli geben wir einen Beweis in einem Spezialfall (Satz 4.4.2). Auch gibt es eine Überschneidung mit der Vorlesung "Analysis 3" in der die klassische Theorie der Fourierreihen behandelt wird. Diese wird hier weder vorausgesetzt, noch näher besprochen.

Schließlich ist anzumerken, dass die vorliegende Version des Skriptes die erste ist. Es ist also damit zu rechnen, dass sich im Skript noch zahlreiche Tippfehler, wenn nicht gar mathematische Ungenauigkeiten, befinden. Wenn man mich auf solche aufmerksam macht (`markus.kunze@uni-ulm.de`) so wäre ich sehr dankbar.

M. K.

Inhaltsverzeichnis

Kapitel 1

Grundlagen

In der Funktionalanalysis studiert man unendlichdimensionale Vektorräume. Anders als in der linearen Algebra sind diese Räume zusätzlich mit einer *Norm* ausgestattet, sodass man auch einen Konvergenzbegriff hat. In diesem Kapitel führen wir die zentralen Begriffe *normierter Raum* und *Banachraum* ein und stellen die wichtigsten Beispiele vor. Anschließend studieren wir stetige, lineare Abbildungen zwischen solchen Räumen. Diese Abbildungen respektieren aufgrund ihrer Stetigkeit nicht nur die lineare, sondern auch die topologische Struktur.

1.1 Normierte Räume und Banachräume

Definition 1.1.1. Es sei X ein Vektorraum über $\mathbb{K} = \mathbb{C}$ oder über $\mathbb{K} = \mathbb{R}$. Eine *Norm* auf X ist eine Abbildung $\|\cdot\| : X \to [0, \infty)$ derart, dass für alle $x, y \in X$ und $\lambda \in \mathbb{K}$ folgende Aussagen gelten.

(N1) *Definitheit*: Es ist $\|x\| = 0$ genau dann, wenn $x = 0$.

(N2) *Homogenität*: Es ist $\|\lambda x\| = |\lambda| \cdot \|x\|$.

(N3) *Dreiecksungleichung*: Es ist $\|x + y\| \leq \|x\| + \|y\|$.

Das Paar $(X, \|\cdot\|)$ heißt *normierter Raum*. Gelten lediglich (N2) und (N3), so nennt man $\|\cdot\|$ eine *Halbnorm* auf X und $(X, \|\cdot\|)$ einen *halbnormierten Raum*.

Ist $(X, \|\cdot\|)$ ein normierter Raum, so definiert $d(x, y) := \|x - y\|$ eine *Metrik* auf X, sodass uns alle Begriffe die auf metrischen Räumen definiert sind, zur Verfügung stehen. Insbesondere können wir von offenen und abgeschlossenen Mengen, von stetigen Funktionen und von konvergenten Folgen sprechen. Wir wiederholen nochmals die Definition von konvergenten Folgen und von Cauchyfolgen in unserer Situation.

Definition 1.1.2. Es sei $(X, \|\cdot\|)$ ein normierter Raum und $(x_n)_{n\in\mathbb{N}}$ eine Folge in X.

(1) $(x_n)_{n\in\mathbb{N}}$ heißt *Cauchy Folge*, falls es zu jedem $\varepsilon > 0$ einen Index $n_0 \in \mathbb{N}$ gibt, sodass $\|x_n - x_m\| \leq \varepsilon$ für alle $n, m \geq n_0$.

(2) $(x_n)_{n\in\mathbb{N}}$ heißt *konvergent*, falls es ein $x \in X$ gibt, sodass es zu jedem $\varepsilon > 0$ einen Index $n_0 \in \mathbb{N}$ gibt, sodass $\|x_n - x\| \leq \varepsilon$ für alle $n \geq n_0$. In diesem Fall heißt x *Grenzwert* der Folge $(x_n)_{n\in\mathbb{N}}$. Wir schreiben $x = \lim_{n\to\infty} x_n$ oder $x_n \to x$.

Ein normierter Raum, in dem jede Cauchyfolge konvergiert, heißt *vollständiger normierter Raum* oder *Banachraum*.

Lemma 1.1.3. *Es sei $(x_n)_{n\in\mathbb{N}}$ eine Folge im normierten Raum $(X, \|\cdot\|)$.*

(1) Gelten $x_n \to x$ und $x_n \to y$, so ist $x = y$.

(2) Wenn $x_n \to x$, dann $\|x_n\| \to \|x\|$.

(3) Wenn $(x_n)_{n\in\mathbb{N}}$ konvergiert, so ist $(x_n)_{n\in\mathbb{N}}$ eine Cauchyfolge.

(4) Ist $(x_n)_{n\in\mathbb{N}}$ eine Cauchyfolge, so ist $(x_n)_{n\in\mathbb{N}}$ beschränkt*, d.h.* $\sup_{n\in\mathbb{N}} \|x_n\| < \infty$.

Beweis. (1) Sei $x_n \to x$ und $x_n \to y$. Zu $\varepsilon > 0$ finden wir Indizes n_1 und n_2 derart, dass $\|x_n - x\| \leq \varepsilon$ für $n \geq n_1$ und $\|x_n - y\| \leq \varepsilon$ für $n \geq n_2$. Unter Verwendung der Dreiecksungleichung und der Homogenität folgt also, dass

$$\begin{aligned}\|x - y\| &= \|x - x_{n_1+n_2} + x_{n_1+n_2} - y\| \leq \|x - x_{n_1+n_2}\| + \|x_{n_1+n_2} - y\| \\ &= \|x - x_{n_1+n_2}\| + |-1|\|y - x_{n_1+n_2}\| \leq 2\varepsilon\end{aligned}$$

Da $\varepsilon > 0$ beliebig war, folgt $\|x - y\| = 0$ also $x = y$ wegen der Definitheit von $\|\cdot\|$.

(2) Es folgt aus der Dreiecksungleichung, dass $\|x_n\| \leq \|x_n - x\| + \|x\|$, also $\|x_n\| - \|x\| \leq \|x_n - x\|$. In Verbindung mit der Homogenität sieht man genauso, dass $\|x\| - \|x_n\| \leq \|x_n - x\|$, insgesamt also $|\|x_n\| - \|x\|| \leq \|x_n - x\|$. Daraus folgt die Behauptung.

(3) Wenn (x_n) konvergiert, etwa gegen x, so findet man zu gegebenem $\varepsilon > 0$ einen Index n_0 derart, dass $\|x_n - x\| = \|x - x_n\| \leq \varepsilon/2$ für alle $n \geq n_0$. Insbesondere ist für $n, m \geq n_0$ stets

$$\|x_n - x_m\| \leq \|x_n - x\| + \|x - x_m\| \leq \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Dies zeigt, dass (x_n) eine Cauchyfolge ist.

(4) Ist (x_n) eine Cauchyfolge, so können wir N so wählen, dass $\|x_n - x_N\| \leq 1$ für alle $n \geq N$. Ist $C := \max\{\|x_j\| : j = 1, \ldots, N\} + 1$, so ist $\|x_n\| \leq C$ für alle $n \in \mathbb{N}$. □

Bemerkung 1.1.4. In halbnormierten Räumen kann man immer noch von Konvergenz und Vollständigkeit reden. Allerdings sind die Grenzwerte von konvergenten Folgen nicht mehr notwendigerweise eindeutig.

Wir geben nun einige Beispiele von Banachräumen. Der Beweis, dass ein normierter Raum X vollständig ist (also ein Banachraum ist) erfolgt meist in drei Schritten:

(i) Gegeben eine Cauchyfolge findet man einen *Kandidaten* für den Grenzwert. Typischerweise verwendet man hier die Vollständigkeit des Skalarkörpers und nimmt Grenzwerte "in einem schwächeren Sinn", etwa punktweise.

(ii) Man zeigt, dass der gefundene Kandidat in X liegt.

(iii) Man zeigt, dass die Folge tatsächlich bezüglich der Norm gegen den Kandidaten konvergiert.

Satz 1.1.5. *Es sei Ω eine Menge und*

$$\mathcal{F}_b(\Omega) := \{f : \Omega \to \mathbb{K} : f \text{ ist beschränkt}\}.$$

Dann ist $\mathcal{F}_b(\Omega)$ ein Banachraum bezüglich der Supremumsnorm *(!) $\|\cdot\|_\infty$, gegeben durch*

$$\|f\|_\infty := \sup_{x\in\Omega} |f(x)|.$$

Beweis. Es sei (f_n) eine Cauchyfolge bezüglich $\|\cdot\|_\infty$. Für $x \in \Omega$ ist dann wegen

$$|f_n(x) - f_m(x)| \leq \sup_{x\in\Omega} |f_n(x) - f_m(x)| = \|f_n - f_m\|_\infty$$

auch $(f_n(x))$ eine Cauchyfolge in $\mathbb{K}$, also (wegen der Vollständigkeit von $\mathbb{K}$) konvergent. Es sei $f(x) := \lim_{n\to\infty} f_n(x)$. (Dies ist unser "Kandidat" für den Grenzwert.)

Dann ist f beschränkt, also ein Element von $\mathcal{F}_b(\Omega)$. In der Tat ist wegen Lemma 1.1.3(4) die Folge f_n beschränkt. Also gibt es ein $C \geq 0$ mit $|f_n(x)| \leq \|f_n\|_\infty \leq C$ für alle $x \in \Omega$ und $n \in \mathbb{N}$. Weil das Intervall $(-\infty, C]$ abgeschlossen ist, folgt, dass auch $|f(x)| = \lim |f_n(x)| \leq C$ ist. Somit ist f beschränkt.

Es bleibt zu zeigen, dass $\|f_n - f\|_\infty \to 0$. Sei hierzu $\varepsilon > 0$ gegeben. Zunächst wählen wir n_0 derart, dass $\|f_n - f_m\|_\infty \leq \varepsilon$ ür alle $n, m \geq n_0$ gilt. Sei nun $x \in \Omega$. Nach Wahl von f gibt es einen Index $n_1(x)$ so, dass $|f_n(x) - f(x)| \leq \varepsilon$ für alle $n \geq n_1(x)$. Ohne Einschränkung dürfen wir annehmen, dass $n_1(x) \geq n_0$ ist. Für $n \geq n_0$ gilt dann

$$\begin{aligned}|f_n(x) - f(x)| &\leq |f_n(x) - f_{n_1(x)}(x)| + |f_{n_1(x)}(x) - f(x)| \\ &\leq \|f_n - f_{n_1(x)}\|_\infty + \varepsilon \leq 2\varepsilon.\end{aligned}$$

Da $x \in \Omega$ beliebig war, folgt $\|f_n - f\|_\infty \leq 2\varepsilon$ für alle $n \geq n_0$. Dies zeigt, dass $f_n \to f$. □

Wählt man für $\Omega = \{1, \ldots, d\}$, so kann man den Raum $\mathcal{F}_b(\Omega)$ mit $\mathbb{K}^d$, dem üblichen d-dimensionalen Koordinatenraum identifizieren. Die Norm $\|x\|_\infty$ ist in diesem Fall gerade gegeben durch

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

Wählt man $\Omega = \mathbb{N}$, so erhält man einen anderen wichtigen Spezialfall, den Raum ℓ^∞ der beschränkten Folgen. Dies ist also auch ein Banachraum bezüglich der Supremumsnorm.

Korollar 1.1.6. *Es sei*

$$\ell^\infty := \{\mathbf{x} := (x_n)_{n\in\mathbb{N}} : (x_n) \textit{ ist eine beschränkte Folge in } \mathbb{K}\}.$$

Dann ist ℓ^∞ ein Banachraum bezüglich der Norm $\|\mathbf{x}\|_\infty := \sup_n |x_n|$.

Bemerkung 1.1.7. Es ist leicht zu sehen, dass eine Folge in $\mathbb{K}^d$ genau dann konvergiert, wenn jede Komponente konvergiert. Dies ist in ℓ^∞ nicht der Fall, wie man an der Folge $\mathbf{x}_n := (0, 0, \ldots, 0, 1, 1, \ldots)$ (zuerst n Nullen, dann nur noch Einsen) sieht. Diese Folge konvergiert komponentenweise gegen den Nullvektor $0 = (0, 0, 0, \ldots)$, aber $\|\mathbf{x}_n - 0\|_\infty = 1 \not\to 0$.

Ist X ein Teilraum eines Banachraumes, so ist es wesentlich einfacher zu zeigen, dass X selbst ein Banachraum ist.

Lemma 1.1.8. *Es sei $(Y, \|\cdot\|)$ ein Banachraum und $X \subset Y$ ein Teilraum. Dann ist $(X, \|\cdot\|)$ genau dann ein Banachraum, wenn X abgeschlossen in Y ist.*

Beweis. Angenommen, $(X, \|\cdot\|)$ ist ein Banachraum. Ist x_n eine Folge in X, die gegen $y \in Y$ konvergiert, so ist wegen Lemma 1.1.3(3) (x_n) eine Cauchyfolge in Y, also auch in X. Wegen der Vollständigkeit von X gilt $x_n \to x$ für ein $x \in X$. Aus Lemma 1.1.3(1) folgt $x = y$, also $y \in X$.

Ist umgekehrt X abgeschlossen in Y und (x_n) eine Cauchyfolge in X, so ist (x_n) auch eine Cauchyfolge in Y, also, wegen Vollständigkeit, konvergent. Es sei etwa $x_n \to y$. Da X abgeschlossen ist, ist $y \in X$ und daher (x_n) als Folge in X konvergent. □

Lemma 1.1.8 besagt, dass man sich in obigem “Kochrezept” auf Punkt (ii) beschränken kann, wenn der gegebene Raum ein Teilraum eines Banachraumes ist. In der Tat ist, wegen der Eindeutigkeit des Grenzwertes, der einzig mögliche Kandidat für den Grenzwert derjenige, den man im größeren Raum findet. Gegen diesen konvergiert die Folge auch in der richtigen Norm. Es bleibt lediglich zu klären, ob der Grenzwert im richtigen Raum liegt.

Korollar 1.1.9. *Es sei Ω ein metrischer Raum. Insbesondere ist der Raum $C_b(\Omega)$ der beschränkten, stetigen Funktionen auf Ω ein Banachraum bezüglich der Supremumsnorm.*

Beweis. Offensichtlich ist $C(\Omega) \subset \mathcal{F}_b(\Omega)$. Wegen Lemma 1.1.8 genügt es also zu zeigen, dass $C(\Omega)$ abgeschlossen in $\mathcal{F}_b(\Omega)$ ist. Sei hierzu f_n eine Folge in $C(\Omega)$ die gegen $f \in \mathcal{F}_b(\Omega)$ konvergiert. Zu gegebenem ε gibt es also ein n_0 sodass $|f_n(x) - f(x)| \leq \varepsilon$ für alle $n \geq n_0$ und alle $x \in \Omega$. Das bedeutet, dass f_n *gleichmäßig* gegen f konvergiert. Da der gleichmäßige Grenzwert stetiger Funktionen stetig ist, ist $f \in C_b(\Omega)$, also $C(\Omega)$ abgeschlossen in $\mathcal{F}_b(\Omega)$. $\square$

Bemerkung 1.1.10. Ist $\Omega = K$ kompakt, so ist jede stetige Funktion auf K beschränkt. In diesem Fall lässt man den index b häufig weg und schreibt lediglich $C(K)$ anstelle von $C_b(K)$.

Korollar 1.1.11. *Der Raum c_0 der Nullfolgen ist ein Banachraum bezüglich der Supremumsnorm.*

Beweis. Wegen Lemma 1.1.8 genügt es zu zeigen, dass c_0 in ℓ^∞ abgeschlossen ist. Sei hierzu $(\mathbf{x}_n)$ eine Folge in c_0, die gegen $\mathbf{x} \in \ell^\infty$ konvergiert. Wir schreiben $\mathbf{x}_n = (x_n^{(j)})_{j\in\mathbb{N}}$ und $\mathbf{x} = (x^{(j)})_{j\in\mathbb{N}}$. Es ist zu zeigen, dass $x^{(j)} \to 0$ für $j \to \infty$. Zu gegebenem $\varepsilon > 0$ wählen wir n_0 derart, dass $\|\mathbf{x}_n - \mathbf{x}\|_\infty \leq \varepsilon$ für $n \geq n_0$. Das ist möglich, da $\mathbf{x}_n \to \mathbf{x}$. Weil $\mathbf{x}_{n_0} \in c_0$ gibt es j_0 mit $|x_{n_0}^{(j)}| \leq \varepsilon$ für $j \geq j_0$. Für solche j folgt dann

$$|x^{(j)}| \leq |x^{(j)} - x_{n_0}^{(j)}| + |x_{n_0}^{(j)}| \leq \|\mathbf{x} - \mathbf{x}_n\|_\infty + \varepsilon \leq 2\varepsilon.$$

Somit ist $\mathbf{x} \in c_0$. $\square$

Für $a < b$ können wir auf dem Raum $C([a,b])$ auch folgende Norm betrachten:

$$\|f\|_1 := \int_a^b |f(t)|\,dt. \tag{1.1}$$

Aus der Linearität des Integrals folgen Homogenität und Dreiecksungleichung leicht. Für die Definitheit genügt es zu bemerken, dass eine *stetige* Funktion $f \geq 0$ die nicht Null ist, positives Integral hat. In der Tat, ist $f(t_0) \geq 2\varepsilon > 0$, so gibt es wegen der Stetigkeit von f ein $\delta > 0$ mit $f(t) \geq \varepsilon$ für $t \in (t_0 - \delta, t_0 + \delta)$. Wir können annehmen, dass entweder $[t_0, t_0 + \delta) \subset [a,b]$ oder $(t_0 - \delta, t_0] \subset [a,b]$, sonst verkleinern wir δ. Dann gilt

$$\int_0^1 f(t)\,dt \geq \int_0^1 \varepsilon \mathbb{1}_{(t_0-\delta,t_0+\delta)}(t)\,dt = \varepsilon\delta > 0,$$

also ist $\|\cdot\|_1$ definit. Allerdings ist $(C([a,b]), \|\cdot\|_1)$ *kein* Banachraum. Wir zeigen dies für $a = -1, b = 1$. Den allgemeinen Fall beweist man analog. Es sei

$$f_n(t) = nt\mathbb{1}_{(0,n^{-1})} + \mathbb{1}_{[n^{-1},1]} \in C([-1,1]).$$

Dann ist $f_n \leq f_m$ für $m > n$, also

$$\|f_n - f_m\|_1 = \int_{-1}^1 f_m(t) - f_n(t)\,dt = \int_0^{\frac{1}{m}} mt - nt\,dt + \int_{\frac{1}{m}}^{\frac{1}{n}} 1 - nt\,dt$$

$$= \frac{m-n}{2m^2} + \frac{1-n}{2n^2} - \frac{1-n}{2m^2} = \frac{1}{2}\Big(\frac{1}{n^2} - \frac{1}{m^2}\Big) + \frac{1}{2}\Big(\frac{1}{m} - \frac{1}{n}\Big).$$

Weil sowohl $\frac{1}{2}\frac{1}{n^2}$, als auch $\frac{1}{2}\frac{1}{n}$ in $\mathbb{R}$ konvergent, also Cauchyfolgen sind, folgt, dass auch f_n eine Cauchyfolge in $(C([-1,1]), \|\cdot\|_1)$ ist. Allerdings konvergiert f_n in $(C([-1,1]), \|\cdot\|_1)$ nicht.

Wäre nämlich f ein Grenzwert der Folge f_n, so müsste $f(x) = 1$ für $x \in (0,1]$ und $f(x) = 0$ für $x \in [-1,0)$ sein, was für eine stetige Funktion unmöglich ist. Wenn nämlich $f(x_0) \neq 1$ für $x_0 > 1$ wäre, so müsste wegen der Stetigkeit von f_n und f in x_0 bereits $|f_n(x) - f(x)| \geq \varepsilon$ für alle x in einer Umgebung von x_0 für genügend große n. Dann wäre aber, wie beim Beweis der Definitheit, $\|f_n - f\|_1 \geq c > 0$ für große n, was nicht sein kann. Analog zeigt man, dass $f_n(x) = 0$ für $x \in [-1,0)$. Insgesamt haben wir gezeigt

Satz 1.1.12. *Der normierte Raum* $(C([a,b]), \|\cdot\|_1)$ *ist nicht vollständig.*

Somit "fehlen" in $C([a,b])$ Elemente um diesen Raum bezüglich $\|\cdot\|_1$ vollständig zu machen. Man könnte nun versuchen, den Raum $C([a,b])$ durch den Raum $R([a,b])$ der auf $[a,b]$ Riemann-integrierbaren Funktionen zu ersetzen. In der Tat ist $f := \mathbb{1}_{(0,1]} \in R([-1,1])$ und es gilt für die oben betrachtete Folge f_n, dass $f_n \to f$ bezüglich $\|\cdot\|_1$. Allerdings ist $\|\cdot\|_1$ auf $R([a,b])$ keine Norm mehr, sondern lediglich eine Halbnorm (etwa ist $\|\mathbb{1}_{\{0\}}\| = 0$ obwohl $\mathbb{1}_{\{0\}} \neq 0$ ist).

Dieses Problem könnte man dadurch lösen, dass man Funktionen mit gleicher Norm *identifiziert*, d.h. man geht zu Äquivalenzklassen über. Dies kann man in der Tat machen, allerdings ist der Raum, den man erhält immer noch unvollständig. Um vollständige Räume integrierbarer Funktionen zu erhalten, bedient man sich der *Lebesgue'schen Integrationstheorie.*

Sei hierzu (Ω, Σ, μ) ein Maßraum, d.h. Σ ist eine σ-Algebra auf Ω und $\mu : \Sigma \to [0,\infty]$ ist ein Maß auf Σ, also eine Abbildung derart, dass für paarweise disjunkte Mengen $A_n \in \Sigma$ stets $\mu(\bigcup_n A_n) = \sum_\mu (A_n)$ gilt.

Proposition 1.1.13. *Für* $p \in [1,\infty)$ *ist die Menge* $\mathscr{L}^p(\Omega, \Sigma, \mu)$, *definiert als*

$$\mathscr{L}^p(\Omega) := \{f : \Omega \to \mathbb{K} \text{ messbar} : \int_\Omega |f|^p \, d\mu < \infty\}$$

ein Vektorraum. Weiter ist $\|\cdot\|_p^*$, *definiert durch*

$$\|f\|_p^* := \Big(\int_\Omega |f(x)|^p \, d\mu(x)\Big)^{\frac{1}{p}},$$

eine Halbnorm auf $\mathscr{L}^p(\Omega, \Sigma, \mu)$.

Ist der Maßraum (Ω, Σ, μ) aus dem Zusammenhang klar, so schreiben wir oft nur $\mathscr{L}^p(\Omega)$ oder sogar lediglich $\mathscr{L}^p$.

Der Beweis von Proposition 1.1.13 beruht größtenteils auf elementaren Eigenschaften messbarer Funktionen. Lediglich die Dreiecksungleichung im Fall $p \in (1,\infty)$ bedarf größerer Aufmerksamkeit. Wir wiederholen die Details aufgrund der Wichtigkeit für die Funktionalanalysis.

Satz 1.1.14. *Es sei* (Ω, Σ, μ) *ein Maßraum,* $p, q \in (1,\infty)$ *mit* $\frac{1}{p} + \frac{1}{q} = 1$.

(1) (Hölderungleichung) Ist $f \in \mathscr{L}^p(\Omega)$ *und* $g \in \mathscr{L}^q(\Omega)$, *so ist* $fg \in \mathscr{L}^1(\Omega)$. *Weiter ist*

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

(2) (Minkowskiungleichung) Sind $f, g \in \mathscr{L}^p$, so ist $f + g \in \mathscr{L}^p(\Omega)$ und

$$\|f + g\|_p^* \leq \|f\|_p^* + \|g\|_p^*.$$

Beweis. (1) Zunächst zeigen wir, dass $x^r y^{1-r} \leq rx + (1-r)y$ für $x, y \geq 0$ und $r \in (0,1)$. Ist $x = 0$ oder $y = 0$, so ist dies offensichtlich. Ansonsten ist die Behauptung äquivalent zur Konkavität des Logarithmus:

$$\log(x^r y^{1-r}) = r \log x + (1-r) \log y \leq \log(rx + (1-r)y).$$

Diese wiederum folgt aus der Tatsache, dass $\log''(t) = -t^{-2} < 0$ für $t > 0$.

Seien nun $f \in \mathscr{L}^p(\Omega)$ und $g \in \mathscr{L}^q(\Omega)$. Definiere $A := (\|f\|_p^*)^p$ und $B := (\|g\|_q^*)^q$. Ohne Einschränkung seien $A, B > 0$. Verwenden wir obige Aussage für $r = p^{-1}$ (also $1 - r = 1 - p^{-1} = q^{-1}$) und $x = A^{-1}|f(\omega)|^p$, $y = B^{-1}|g(\omega)|^q$ an, so folgt

$$\Big(\frac{|f(\omega)|^p}{A}\Big)^{\frac{1}{p}} \Big(\frac{|g(\omega)|^q}{B}\Big)^{\frac{1}{q}} \leq \frac{1}{p}\frac{f(\omega)^p}{A} + \frac{1}{q}\frac{g(\omega)^q}{B}.$$

Da A und B nach Voraussetzung endlich sind, folgt bei Integration über Ω die Integrierbarkeit von fg aus der Monotonie des Integrals. Weiterhin folgt

$$\frac{1}{A^{\frac{1}{p}} B^{\frac{1}{q}}} \|fg\|_1^* \leq \frac{1}{p}\frac{1}{A}(\|f\|_p^*)^p + \frac{1}{q}\frac{1}{B}(\|g\|_q^*)^q.$$

Nehmen wir an, dass $A = B = 1$ ist, so folgt $\|fg\|_1^* \leq p^{-1} + q^{-1} = 1 = \|f\|_p^* \|g\|_q^*$. Für den allgemeinen Fall ersetzen wir f durch $\tilde{f} = A^{-1} f$ und $\tilde{g} = B^{-1} g$.

(2) Aus der Abschätzung $|f + g|^p \leq 2^p(|f|^p + |g|^p)$ folgt sofort, dass mit f und g die Funktion $f + g$ auch p-integierbar ist.

Sei nun $q^{-1} = 1 - p^{-1} = \frac{p-1}{p}$ also $(p-1)q = p$. Aus Teil (1) folgt

$$\begin{aligned}
\big(\|f + g\|_p^*\big)^p &= \int_\Omega |f(\omega) + g(\omega)|^p \, d\mu(\omega) \\
&\leq \int_\Omega |f(\omega)||f(\omega) + g(\omega)|^{p-1} \, d\mu(\omega) + \int_\Omega |g(\omega)||f(\omega) + g(\omega)|^{p-1} \, d\mu(\omega) \\
&\leq (\|f\|_p^* + \|g\|_p^*)(\|(f+g)^{p-1}\|_q^* \\
&= (\|f\|_p^* + \|g\|_q^*)(\|f + g\|_p^*)^{p-1},
\end{aligned}$$

was äquivalent zur Aussage ist. □

Um die Vollständigkeit von $\mathscr{L}^p(\Omega)$ zu zeigen, verwenden wir folgendes Lemma.

Lemma 1.1.15. *Ein halbnormierter Raum $(X, \|\cdot\|)$ ist genau dann vollständig, wenn jede absolut konvergente Reihe konvergiert, d.h. ist (x_n) eine Folge in X mit $\sum_n \|x_n\| < \infty$, so gibt es ein Element $x \in X$ mit $\sum_{k=1}^n x_k \to x$.*

Beweis. Aufgrund der Dreiecksungleichung ist $\| \sum_{k=n}^m x_k\| \leq \sum_{k=n}^m \|x_k\|$. Ist also die Reihe der x_k absolut konvergent, so bilden die Partialsummen $\sum_{k=1}^n x_k$ eine Cauchyfolge. Insbesondere konvergieren die Partialsummen wenn X vollständig ist.

Nimm nun umgekehrt an, dass in X jede absolut konvergente Reihe konvergiert. Sei weiter eine Cauchyfolge (x_k) gegeben. Zu $\varepsilon_k := 2^{-k}$ finden wir ein N_k derart, dass

$$\|x_n - x_m\| \leq 2^{-k}$$

für $n, m \geq N_k$. Insbesondere finden wir eine Teilfolge x_{n_k} mit

$$\|x_{n_{k+1}} - x_{n_k}\| \leq 2^{-k}$$

für alle $k \in \mathbb{N}$. Für $y_k := x_{n_{k+1}} - x_{n_k}$ ist also $\sum_{k=1}^{\infty} \|y_k\| < \infty$. Nach Voraussetzung konvergieren die Partialsummen gegen ein $y \in X$. Es ist also

$$\sum_{k=1}^{r} y_k = x_{n_{r+1}} - x_1 \to y$$

und daher $x_{n_r} \to y + x_1$. Somit hat (x_n) eine konvergente Teilfolge und ist daher selbst konvergent. □

Satz 1.1.16. *Der halbnormierte Raum* $(\mathscr{L}^p(\Omega), \|\cdot\|_p^*)$ *ist vollständig.*

Beweis. Wir benutzen Lemma 1.1.15 und zeigen, dass in $\mathscr{L}^p(\Omega)$ jede absolut konvergente Reihe konvergiert. Sei hierzu eine Folge $(f_n) \subset \mathscr{L}^p(\Omega)$ mit $\sum_{k=1}^{\infty} \|f_k\|_p^* =: c < \infty$ gegeben. Wir definieren $h_n := \sum_{k=1}^{n} |f_k|$ und $h_\infty := \sup h_n$. Dann ist h_∞ als punktweiser Grenzwert messbarer Funktionen messbar (nimmt aber eventuell den Wert ∞ an). Aus der Dreiecksungleichung folgt

$$\|h_n\|_p^* \leq \sum_{k=1}^{n} \|f_k\|^* \leq c < \infty.$$

Aus dem Satz über die monotone Konvergenz folgt

$$\int_\Omega h_\infty^p \, d\mu = \sup_{n \in \mathbb{N}} \int_\Omega h_n^p \, d\mu \leq c^p < \infty.$$

Insbesondere ist h_∞ außerhalb einer Nullmenge (etwa außerhalb der Nullmenge N) endlich. Für $\omega \in \Omega \setminus N$ ist also

$$\sum_{k=1}^{\infty} |f_k(\omega)| = h_\infty(\omega) < \infty.$$

Weil der Skalarkörper $\mathbb{K}$ vollständig ist, gibt es für solche ω ein $f(\omega)$ mit

$$f(\omega) = \sum_{k=1}^{\infty} f_k(\omega).$$

Für $\omega \in N$ setzen wir noch $f(\omega) = 0$. Somit haben wir eine messbare Funktion f definiert. Nach Konstruktion gilt $|f| \leq h_\infty$ und daher liegt mit h_∞ auch f in $\mathscr{L}^p(\Omega)$. Es bleibt nur noch zu zeigen, dass $\|f - \sum_{k=1}^{n} f_k\|_p^* \to 0$. Sei hierzu $g_n := |f - \sum_{k=1}^{n-1} f_k|^p = |\sum_{k=n}^{\infty} f_k|^p$. Nach Konstruktion gilt $g_n \to 0$ fast überall. Weiter ist $0 \leq g_n \leq h_\infty^p$ und die letzte Funktion ist integrierbar. Nach dem Satz über die dominierte Konvergenz gilt also $\|g_n\|_1^* \to 0$, was zu zeigen war. □

Nun skizzieren wir noch, wie man aus dem halbnormierten Raum $\mathscr{L}^p(\Omega)$ einen normierten Raum $L^p(\Omega)$ konstruiert. Hierzu identifizieren wir Funktionen f und g mit $\|f - g\|_p^* = 0$. Dies ist äquivalent zu $f = g$ fast überall. Wir bilden Äquivalenzklassen

$$[f] = \{g \in \mathscr{L}^p(\Omega) \, : \, f = g \text{ fast überall}\}.$$

Sodann fassen wir diese Äquivalenzklassen als unsere primären Objekte auf:

$$L^p(\Omega) := \{[f] : f \in \mathscr{L}^p(\Omega)\}.$$

Die Vektorraumoperationen werden durch $[f] + [g] := [f + g]$ und $\lambda[f] = [\lambda f]$ erklärt (man überzeuge sich, dass dies wohldefiniert ist!). Zu guter Letzt definieren wir noch

$$\|[f]\|_p := \|f\|_p^*$$

Dies ist wohldefiniert, denn wenn $f = g$ fast überall ist, so ist $\|f\|_p^* = \|g\|_p^*$. Außerdem ist dies eine Norm, denn ist $\|[f]\|_p = 0$, so muss $f = 0$ fast überall sein, also $[f] = [0]$. Die anderen Eigenschaften einer Norm sind klar. Aus Satz 1.1.16 erhalten wir nun

Korollar 1.1.17. *Der Raum $(L^p(\Omega), \|\cdot\|_p)$ ist ein Banachraum.*

Bemerkung 1.1.18. Wie üblich behalten wir die Schreibweise $[f]$ für Elemente von $L^p(\Omega)$ *nicht* bei, sondern schreiben einfach f und operieren mit den Elementen von $L^p(\Omega)$ als wären sie Funktionen. Dies führt meist nicht zu Problemen, solange man beachtet, dass alle Gleichheiten lediglich fast überall zu verstehen sind.

Wir erwähnen noch einen wichtigen Spezialfall.

Korollar 1.1.19. *Der Raum ℓ^p der p-summierbaren Folgen*

$$\ell^p = \{\mathbf{x} = (x_n) \subset \mathbb{K} : \sum_{k=1}^{\infty} |x_k|^p < \infty\}$$

ist ein Banachraum bezüglich der Norm

$$\|\mathbf{x}\|_p = \Big(\sum_{k=1}^{\infty} |x_k|^p\Big)^{\frac{1}{p}}.$$

Beweis. Dies folgt aus Korollar 1.1.17 für den Maßraum $(\Omega, \Sigma, \mu) = (\mathbb{N}, \mathscr{P}(\mathbb{N}), \zeta)$, wobei ζ das Zählmaß ist. □

1.2 Beschränkte Operatoren und Funktionale

Definition 1.2.1. Es seien X, Y normierte Räume. Ein *beschränkter Operator* (oder auch *stetiger Operator*) von X nach Y ist eine lineare Abbildung $T : X \to Y$ die stetig ist. Gilt also $x_n \to x$ in X, so folgt $T(x_n) \to T(x)$ in Y. Wir schreiben $\mathscr{L}(X, Y)$ für den Vektorraum (!) der beschränkten Operatoren von X nach Y. Für $Y = \mathbb{K}$ spricht man auch von *stetigen Funktionalen* und nennt den Raum $\mathscr{L}(X, \mathbb{K})$ den *Dualraum* von X. Notation: $X^* := \mathscr{L}(X, \mathbb{K})$.

Bemerkung 1.2.2. Sind X und Y endlichdimensionale Räume, so ist jede lineare Abbildung T stetig. Nach Wahl von Basen kann man eine lineare Abbildung mit ihrer Darstellungsmatrix A identifizieren, d.h. $T(x) = Ax$. Es ist üblich auch in unendlicher Dimension Tx statt $T(x)$ zu schreiben.

Proposition 1.2.3. *Es seien $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ normierte Räume und T eine lineare Abbildung von X nach Y. Folgende Aussagen sind äquivalent.*

(i) T ist stetig, d.h. $x_n \to x$ impliziert $Tx_n \to Tx$.

(ii) T *ist stetig in* 0, *d.h.* $x_n \to 0$ *impliziert* $Tx_n \to 0$.

(iii) Es gibt eine Konstante L *derart, dass* $\|Tx\|_Y \le L\|x\|_X$ *für alle* $x \in X$ *ist.*

(iv) $\|T\|_{\mathrm{op}} := \sup\{\|Tx\|_Y : \|x\|_X \le 1\} < \infty$.

Beweis. (i) $\Rightarrow$ (ii) ist trivial.

Angenommen, es gäbe keine Konstante L wie in (iii). Dann gäbe es zu $n \in \mathbb{N}$ ein Element x_n mit $\|Tx_n\| \ge n\|x_n\|$. Wir können ohne Einschränkung annehmen, dass $\|x_n\| = 1$ ist, sonst teilen wir durch die Norm. Es folgt $y_n := n^{-1}x_n \to 0$. Andererseits ist wegen der Linearität von T aber $Ty_n = n^{-1}Tx_n$, und somit $\|Ty_n\| = n^{-1}\|Tx_n\| \ge 1$, also $Ty_n \not\to 0$, im Widerspruch zu (ii). Dies zeigt (ii) $\Rightarrow$ (iii).

(iii) $\Leftrightarrow$ (iv): Es gilt offensichtlich $\|T\|_{\mathrm{op}} \le L < \infty$ für jedes L wie in (iii). Setzen wir $\tilde{x} = \|x\|_X^{-1}x$ für $x \neq 0$, so gilt umgekehrt

$$\|Tx\|_Y = \|x\|_X\|T\tilde{x}\|_Y \le \|T\|_{\mathrm{op}}\|x\|_X,$$

denn $\|\tilde{x}\|_X = 1$. Insgesamt ist also $\|T\|_{\mathrm{op}}$ das Infimum aller L für die die Abschätzung in (iii) gilt.

(iii) $\Rightarrow$ (i) Ist $x_n \to x$, so gilt $\|x_n - x\|_X \to 0$. Wegen (iii) folgt

$$\|Tx_n - Tx\|_Y = \|T(x_n - x)\|_Y \le L\|x_n - x\|_X \to 0,$$

also $Tx_n \to Tx$. □

Proposition 1.2.4. *Es seien* X *und* Y *normierte Räume. Dann definiert* $\|\cdot\|_{\mathrm{op}}$ *eine Norm auf* $\mathscr{L}(X,Y)$ *– die sogenannte* Operatornorm. *Ist* Y *vollständig, so ist auch* $\mathscr{L}(X,Y)$ *vollständig. Insbesondere ist* X^* *stets ein Banachraum.*

Beweis. (ÜA) □

Im Folgenden werden wir meist nicht zwischen der Norm auf den Räumen X und Y und der Operatornorm auf $\mathscr{L}(X,Y)$ unterscheiden, da aus dem Zusammenhang klar ist, welche Norm gemeint ist.

Die Operatornorm hat eine weitere wichtige Eigenschaft, sie ist *submultiplikativ.*

Lemma 1.2.5. *Es seien* X, Y *und* Z *normiertere Räume. Sind* $T \in \mathscr{L}(X,Y)$ *und* $S \in \mathscr{L}(Y,Z)$, *so ist* $ST := S \circ T \in \mathscr{L}(X,Z)$ *und* $\|ST\| \le \|S\| \cdot \|T\|$.

Beweis. Für $x \in Z$ ist $\|STx\| \le \|S\|\|Tx\| \le \|S\|\|T\|\|x\|$, woraus sofort die Behauptung folgt. □

Beispiel 1.2.6. (Multiplikationsoperatoren)

Es sei $\mathbf{m} \in \ell^\infty$. Dann definiert $M : \ell^p \to \ell^p$, $M\mathbf{x} = (m_n x_n)_{n\in\mathbb{N}}$ einen beschränkten Operator auf ℓ^p. Weiter ist $\|M\| = \|\mathbf{m}\|_\infty$.

Beweis. Offensichtlich ist M linear. Weiter gilt

$$\|M\mathbf{x}\|_p^p = \sum_{k=1}^\infty |m_k x_k|^p \le \sum_{k=1}^\infty \sup_{j\ge 1} |m_n|^p |x_k|^p = \|\mathbf{m}\|_\infty^p \|\mathbf{x}\|_p^p.$$

Es folgt, dass M stetig ist mit $\|M\| \le \|\mathbf{m}\|_\infty$. Andererseits gilt für $\mathbf{e}_n = (\delta_{kn})_{k\in\mathbb{N}}$ dass $\|\mathbf{e}_n\|_p = 1$. Somit ist $\|M\| \ge \|M\mathbf{e}_n\|_p = |m_n|$ für alle $n \in \mathbb{N}$. Es folgt $\|M\| \ge \|\mathbf{m}\|_\infty$. □

Beispiel 1.2.7. (Punktauswertungen)
Wir betrachten den Raum $C(K)$ für ein Kompaktum $K \subset \mathbb{R}^d$. Für $x_0 \in K$ ist die Abbildung $\delta_{x_0} : C(K) \to \mathbb{K}$, definiert durch $\delta_{x_0}(f) = f(x_0)$, ein beschränktes Funktional auf $C(K)$.

Beweis. Offenbar ist δ_{x_0} linear. Weiter ist

$$|\delta_{x_0}(f)| = |f(x_0)| \le \sup_{x\in K} |f(x)| = \|f\|_\infty.$$

Dies zeigt, dass δ_{x_0} beschränkt ist mit $\|\delta_{x_0}\| \le 1$. Setzt man $f \equiv 1$ ein, so sieht man, dass $\|\delta_{x_0}\| = 1$ ist. □

Wir betrachten nun noch eine für die Anwendungen besonders wichtige Klasse von Beispielen – die sogenannten Integraloperatoren. Diese Operatoren treten in natürlicher Weise in Integralgleichungen auf und das Studium solcher Gleichungen war eine der Hauptmotivationen für die Entwicklung der Funktionalanalysis.

Beispiel 1.2.8. (Integraloperatoren)
Wir betrachten den Raum $L^2((0,1),\lambda)$, wobei λ das Lebesguemaß auf der Borel σ-Algebra von $(0,1)$ ist. Sei weiter $k \in L^2((0,1)^2, \lambda^2)$, wobei λ^2 das zweidimensionale Lebesguemaß ist. Es folgt aus dem Satz von Fubini, dass $k(x,\cdot) \in L^2((0,1),\lambda)$ für fast alle $x \in (0,1)$. Aus der Hölder'schen Ungleichung folgt, dass für $f \in L^2(0,1)$ die Funktion $k(x,\cdot)f(\cdot)$ in $L^1(0,1)$ liegt für fast alle $x \in (0,1)$. Somit können wir für $f \in L^2(\Omega)$ und $x \in (0,1)$

$$T_k f(x) := \int_0^1 k(x,y)f(y)\,d\lambda(y)$$

definieren. Zunächst ist $(T_k f)(x)$ nur für fast alle x wohl-definiert, für andere x setzen wir $T_k f(x) = 0$. Als Konsequenz des Satzes von Fubini ist $T_k f(x)$ messbar. Mit Hölder und Fubini folgt

$$\begin{aligned}
\|T_k f\|_2^2 &= \int_0^1 \Big| \int_0^1 k(x,y)f(y)\,d\lambda(y)\Big|^2 d\lambda(x)\\
&\le \int_0^1 \Big(\int_0^1 |k(x,y)|\cdot|f(y)|\,d\lambda(y)\Big)^2 d\lambda(x)\\
&\le \int_0^1 \Big(\int_0^1 |k(x,y)|^2\,d\lambda(y)\Big)\Big(\int_0^1 |f(y)|\,d\lambda(y)\Big)\,d\lambda(x)\\
&\le \|f\|_2^2 \int_0^1 \int_0^1 |k(x,y)|^2\,d\lambda(x)d\lambda(y).
\end{aligned}$$

Somit ist T_k beschränkt mit $\|T_k\| \le \|k\|_{L^2((0,1)^2)}$. Allerdings gilt hier im Allgemeinen keine Gleichheit.

Beispiel 1.2.9. (Ein unstetiger linearer Operator)
Wir betrachten den Raum $C^1([0,1])$ der stetig differenzierbaren Funktionen auf $[0,1]$, versehen mit der Supremumsnorm. Dann ist der lineare Operator $D : C^1([0,1]) \to C([0,1])$, $Df := f'$ nicht stetig. In der Tat konvergiert $f_n(t) := n^{-1}t^n$ gegen 0, aber $Df_n(t) = t^{n-1}$ erfüllt $\|Df_n\| \equiv 1$, insbesondere gilt $Df_n \not\to 0$.

Man kann den Operator D auch als Operator auf $C([0,1])$ auffassen, der allerdings nicht auf dem gesamten Raum definiert ist, sondern nur auf dem Teilraum $C^1([0,1])$. In dieser Situation ist man häufig (insbesondere in Anwendungen auf partielle Differentialgleichungen).

Definition 1.2.10. Es seien X und Y normierte Räume. Ein *Operator* ist eine lineare Abbildung $A : \mathrm{def}(A) \to Y$, wobei der *Definitionsbereich* $\mathrm{def}(A)$ ein Teilraum von X ist. Der Operator A heißt abgeschlossen, falls für jede Folge $(x_n) \subset \mathrm{def}(A)$ aus $x_n \to x$ und $Ax_n \to y$ stets $x \in \mathrm{def}(A)$ und $Ax = y$ folgt.

Beispiel 1.2.11. Auf $X = C([a,b])$ ist der Operator D, gegeben durch $\mathrm{def}(D) = C^1([0,1])$ und $Df = f'$ ist abgeschlossen. Ist nämlich f_n eine Folge stetig differenzierbarer Funktionen und konvergieren $f_n \to f$ und $f_n' \to g$ gleichmässig, so ist f stetig differenzierbar und $f' = g$, wie man in der Analysis lernt.

Definition 1.2.12. Es seien $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ normierte Räume. Ein *Isomorphismus* ist eine bijektive Abbildung $T : X \to Y$ derart, dass T und T^{-1} beschränkte Operatoren sind. Gibt es einen Isomorphismus zwischen X und Y, so sagt man X und Y seien *isomorph*. Eine *Isometrie* ist eine Abbildung $T : X \to Y$ mit $\|Tx\|_Y = \|x\|_X$ für alle $x \in X$. Ist T ein Isomorphismus und eine Isometrie, so nennt man T einen *isometrischen Isomorphismus* und sagt X und Y seien *isometrisch isomorph*.

Bemerkung 1.2.13. Sind durch $\|\cdot\|_1$ und $\|\cdot\|_2$ zwei Normen auf dem Vektorraum X gegeben, so ist die Identität ein genau dann ein Isomorphismus zwischen $(X, \|\cdot\|_1)$ und $(X, \|\cdot\|_2)$, wenn es Konstanten $c_1, c_2 > 0$ gibt mit

$$c_1\|x\|_1 \leq \|x\|_2 \leq c_2\|x\|_1$$

für alle $x \in X$. In diesem Fall sagt man, die Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ sind *äquivalent*.

Auf endlichdimensionalen Vektorräumen sind je zwei Normen äquivalent. Manchmal beweist man dieses Ergebnis in Analysis 2. Wir geben der Vollständigkeit halber einen Beweis an.

Proposition 1.2.14. *Ist X ein endlichdimensionaler (reeller oder komplexer) Vektorraum, so sind je zwei Normen auf X äquivalent.*

Beweis. Es sei $\mathfrak{v} = (v_1, \ldots, v_d)$ eine Basis von X und $T_\mathfrak{v} : \mathbb{K}^d \to X$ definiert durch $T_\mathfrak{v}\alpha := \sum_{j=1}^d \alpha_j v_j$. Dann ist $T_\mathfrak{v}$ bijektiv. Wir definieren die Norm (!) $\|\cdot\|_e$ auf X durch $\|x\|_e := \|T_\mathfrak{v}^{-1}x\|_\infty$. Ist nun $\|\cdot\|$ eine Norm auf X so genügt es zu zeigen, dass $\|\cdot\|$ und $\|\cdot\|_e$ äquivalent sind.

Beachte, dass $T_\mathfrak{v}$ eine stetige Abbildung von $(\mathbb{K}^d, \|\cdot\|_\infty)$ nach $(X, \|\cdot\|)$ ist. In der Tat ist nämlich

$$\|T_\mathfrak{v}\alpha - T_\mathfrak{v}\beta\| = \Big\|\sum_{k=1}^d (\alpha_k - \beta_k)v_k\Big\| \leq \|\alpha - \beta\|_\infty \sum_{k=1}^n \|v_k\|.$$

Also ist auch die Abbildung $f := \|\cdot\| \circ T_\mathfrak{v}$ stetig. Da die Menge $K := \{\alpha : \|\alpha\|_\infty = 1\}$ kompakt ist, nimmt f auf K Minimum und Maximum an, es gibt also c_1, c_2 mit $c_1 \leq f(\alpha) \leq c_2$, also $c_1 \leq \|T_\mathfrak{v}\alpha\| \leq c_2$ für alle $x = T_\mathfrak{v}\alpha$ mit $\|x\|_e = 1$. Beachte, dass $c_1 > 0$ ist, denn sonst gäbe es ein $\alpha \in K$ mit $\|T_\mathfrak{v}\alpha\| = 0$. Also, weil $\|\cdot\|$ definit ist, wäre $T_\mathfrak{v}\alpha = 0$, was $\alpha = 0$ nach sich zieht, denn $T_\mathfrak{v}$ ist eine bijektive lineare Abbildung. Dies widerspricht aber der Bedingung $\|\alpha\| = 1$.

Verwendet man nun die Homogenität von $\|\cdot\|$, so folgt

$$c_1\|T_\mathfrak{v}\alpha\|_e \leq \|T_\mathfrak{v}\alpha\| \leq c_2\|T_\mathfrak{v}\alpha\|_e$$

für alle $\alpha \in \mathbb{K}^d$. □

Sind zwei Räume isometrisch isomorph, so sind sie (vom mathematischen Standpunkt her) als normierte Räume nicht zu unterscheiden. Die Objekte sind lediglich anders benannt worden. Wir geben hierzu ein Beispiel, in dem wir den Dualraum eines normierten Raumes identifizieren.

Satz 1.2.15. *Der Dualraum von ℓ^1 ist isometrisch isomomorph zu ℓ^∞.*

Beweis. Wir definieren $\Phi : \ell^\infty \to (\ell^1)^*$ durch $\big[\Phi(\mathbf{x})\big](\mathbf{y}) := \sum_{k=1}^n x_k y_k$, d.h. das Funktional $\Phi(\mathbf{x})$ bildet das Element $\mathbf{y}$ auf $\sum_{k=1}^\infty x_k y_k$ ab. Beachte zunächst, dass

$$|\Phi(\mathbf{x})(\mathbf{y})| = \Big|\sum_{k=1}^\infty x_k y_k\Big| \leq \sum_{k=1}^\infty |y_k| \sup_{j\geq 1} |x_j| = \|\mathbf{x}\|_\infty \|\mathbf{y}\|_1$$

ist. Somit ist $\Phi(\mathbf{x})$ in der Tat ein beschränktes Funktional und es gilt $\|\Phi(\mathbf{x})\|_{(\ell^1)^*} \leq \|\mathbf{x}\|_\infty$. Indem man die Vektoren $\mathbf{e}_n$ einsetzt sieht man wie in Beispiel 1.2.6, dass hier Gleichheit gilt. Insbesondere folgt, dass Φ injektiv ist. Es bleibt somit nur noch zu zeigen, dass Φ surjektiv ist.

Sei hierzu $\varphi \in (\ell^1)^*$. Wir setzen $x_k := \varphi(\mathbf{e}_k)$. Dann ist $|x_k| \leq \|\varphi\|\|\mathbf{e}_k\|_1 = \|\varphi\|$. Insbesondere ist $\mathbf{x} = (x_k) \in \ell^\infty$. Sei nun $\mathbf{y} = (y_k) \in \ell^1$ gegeben. Wir setzen $\mathbf{y}_n := \sum_{j=1}^n y_j \mathbf{e}_j$. Wegen der Linearität von φ und $\Phi(\mathbf{x})$ ist

$$\varphi(\mathbf{y}_n) = \sum_{j=1}^n y_j \varphi(\mathbf{e}_j) = \sum_{j=1}^n y_j x_j = \big[\Phi(\mathbf{x})\big](\mathbf{y}_n).$$

Nun beachte, dass $\mathbf{y}_n \to \mathbf{y}$ in ℓ^1. Wegen der Stetigkeit von φ und $\Phi(\mathbf{x})$ folgt

$$\varphi(\mathbf{y}) = \lim_{n\to\infty} \varphi(\mathbf{y}_n) = \lim_{n\to\infty} \big[\Phi(\mathbf{x})\big](\mathbf{y}_n) = \big[\Phi(\mathbf{x})\big](\mathbf{y}).$$

Somit gilt $\varphi = \Phi(\mathbf{x})$. Dies zeigt die Surjektivität von Φ. $\square$

Kapitel 2

Hilberträume

Wir betrachten nun eine spezielle Klasse von Banachräumen, nämlich *Hilberträume*, in denen eine zusätzliche Struktur – ein Skalarprodukt – vorhanden ist. Hilberträume sind den endlichdimensionalen Räumen sehr ähnlich. Beispielsweise kann man immer in einer Orthonormalbasis rechnen. Für uns besonders wichtig ist der Projektionssatz 2.3.1, der besagt dass es in einer konvexen abgeschlossenen Menge stets genau einen Punkt gibt, der einem vorgegebenen anderen Punkt am nächsten liegt. Dies hat erstaunliche Konsequenzen. Beispielsweise kann man einen Hilbertraum stets mit seinem Dualraum identifizieren.

2.1 Unitäre Räume

Definition 2.1.1. Es sei H ein reeller oder komplexer Vektorraum. Ein *Skalarprodukt* auf H ist eine Abbildung $\langle\,\cdot\,,\,\cdot\,\rangle : H \times H \to \mathbb{K}$ derart, dass für $x, y, z \in H$ und $\lambda \in \mathbb{K}$ folgende Eigenschaften gelten.

(SP1) *Definitheit*: $\langle x, x\rangle \geq 0$ und $\langle x, x\rangle = 0$ impliziert $x = 0$.

(SP2) *Symmetrie*: $\langle x, y\rangle = \overline{\langle y, x\rangle}$.

(SP3) *Linearität*: $\langle x + \lambda y, z\rangle = \langle x, z\rangle + \lambda\langle y, z\rangle$.

Das Paar $(H, \langle\,\cdot\,,\,\cdot\,\rangle)$ heißt *unitärer Raum* oder *Prähilbertraum*.

Bemerkung 2.1.2. Wegen (SP2) ist $\langle x, x\rangle = \overline{\langle x, x\rangle}$ also $\langle x, x\rangle \in \mathbb{R}$. Aus (SP2) und (SP3) folgt, dass das Skalarprodukt in der zweiten Komponente "fast linear" ist, es gilt nämlich $\langle x, y + \lambda z\rangle = \langle x, y\rangle + \bar{\lambda}\langle x, z\rangle$. Man muss also Skalare "konjugiert komplex herausziehen". Manchmal nennt man eine solche Abbildung *antilinear*. Vom Skalarprodukt selbst sagt man, es sei *sesquilinear* (sesqui = $1\frac{1}{2}$). Ist der Skalarkörper $\mathbb{R}$, so muss man nicht konjugieren und das Skalarprodukt ist in beiden Komponenten linear, also *bilinear*.

Beispiele 2.1.3. (1) Auf $H = \mathbb{K}^d$ ist durch

$$\langle x, y\rangle = \sum_{k=1}^{d} x_k \bar{y}_k$$

ein Skalarprodukt erklärt.

(2) Auf $H = \ell^2$ ist durch

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{k=1}^{\infty} x_k \bar{y}_k$$

ein Skalarprodukt erklärt.

(3) Auf $C([a,b])$ ist durch

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

ein Skalarprodukt erklärt.

(4) Ist (Ω, Σ, μ) ein Maßraum, so definiert

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

ein Skalarprodukt auf $L^2(\Omega, \Sigma, \mu)$.

Interpretation 2.1.4. Wir betrachten den Raum $H = \mathbb{R}^2$ mit dem Standardskalarprodukt $\langle x, y \rangle = x_1 y_2 + x_2 y_2$. Stellen wir uns den Vektor $x = (x_1, x_2)$ als Pfeil vom Koordinatenursprung nach (x_1, x_2) vor, so folgt aus dem Satz des Pythagoras, dass $\sqrt{\langle x, x \rangle}$ die *Länge* des Vektors ist.

Seien nun $x, y \in H$ mit $\langle x, x \rangle = \langle y, y \rangle = 1$ gegeben. Dann gibt es Winkel $\phi, \psi \in [0, 2\pi)$ derart, dass $x = (\cos\phi, \sin\phi)$ und $y = (\cos\psi, \sin\psi)$. Aus den Additionstheoremen für den Kosinus folgt, dass

$$\langle x, y \rangle = \cos\phi\cos\psi + \sin\phi\sin\psi = \cos(\phi - \psi).$$

Somit ist $\langle x, y \rangle$ der *Kosinus des Winkels zwischen x und y*. Insbesondere ist $\langle x, y \rangle = 0$ genau dann, wenn x und y senkrecht aufeinander stehen.

Wir wollen das Skalarprodukt in allgemeinen (auch abstrakten) unitären Räumen ähnlich interpretieren. Natürlich gibt es in solchen Räumen a priori keinen Begriff von Länge und Winkel. Wir definieren:

Definition 2.1.5. Es sei $(H, \langle\,\cdot\,,\,\cdot\,\rangle)$ ein unitärer Raum. Wir setzen $\|x\| := \sqrt{\langle x, x \rangle}$. Weiter sagen wir dass x und y *senkrecht aufeinander stehen*, falls $\langle x, y \rangle = 0$; Notation: $x \perp y$. Für eine Menge $M \subset H$ setzen wir $M^\perp := \{x \in H : x \perp y \;\forall\, y \in M\}$.

Beachte, dass es bislang nicht gezeigt wurde, dass $\|\cdot\|$ eine Norm auf H definiert. Allerdings folgen die Definitheit und die Homogenität von $\|\cdot\|$ direkt aus den Eigenschaften eines Skalarprodukts. Es bleibt also die Dreiecksungleichung zu zeigen. Zunächst zeigen wir jedoch

Satz 2.1.6. *(Satz von Pythagoras)*
Sind x, y Elemente eines unitären Raumes und $x \perp y$, so ist

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

Beweis. Es ist

$$\|x + y\|^2 = \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle = \|x\|^2 + 0 + 0 + \|y\|^2 \qquad \square$$

Satz 2.1.7. *(Cauchy–Schwarz Ungleichung)*
Sind x, y Elemente eines unitären Raumes, so gilt

$$|\langle x, y\rangle| \leq \|x\|\|y\|,$$

wobei Gleichheit genau dann gilt, wenn x und y linear abhängig sind.

Beweis. Sei zunächst $\|y\| = 1$. Wir setzen $Px := \langle x, y\rangle y$. Dann ist $x - Px \perp Px$. In der Tat ist

$$\langle x - Px, Px\rangle = \langle x, Px\rangle - \langle Px, Px\rangle = \overline{\langle x, y\rangle}\langle x, y\rangle - \langle x, y\rangle\overline{\langle x, y\rangle}\langle y, y\rangle = 0.$$

Nach dem Satz von Pythagoras gilt also

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

Hierbei gilt Gleichheit genau dann, wenn $\|x - Px\| = 0$ gilt, also, wegen Definitheit des Skalarprodukts, genau dann, wenn $x = \langle x, y\rangle y$ gilt. In diesem Fall sind also x und y linear abhängig.

Für beliebiges $y \neq 0$ wende den ersten Teil auf $\tilde{y} = \|y\|^{-1}y$ an. Sind schließlich x und y linear abhängig, so ist es einfach zu sehen, dass Gleichheit in der Cauchy-Schwarz Ungleichung gilt. □

Korollar 2.1.8. $\|\cdot\|$ *ist eine Norm.*

Beweis. Wie bereits erwähnt, bleibt lediglich die Dreiecksungleichung zu zeigen. Aus der Cauchy–Schwarz Ungleichung folgt jedoch sofort

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

was äquivalent zur Dreiecksungleichung ist. □

Korollar 2.1.9. *Das Skalarprodukt ist stetig, d.h. aus $x_n \to x$ und $y_n \to y$ folgt $\langle x_n, y_n\rangle \to \langle x, y\rangle$.*

Beweis. Gilt $x_n \to x$ und $y_n \to y$, so sind die Folgen x_n und y_n beschränkt, etwa durch C. Es folgt

$$|\langle x_n, y_n\rangle - \langle x, y\rangle| = |\langle x_n - x, y_n\rangle + \langle x, y_n - y\rangle| \leq C\|x_n - x\| + C\|y_n - y\| \to 0. \qquad \square$$

Definition 2.1.10. Ein unitärer Raum $(H, \langle\,\cdot\,,\,\cdot\,\rangle)$ heißt *Hilbertraum*, falls H mit der Norm $\|\cdot\| := \sqrt{\langle\,\cdot\,,\,\cdot\,\rangle}$ vollständig, also ein Banachraum, ist.

Beispiele 2.1.11. Die unitären Räume $\mathbb{K}^d$, ℓ^2 und – für jeden Maßraum (Ω, Σ, μ) – der Raum $L^2(\Omega)$ sind Hilberträume. Der Raum $C([a, b])$ mit dem kanonischen Skalarprodukt ist kein Hilbertraum.

2.2 Orthonormalbasen

Definition 2.2.1. Es sei H ein unitärer Raum und $I = \mathbb{N}$ oder $I = \mathbb{Z}$ oder $I = \{1, \ldots, d\}$ eine Indexmenge. Eine Familie $\{e_j : j \in I\}$ heißt *orthonormal*, falls

$$\langle e_i, e_j \rangle = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}.$$

Es gilt also $\|e_i\| = 1$ für $i \in I$ und $e_i \perp e_j$ für $i \neq j$. Wir sagen auch $\{e_j : j \in I\}$ ist ein *Orthonormalsystem*.

Beispiel 2.2.2. (Endliches Orthonormalsystem)
Es sei H ein unitärer Raum und $\{e_1, \ldots, e_d\}$ eine orthonormale Familie. Weiter sei $E_d := \operatorname{span}\{e_1, \ldots, e_d\}$. Ist $x = \sum_{j=1}^d \lambda_j e_j \in E_d$, so folgt

$$\langle x, e_k \rangle = \sum_{j=1}^d \lambda_j \langle e_j, e_k \rangle = \lambda_k.$$

Insbesondere ist

$$x = \sum_{k=1}^d \langle x, e_k \rangle e_k$$

die eindeutige Entwicklung von x in der Basis (!) $e_1, \ldots, e_d$ von E_d.

Beispiel 2.2.3. (Trigonometrische Funktionen)
Betrachte den (komplexen) unitären Raum $L^2((0, 2\pi), \frac{dt}{2\pi})$. Für $k \in \mathbb{Z}$ sei $e_k(t) := e^{ikt}$. Dann ist

$$\langle e_k, e_l \rangle = \frac{1}{2\pi} \int_0^{2\pi} e^{ikt} e^{-ilt}\, dt = \begin{cases} \left[\frac{1}{2\pi i(k-l)} e^{i(k-l)t} \right|_0^{2\pi} = 0 & k \neq l \\ \frac{1}{2\pi} \int_0^{2\pi} 1\, dt = 1 & k = l. \end{cases}$$

Somit ist $\{e_k\}_{k \in \mathbb{Z}}$ ein Orthonormalsystem.

Es stellt sich die Frage, ob auch im zweiten Beispiel eine "Entwicklung in der Basis" gilt, ob also

$$x = \sum_{k \in \mathbb{Z}} \langle x, e_k \rangle e_k$$

gilt. Beachte, dass falls diese Gleichung für alle x gilt, jedes Element des Raumes durch (endliche) Linearkombinationen der e_k's approximiert werden kann. Wir geben dieser Eigenschaft einen Namen.

Definition 2.2.4. Es sei $(X, \|\cdot\|)$ ein normierter Raum. Eine Teilmenge M von X heißt *dicht*, falls $\overline{M} = X$ ist, also jeder Punkt aus X Grenzwert einer Folge in M ist. Eine Teilmenge U von X heißt *total*, falls $\operatorname{span} U$ dicht in X ist.

Lemma 2.2.5. *Es sei H ein unitärer Raum und $M \subset H$ dicht. Gilt $\langle x, y \rangle = 0$ für alle $y \in M$, so ist $x = 0$.*

Beweis. Es gibt eine Folge x_n in M, die gegen x konvergiert. Wegen der Stetigkeit des Skalarprodukts folgt $0 = \langle x, x_n \rangle \to \langle x, x \rangle = \|x\|^2$. Somit ist $\|x\| = 0$ und daher $x = 0$. □

Lemma 2.2.6. *$C([a,b])$ ist dicht in $L^p((a,b),\lambda)$ für $1 \le p < \infty$.*

Beweis. Es sei X der Abschluss von $C([a,b])$ in $L^p((a,b),\lambda)$. Dann ist X ein Untervektorraum von $L^p((a,b),\lambda)$ (ÜA). Es sei $\mathscr{G} := \{A \in \mathscr{B}((a,b)) : \mathbb{1}_A \in X\}$ die Menge derjenigen Borelmengen, deren Indikator in X liegt. Dann gehört jedes abgeschlossene Intervall $[a',b']$ zu $\mathscr{G}$, denn der Indikator einer solchen Menge ist punktweiser Grenzwert einer beschränkten Folge stetiger Funktionen, gehört also nach dem Satz über die dominierte Konvergenz zum Abschluss der stetigen Funktionen.

Wir behaupten, dass $\mathscr{G} = \mathscr{B}((a,b))$. Weil die abgeschlossenen Intervalle ein durschnittsstabiler Erzeuger der Borel σ-Algebra sind, genügt es nach Dynkin's π-λ Theorem zu zeigen, dass $\mathscr{G}$ ein Dynkin System ist. Offensichtlich ist $(a,b) \in \mathscr{G}$, da $\mathbb{1}_{[a,b]} \in C([a,b])$ ist. Konvergiert $f_n \to \mathbb{1}_A$, so konvergiert $1 - f_n \to \mathbb{1}_{A^c}$. Es folgt, dass $\mathscr{G}$ mit A auch A^c enthält. Sind schliesslich A_n paarweise disjunkt und in $\mathscr{G}$ enhalten, so ist $\sum_{j=1}^n \mathbb{1}_{A_j} = \mathbb{1}_{\bigcup_{j=1}^n A_j} \in X$, da X ein Vektorraum ist. Da X auch abgeschlossen ist, ist auch $\mathbb{1}_{\bigcup_{j=1}^\infty A_j} = \lim \mathbb{1}_{\bigcup_{j=1}^n A_j} \in X$. Somit ist $\mathscr{G}$ in der Tat ein Dynkinsystem und es folgt dass $\mathbb{1}_A$ für jede Borelmenge in X liegt.

Weil X ein Vektorraum ist, folgt dass X jede einfache Funktion enthält. Weil aber jede L^p-funktion der Grenzwert einfacher Funktionen ist, muss der abgeschlossene Raum X jede L^p Funktion enthalten, also $X = L^p(a,b)$. □

Bemerkung 2.2.7. Der gleiche Beweis zeigt, dass jeder Teilraum V von $L^p((a,b),\lambda)$ mit der Eigenschaft, dass es für $a < a' < b' < b$ eine Folge f_n in V gibt mit $f_n \to \mathbb{1}_{[a',b']}$ in $L^p((a,b))$, dicht ist. Beispiele für solche Räume V sind beispielsweise $C_c((a,b)$, die stetigen Funktionen mit kompaktem Träger (d.h. $f(t) = 0$ in einer Umgebung von a und b) oder aber Funktionen mit höherer Regularität, etwa C^k-Funktionen.

Definition 2.2.8. Es sei H ein unitärer Raum. Eine *Orthonormalbasis* (ONB) ist ein totales Orthonormalsystem.

Bemerkung 2.2.9. Man kann zeigen, dass die Trigonometrischen Funktionenen $e_k(t) := e^{ikt}$ total in $L^2(0,2\pi)$ sind. Hierzu genügt es wegen Lemma 2.2.6 zu zeigen, dass man jede periodische, stetige Funktion (gleichmäßig) durch trigonometrische Polynome $\sum_{k=-n}^n a_k e^{ikt}$ approximieren kann. Dies folgt aus dem *Weierstraß'schen Approximationssatz.* Es gibt aber auch andere Möglichkeiten, dies zu zeigen.

Satz 2.2.10. *Es sei H ein uniärer Raum, $I = \mathbb{N}$, $I = \mathbb{Z}$ oder $I = \{1,\ldots,d\}$ und $(e_j)_{j\in I}$ eine Orthonormalbasis. Dann gilt für alle $x, y \in H$*

(1) $x = \sum_{j\in I} \langle x, e_j\rangle e_j$,

(2) $\|x\|^2 = \sum_{j\in I} |\langle x, e_j\rangle|^2$, *(Parseval Gleichung)*

(3) $\langle x, y\rangle = \sum_{j\in I} \langle x, e_j\rangle\langle e_j, y\rangle$.

Zum Beweis verwenden wir folgendes Lemma.

Lemma 2.2.11. *Es seien X, Y normierte Räume, T, T_n Elemente von $\mathscr{L}(X,Y)$. Weiter seien $\|T\|, \sup_{n\in\mathbb{N}} \|T_n\| \le C$ für eine geeignete Konstante $C > 0$. Schließlich gebe es eine dichte Teilmenge M von X, sodass $T_n x \to Tx$ für alle $x \in M$. Dann gilt $T_n x \to Tx$ für alle $x \in X$.*

Beweis. Es sei $x \in X$. Zu gegebenem $\varepsilon > 0$, finden wir ein $y \in M$ mit $\|x - y\| \leq \varepsilon/(3C)$. Weiter finden wir einen Index n_0 derart, dass $\|T_n y - Ty\| \leq \varepsilon/3$ für alle $n \geq n_0$. Für solche n gilt

$$\begin{aligned}\|T_n x - Tx\| &\leq \|T_n x - T_n y\| + \|T_n y - Ty\| + \|Ty - Tx\| \\ &\leq \|T_n\|\|x - y\| + \varepsilon/3 + \|T\|\|y - x\| \\ &\leq C\varepsilon/(3C) + \varepsilon/3 + C\varepsilon/(3C) = \varepsilon.\end{aligned}$$

Daher gilt $T_n x \to Tx$ wie behauptet. □

Beweis von Satz 2.2.10. (1) Es sei $I = \{i_n : n \in \mathbb{N}\}$ und $E_n := \operatorname{span}\{e_{i_1}, \ldots, e_{i_n}\}$. Wir setzen

$$P_n x := \sum_{k=1}^{n} \langle x, e_{i_k} \rangle e_{i_k}.$$

Dann ist $P_n x \in E_n$ und $x - P_n x \in E_n^\perp$. In der Tat gilt für $y = \sum_{l=1}^{n} \lambda_l e_{i_l} \in E_n$ stets

$$\langle x - P_n x, y \rangle = \sum_{l=1}^{n} \bar{\lambda}_l \langle x, e_{i_l} \rangle - \sum_{k=1}^{n} \sum_{l=1}^{n} \langle x, e_{i_k} \rangle \bar{\lambda}_l \langle e_{i_k}, e_{i_l} \rangle = 0.$$

Somit folgt aus dem Satz von Pythagoras, dass

$$\|x\|^2 = \|x - P_n x\|^2 + \|P_n x\|^2 \geq \|P_n x\|^2$$

und daher $\|P_n\| \leq 1$ für alle $n \in \mathbb{N}$. Für $x \in \operatorname{span}\{e_i : i \in I\}$ ist $P_n x \to x$, denn da e_i ein Orthonormalsystem ist, ist für solche x immer $P_n x \equiv x$ für genügend große n. Aus Lemma 2.2.11 folgt dass $P_n x \to x$ für alle $x \in H$. Das ist gerade Teil (1).

(2) Da die Norm stetig ist, folgt

$$\begin{aligned}\|x\|^2 &= \lim_{n\to\infty} \|P_n x\|^2 \\ &= \lim_{n\to\infty} \Big\langle \sum_{k=1}^{n} \langle x, e_{i_k} \rangle e_{i_k}, \sum_{l=1}^{n} \langle x, e_{i_l} \rangle e_{i_l} \Big\rangle \\ &= \lim_{n\to\infty} \sum_{k=1}^{n} |\langle x, e_{i_k} \rangle|^2 = \sum_{i\in I} |\langle x, e_i \rangle|^2.\end{aligned}$$

(3) folgt genauso, da das Skalarprodukt stetig ist. □

Bemerkung 2.2.12. Die Aussage von Satz 2.2.10 gilt auch für beliebige Indexmengen. Allerdings muss man in diesem Fall zunächst definieren, was unter einer “Reihe” mit eventuell überabzählbar vielen Glieder zu verstehen ist.

Bemerkung 2.2.13. Aus Satz 2.2.10 folgt insbesondere folgendes Resultat. Es sei H ein unitärer Raum mit ONB $(e_n)_{n\in\mathbb{N}}$. Definieren wir $U : H \to \ell^2$ durch $Ux = (\langle x, e_n \rangle)_{n\in\mathbb{N}}$, so ist dies wohldefiniert und sogar isometrisch, es gilt $\|Ux\|_{\ell^2} = \|x\|_H$ und sogar $\langle Ux, Uy \rangle_{\ell^2} = \langle x, y \rangle_H$. Es stellt sich die Frage, ob U auch bijektiv, also ein isometrischer Isomorphismus ist.

Definition 2.2.14. Es seien H_1, H_2 unitäre Räume. Ein *unitärer Operator* ist eine lineare Abbildung $U : H_1 \to H_2$ derart, dass $\langle Ux, Uy \rangle_{H_2} = \langle x, y \rangle_{H_1}$.

Lemma 2.2.15. *Es seien H und U wie in Bemerkung 2.2.13. Dann ist U genau dann ein isometrischer Isomorphismus, wenn H ein Hilbertraum ist.*

Beweis. Ist U surjektiv, also U ein isometrischer Isomorphismus, so muss H vollständig sein. Ist nämlich x_n eine Cauchyfolge in H, so ist Ux_n eine Cauchyfolge in ℓ^2, da U isometrisch ist. Weil ℓ^2 vollständig ist, gilt $Ux_n \to \mathbf{x}$ für ein geeignetes $\mathbf{x} \in \ell^2$. Wenn aber U surjektiv ist, ist $\mathbf{x} = Ux$ für ein $x \in H$ und dann $\|x_n - x\|_H = \|Ux_n - U\mathbf{x}\|_{\ell^2} \to 0$ und daher $x_n \to x$.

Ist umgekehrt H vollständig, so ist U surjektiv. Ist nämlich $\mathbf{x} \in \ell^2$, so gilt $\mathbf{x}_n := \sum_{j=1}^n x_n^{(j)} \mathbf{e}_j \to \mathbf{x}$. Offensichtlich ist aber $\mathbf{x}_n \in \mathrm{rg} U$, etwa $\mathbf{x}_n = Ux_n$. Weil $\mathbf{x}_n$ eine Cauchyfolge ist, ist auch x_n eine Cauchyfolge, also wegen der Vollständigkeit konvergent, etwa $x_n \to x$. Wegen Stetigkeit muss dann aber $\mathbf{x} = Ux$ sein, also U surjektiv. □

Korollar 2.2.16. *Sind H_1 und H_2 Hilberträume mit Orthonormalbasen $(e_k)_{k\in\mathbb{N}}$ bzw. $(v_k)_{k\in\mathbb{N}}$ so ist $U : H_1 \to H_2$ gegeben durch*

$$U : x = \sum_{k=1}^{\infty} x_k e_k \mapsto \sum_{k=1}^{\infty} x_k v_k$$

ein isometrischer Isomorphismus.

Beweis. Ist $U_1 : H_1 \to \ell^2$ gegeben durch $U_1 x = (\langle x, e_k \rangle)$ und $U_2 : H_2 \to \ell^2$ gegeben durch $U_2 x = (\langle x, v_k \rangle)$, so ist $U = U_2^{-1} U_1$ isometrischer Isomorphismus als Verknüpfung zweier isometrischer Isomorphismen. □

Bemerkung 2.2.17. Sind H_1 und H_2 unitäre Räume und $U : H_1 \to H_2$ ein isometrischer Isomorphismus, so zeigt man wie in Lemma 2.2.15, dass H_1 genau dann vollständig ist, wenn H_2 vollständig ist.

Es bleibt noch zu klären ob und wann ein unitärer Raum eine abzählbare Orthonormalbasis hat.

Definition 2.2.18. Ein normierter Raum heißt X *separabel*, wenn es eine abzählbare, dichte Menge in X gibt.

Lemma 2.2.19. *Ein normierter Raum ist genau dann separabel, wenn er eine abzählbare, totale Menge enthählt.*

Beweis. Jede abzählbare, dichte Menge ist automatisch total. Ist umgekehrt $\{x_n : n \in \mathbb{N}\}$ eine totale Teilmenge, so sei

$$F := \left\{ \sum_{k=1}^{n} q_k x_k \, n \in \mathbb{N}, q_1, \ldots, q_n \in \mathbb{K}_{\mathbb{Q}} \right\}$$

wobei $\mathbb{K}_{\mathbb{Q}} = \mathbb{Q}$, falls $\mathbb{K} = \mathbb{R}$ und $\mathbb{K}_{\mathbb{Q}} = \mathbb{Q} + i\mathbb{Q}$, falls $\mathbb{K} = \mathbb{C}$ ist. Dann ist F eine abzählbare Menge, die offensichtlich den Aufspann der (x_n) enthält. Da diese total sind, ist der Abschluss von F der gesamte Raum, der somit separabel ist. □

Nun gilt folgender Satz

Satz 2.2.20. *Jeder separable unitäre Raum besitzt eine abzählbare Orthonormalbasis.*

Beweis. Sei H ein unendlichdimensionaler, separabler, unitärer Raum. Dann gibt es eine Folge $\{x_n :\in \mathbb{N}\}$, die total in H ist. Ohne Einschränkung dürfen wir annehmen, dass die x_n linear unabhängig sind. Wir konstruieren nun eine Folge (e_n) die orthonormal ist mit $\operatorname{span}\{x_1, \ldots, x_n\} = \operatorname{span}\{e_1, \ldots, e_n\}$. Daraus folgt die Behauptung.

Wir konstruieren eine solche Folge mit dem *Gram-Schmidt'schen Orthonormalsierungsverfahren.* Setze hierzu $e_1 := \|x_1\|^{-1}x_1$. Sind $e_1, \ldots, e_n$ bereits konstruiert, so setze

$$y_{n+1} := x_{n+1} - \sum_{j=1}^{n} \langle x_{n+1}, e_j \rangle e_j$$

und dann $e_{n+1} := \|y_{n+1}\|^{-1}y_{n+1}$. Dann hat e_{n+1} Norm 1 und $e_{n+1} \perp e_k$ für $k = 1, \ldots, n$ wie man leicht sieht. Offensichtlich hat diese Folge die gewünschten Eigenschaften. $\square$

Korollar 2.2.21. *Jeder separable, unendlichdimensionale Hilbertraum ist isometrisch isomorph zu ℓ^2.*

2.3 Orthogonale Projektion

Es sei H ein Hilbertraum und $C \subset H$ eine abgeschlossene, konvexe Menge, d.h. sind $x, y \in C$ so ist auch $\lambda x + (1 - \lambda)y \in C$ für $\lambda \in [0, 1]$.

Gegeben ein Element x in H wollen wir nun ein Element Px von C finden, welches geringsten Abstand von x hat. Ein solches Element nennt man *Proximum* von x in C.

Satz 2.3.1. *Es sei H ein Hilbertraum und C eine abgeschlossene, konvexe Menge. Dann besitzt jedes Element x von H genau ein Proximum Px in C.*

Beweis. Es sei $x \in H$. Dann gibt es eine Folge $(x_n) \subset C$, mit

$$\|x - x_n\| \to \inf\{\|x - y\| : y \in C\} =: d.$$

Für $u, v \in H$ gilt

$$\|u + v\|^2 + \|u - v\|^2 = \langle u + v, u + v \rangle + \langle u - v, u - v \rangle = 2\|u\|^2 + 2\|v\|^2. \tag{2.1}$$

Wir verwenden dies für $u = x_n - x$ und $v = x - x_m$. Dann ist

$$\begin{aligned}
\|x_n - x_m\|^2 &= 2\|x_n - x\|^2 + 2\|x - x_m\|^2 - \|x_n - x - (x - x_m)\|^2 \\
&= 2\|x_n - x\|^2 + 2\|x - x_m\|^2 - 4\|\frac{1}{2}(x_n + x_m) - x\|^2 \\
&\le 2\|x_n - x\|^2 + 2\|x - x_m\|^2 - 4d^2,
\end{aligned}$$

da $\frac{1}{2}(x_n + x_m) \in C$ liegt. Somit ist

$$\limsup_{m,n\to\infty} \|x_n - x_m\|^2 \le 2d^2 + 2d^2 - 4d^2 = 0,$$

also ist (x_n) eine Cauchyfolge. Weil H vollständig ist, konvergiert x_n, etwa gegen $\bar{x}$ Da C abgeschlossen ist, liegt $\bar{x}$ in C. Weiter ist

$$\|x - \bar{x}\| = \lim_{n\to\infty} \|x - x_n\| = d.$$

Daher ist $\bar{x}$ ein Proximum von x in C.

Angenommen, $\hat{x}$ ist ein weiteres Proximum von x in C, also $\|x-\hat{x}\|^2 = d$. Es folgt aus (2.1), dass

$$\begin{aligned} d = \frac{1}{2}\|x-\bar{x}\|^2 + \frac{1}{2}\|x-\hat{x}\|^2 &= \|x - \frac{1}{2}(\hat{x}+\bar{x})\|^2 + \|\frac{1}{2}(\bar{x}-\hat{x})\|^2 \\ &\geq d + \|\frac{1}{2}(\bar{x}-\hat{x})\|^2. \end{aligned}$$

Somit ist $\|\hat{x}-\bar{x}\| = 0$, also $\hat{x} = \bar{x}$. □

Definition 2.3.2. Sei $x \in H$. Das eindeutige Element $Px \in C$ mit

$$\|x - Px\| = \min_{y \in C} \|x-y\|$$

heißt *orthogonale Projektion von x auf C*. Die Abbildung $P : H \to H$ heißt die *orthogonale Projektion auf C*.

Das folgende Lemma charakterisiert die orthogonale Projektion.

Lemma 2.3.3. *Es sei $\bar{x} \in C$. Dann sind folgende Aussagen äquivalent:*

(1) $\|x-\bar{x}\| \leq \|x-y\|$ für alle $y \in C$, d.h. $\bar{x}$ ist die orthogonale Projektion von x auf C.

(2) $\mathrm{Re}\,\langle x-\bar{x}, y-\bar{x}\rangle \leq 0$ für alle $y \in C$.

Beweis. (1) $\Rightarrow$ (2): Es sei $\bar{x}$ ein Proximum von x in C und $y \in C$. Da C konvex ist, liegt $\bar{x} + t(y-\bar{x})$ für $t \in [0,1]$ in C. Weil $\bar{x}$ ein Proximum ist, gilt

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

für $t \in [0,1]$. Daher gilt für solche t

$$2\mathrm{Re}\,\langle x-\bar{x}, y-\bar{x}\rangle \leq t\|y-\bar{y}\|^2$$

woraus für $t \downarrow 0$ (2) folgt.

(2) $\Rightarrow$ (1): Für $y \in C$ gilt

$$\begin{aligned} \|x-y\|^2 &= \|x-\bar{x}\|^2 + 2\mathrm{Re}\,\langle x-\bar{x}, \bar{x}-y\rangle + \|\bar{x}-y\| \\ &\geq \|x-\bar{x}\|. \end{aligned}$$

□

Wir zeigen nun, dass P stetig, sogar Lipschitz stetig ist. Beachte, dass P im Allgemeinen nicht linear ist.

Satz 2.3.4. *Es sei H ein Hibertraum und P die orthogonale Projektion auf die konvexe, abgeschlossene Menge C. Dann ist $\|Px - Py\| \leq \|x-y\|$.*

Beweis. Es folgt aus Lemma 2.3.3, dass

$$\mathrm{Re}\,\langle x-Px, z-Px\rangle \leq 0 \quad \mathrm{Re}\,\langle y-Py, z-Py\rangle \leq 0$$

für alle $z \in C$. Setzt man $z = Py$, respektive $z = Px$ ein, so erhält man

$$\operatorname{Re}\langle x - Px, Py - Px\rangle \leq 0 \qquad -\operatorname{Re}\langle y - Py, Py - Px\rangle \leq 0.$$

Addiert man diese Ungleichungen, so ergibt sich

$$\operatorname{Re}\langle x - Px + Py - y, Py - Px\rangle = \|Py - Px\|^2 + \operatorname{Re}\langle x - y, Py - Px\rangle \leq 0,$$

und daher, in Verbindung mit der Cauchy–Schwarz Ungleichung,

$$\|Py - Px\|2 \leq \operatorname{Re}\langle y - x, Py - Px\rangle \leq \|y - x\|\|Py - Px\|. \qquad \square$$

Wir betrachten nun noch den Spezialfall dass C ein abgeschlossener Teilraum von H ist. In der Formulierung des folgenden Satzes verwenden wir den Begriff *Hilbertraumsumme*. Sind H_1 und H_2 Hilberträume, so ist die *Hilbertraumsumme* das kartesische Produkt $H_1 \times H_2$, mit der Norm $\|(x_1, x_2)\|^2 := \|x_1\|_{H_1}^2 + \|x_2\|_{H_2}^2$. Wir schreiben $H_1 \oplus_2 H_2$ für die Hilbertraumsumme.

Satz 2.3.5. *Es sei H ein Hilbertraum und H_0 ein abgeschlossener Teilraum von H. Dann ist die orthogonale Projektion P auf H_0 ein linearer, beschränkter Operator mit $\|P\| = 1$ (es sei denn, $H_0 = 0$). Weiterhin ist H isometrisch isomorph zur Hilbertraumsumme $H_0 \oplus_2 H_0^\perp$.*

Beweis. Lemma 2.3.3 impliziert, dass $z = Px$ genau dann, wenn

$$\operatorname{Re}\langle x - z, y - z\rangle \leq 0$$

für alle $y \in H_0$. Da mit y auch $y + z$ ganz H_0 durchläuft (denn H_0 ist ein Unterraum!) ist obiges äquivalent mit $\operatorname{Re}\langle x - z, w\rangle \leq 0$ für alle $w \in H_0$. Ersetzt man w durch $-w$ und iw, so folgt, dass $z = Px$ ist genau dann, wenn $\langle x - z, w\rangle = 0$ für alle $w \in H_0$. Insbesondere liegt also $x - Px$ in $H_0^\perp$.

Da H ein Unterraum ist, liegen mit x_1 und x_2 auch $x_1 + \lambda x_2$ in H_0. Weiter ist

$$\langle (x_1 + \lambda x_2) - (Px_1 + \lambda Px_2), w\rangle = \langle x_1 - Px_1, w\rangle + \lambda\langle x_2 - Px_2, w\rangle = 0$$

für alle $w \in H_0$. Daher ist $P(x_1 + \lambda x_2) = Px_1 + \lambda Px_2$, P also linear. Gemäß Satz 2.3.4 ist P Lipschitzstetig, also $\|P\| \leq 1$. Andererseits ist $P^2 = P$, also $\|P\| = \|P^2\| \leq \|P\|^2$, was sofort $P = 0$ oder $\|P\| \geq 1$ impliziert. Zu guter Letzt ist $Px \perp x - Px$ und daher, nach dem Satz von Pythagoras,

$$\|x\|^2 = \|Px\|^2 + \|x - Px\|^2. \qquad \square$$

Insbesondere haben wir im Beweis des Satzes folgendes gezeigt:

Korollar 2.3.6. *Sei H ein Hilbertraum, H_0 ein abgeschlossener Unterraum und P die orthogonale Projektion auf H_0. Dann ist $z = Px$ genau dann, wenn $x - z \perp H_0$.*

Als Anwendung des Projektionssatzes zeigen wir die Existenz der bedingten Erwartung. Gegeben ist ein Wahrscheinlichkeitsraum $(\Omega, \Sigma, \mathbb{P})$ und eine Unter-σ-Algebra $\mathscr{F}$.

Definition 2.3.7. Ist $X \in L^2(\Omega)$ so heißt eine Zufallsvariable $Y \in L^2(\Omega)$ *bedingte Erwartung von X gegeben $\mathscr{F}$*, falls

(1) Y ist $\mathscr{F}$-messbar.

(2) Für $A \in \mathscr{F}$ ist $\int_A X\, d\mathbb{P} = \int_A Y\, d\mathbb{P}$.

Man schreibt $Y = \mathbb{E}(X|\mathscr{F})$.

Satz 2.3.8. *Es sei $(\Omega, \Sigma, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $\mathscr{F}$ eine Unter-σ-Algebra von Σ. Dann gibt es für jedes $X \in L^2(\Omega)$ eine bedingte Erwartung $\mathbb{E}(X|\mathscr{F})$ gegeben $\mathscr{F}$.*

Beweis. Wir betrachten den Hilbertraum $H = L^2(\Omega, \Sigma, \mathbb{P})$. Der Raum $H_0 := L^2(\Omega, \mathscr{F}, \mathbb{P})$ ist ein abgeschlossener Teilraum von H. Nach Satz 2.3.1 existiert die orthogonale Projektion P von H auf H_0, und wegen Satz 2.3.5 ist P linear. Es ist klar, dass PX $\mathscr{F}$-messbar ist, denn $PX \in H_0$. Weiter ist für $A \in \mathscr{F}$ ist $\mathbb{1}_A \in H_0$. Da also $X - PX \in H_0^\perp$ ist, ist $\langle \mathbb{1}_A, X - PX\rangle = 0$ für solche A. Dies zeigt, dass

$$\int_A X \, d\mathbb{P} = \int_A Y \, d\mathbb{P},$$

also ist PX eine bedingte Erwartung von X gegeben $\mathscr{F}$ ist. □

Nun beschreiben wir den Dualraum eines Hilbertraumes.

Ist $(H, \langle \cdot\, , \, \cdot \rangle)$ ein unitärer Raum und ist $y \in H$ so ist die Abbildung $\varphi_y : H \to \mathbb{K}$, definiert durch $\varphi_y(x) = \langle x, y\rangle$ linear. Es folgt aus der Cauchy–Schwarz Ungleichung $|\varphi_y(x)| \leq \|x\|\|y\|$, dass φ_y sogar beschränkt ist, also ist φ_y ein Element von H^*, dem Dualraum von H, ist. Es stellt sich die Frage, ob jedes Funktional auf H von dieser Form ist. Falls H vollständig ist (also H ein Hilbertraum ist) so ist dies in der Tat der Fall.

Satz 2.3.9. *(Riesz-Fréchet)*

Es sei H ein Hilbertraum und $\varphi \in H^*$. Dann gibt es genau ein $y \in H$ mit $\varphi(x) = \langle x, y\rangle$ für alle $x \in H$. Weiterhin ist $\|y\|_H = \|\varphi\|_{H^*}$.

Beweis. Da φ stetig ist, ist $\ker \varphi =: F$ ein abgeschlossener Teilraum von H. Nach Satz 2.3.5 ist $H = F \oplus_2 F^\perp$. Falls $F^\perp = \{0\}$ ist, ist $\varphi = 0$ und man kann $y = 0$ wählen. Ansonsten findet man ein $y_1 \in F^\perp$ mit $\varphi(y_1) = 1$. Für beliebiges $x \in H$ ist dann $x - \varphi(x)y_1 \in \ker\varphi = F$. Also ist $x - \varphi(x)y_1 \perp y_1$ und es folgt

$$0 = \langle x - \varphi(x)y_1, y_1\rangle = \langle x, y_1\rangle - \varphi(x)\|y_1\|^2.$$

Somit leistet $y := \|y_1\|^{-2}y_1$ das gewünschte.

Wäre z ein weiteres Element von H mit $\varphi(x) = \langle x, z\rangle$ für alle $x \in H$, so ist

$$\|y - z\|^2 = \langle y - z, z\rangle - \langle y - z, y\rangle = \varphi(y - z) - \varphi(y - z) = 0$$

und somit $y = z$. □

Bemerkung 2.3.10. Es folgt, dass die Abbildung $\Phi : H \to H^*$, gegeben durch $\Phi(y) = \varphi_y$ bijektiv, antilinear und isometrisch ist. Ist $\mathbb{K} = \mathbb{R}$, so ist H also isometrisch isomorph zu seinem Dualraum.

2.4 Selbstadjungierte Operatoren

Der Satz von Riesz-Fréchet erlaubt es uns, die Adjungierte eines Operators auf einem Hilbertraum zu definieren.

Lemma 2.4.1. *Es sei H ein Hilbertraum und $T \in \mathscr{L}(H)$. Dann gibt es genau einen Operator $T^* \in \mathscr{L}(H)$ derart, dass*

$$\langle Tx, y\rangle = \langle x, T^*y\rangle$$

für alle $x, y \in H$ gilt. Weiter ist $\|T^*\| = \|T\|$.

Beweis. Es sei $y \in H$. Die Abbildung $x \mapsto \langle Tx, y\rangle$ ist, als Verknüpfung der stetigen Abbildungen φ_y und H stetig. Nach dem Satz von Riesz-Fréchet gibt es ein eindeutiges Element $T^*y \in H$ sodass $\langle Tx, y\rangle = \varphi_y(Tx) = \langle x, T^*y\rangle$ für alle $x \in H$ gilt. Sind nun $y_1, y_2 \in H$ und $\lambda \in \mathbb{K}$, so ist

$$\langle x, T^*y_1 + \lambda T^*y_2\rangle = \langle x, T^*y_1\rangle + \bar{\lambda}\langle x, T^*y_2\rangle = \langle Tx, y_1\rangle + \bar{\lambda}\langle Tx, y_2\rangle = \langle Tx, y_1 + \lambda y_2\rangle.$$

Aus der Eindeutigkeitsaussage im Satz von Riesz-Frechet folgt nun $T^*(y_1 + \lambda y_2) = T^*y_1 + \lambda T^*y_2$, und somit ist T^* linear. Schliesslich ist

$$\begin{aligned}\|T^*\| &= \sup_{\|y\|\leq 1} \|T^*y\| = \sup_{\|y\|\leq 1\|} \sup_{\|x\|\leq 1} |\langle x, T^*y\rangle| \\ &= \sup_{\|x\|\leq 1} \sup_{\|y\|\leq 1\|} |\langle Tx, y\rangle| = \sup_{\|x\|\leq 1} \|Tx\| = \|T\|\end{aligned}$$

woraus einerseits die Beschränktheit von T^* andererseits aber auch die Identität $\|T\| = \|T^*\|$ folgt. □

Definition 2.4.2. Der Operator T^* aus Lemma 2.4.1 heißt *Adjungierte von* T. Ein Operator T heißt *selbstadjungiert*, falls $T = T^*$ ist.

Beispiel 2.4.3. Ist H_0 ein abgeschlossener Teilraum des Hilbertraum H und P die orthogonale Projektion auf H_0, so ist P selbstadjungiert. In der Tat ist $Px \in H_0$ und $x - Px \in H_0^\perp$ für jedes $x \in H$. Somit gilt für $x, y \in H$

$$\langle Px, y\rangle = \langle Px, Py\rangle + \langle Px, y - Py\rangle = \langle Px, Py\rangle = \langle Px, Py\rangle + \langle x - Px, Py\rangle = \langle x, Py\rangle$$

für alle $x, y \in H$.

Wir stellen einige Eigenschaften selbstadjungierter Operatoren zusammen.

Proposition 2.4.4. *Es sei H ein Hilbertraum und $T \in \mathscr{L}(H)$.*

(1) Ist T selbstadjungiert, so ist $\langle Tx, x\rangle \in \mathbb{R}$ für alle $x \in H$. Falls $\mathbb{K} = \mathbb{C}$ ist, so gilt auch die Umkehrung: Ist $\langle Tx, x\rangle \in \mathbb{R}$ für alle $x \in H$, so ist T selbstadjungiert.

(2) ist T selbstadjungiert, so ist $\|T\| = \sup_{\|x\|\leq 1} |\langle Tx, x\rangle|$.

Beweis. (1) Ist T selbstadjungiert, so ist

$$\langle Tx, x\rangle = \langle x, Tx\rangle = \overline{\langle Tx, x\rangle}$$

für alle $x \in H$.

Sei nun $\mathbb{K} = \mathbb{C}$ und T ein beschränkter Operator mit $\langle Tz, z\rangle \in \mathbb{R}$ für alle $z \in H$. Für $x, y \in H$ und $\lambda \in \mathbb{C}$ gilt dann

$$\langle T(x + \lambda y), (x + \lambda y)\rangle = \langle Tx, x\rangle + \bar{\lambda}\langle Tx, y\rangle + \lambda\langle Ty, x\rangle + |\lambda|^2\langle Ty, y\rangle.$$

Durch komplexes konjugieren dieser Gleichung folgt unter Beachtung von $\overline{\langle Tz, z\rangle} = \langle Tz, z\rangle$ und $\overline{\langle Tu, v\rangle} = \langle v, Tu\rangle$ die Gleichung

$$\langle T(x + \lambda y), (x + \lambda y)\rangle = \langle Tx, x\rangle + \lambda\langle y, Tx\rangle + \bar{\lambda}\langle x, Ty\rangle + |\lambda|^2\langle Ty, y\rangle.$$

Subtrahiert man diese Gleichungen voneinander und setzt $\lambda = 1$ resp. $\lambda = -i$ ein, so herhält man

$$\begin{aligned} \langle Tx, y\rangle + \langle Ty, x\rangle &= \langle y, Tx\rangle + \langle x, Ty\rangle \\ i\langle Tx, y\rangle - i\langle Ty, x\rangle &= -i\langle y, Tx\rangle + i\langle x, Ty\rangle. \end{aligned}$$

und daraus $\langle Tx, y\rangle = \langle x, Ty\rangle$ wie gewünscht.

(2) Es ist klar, dass $\|T\| \geq \sup_{\|x\|\leq 1} |\langle Tx, x\rangle| =: C$ gilt. Da T selbstadjungiert is, ist

$$\begin{aligned} \langle T(x+y), x+y\rangle - \langle T(x-y), x-y\rangle &= 2\langle Tx, y\rangle + 2\langle Ty, x\rangle \\ &= 2\langle Tx, y\rangle + 2\overline{\langle Tx, y\rangle} = 4\mathrm{Re}\,\langle Tx, y\rangle \end{aligned}$$

für $x, y \in H$. Mit der Parallelogrammgleichung (2.1) folgt, dass

$$4\mathrm{Re}\,\langle Tx, y\rangle \leq C(\|x+y\|^2 + \|x-y\|^2) = 2C(\|x\|^2 + \|y\|^2).$$

Somit gilt für x, y mit $\|x\| = \|y\| = 1$

$$4\mathrm{Re}\,\langle Tx, y\rangle \leq 4C.$$

Ersetzt man x durch λx für ein geeignetes λ mit $|\lambda| = 1$, so folgt $|\langle Tx, y\rangle| \leq C$ für $\|x\| = \|y\| = 1$ und somit die Behauptung. □

Bemerkung 2.4.5. Proposition 2.4.4 zeigt, wie wichtig die Menge

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

im Falle selbstadjungierter Operatoren T ist. Auch für allgemeine Operatoren auf Hilberträumen ist diese Menge von Interesse und wird *numerischer Wertebereich* genannt.

Kapitel 3

Anwendungen

Wir wenden nun unsere Resultate über Hilberträume auf Anwendungsprobleme an. Genauer studieren wir eindimensionale Varianten von sogenannten elliptischen, partiellen Differentialgleichungen. Zunächst führen wir die relevanten Hilberträume – die Sobolevräume – ein. Dies sind Teilräume von L^2, die aus Funktionen bestehen, die in gewissem Sinne eine schwache Ableitung besitzen. Es ist diese Konstruktion, die es uns erlaubt Hilbertraummethoden zu verwenden, denn die Räume C^k von k-mal stetig differenzierbaren Funktionen sind *keine* Hilberträume. Sodann kann man die Existenz von Lösungen zu gewissen partiellen Differentialgleichungen sehr einfach aus dem Satz von Riesz-Fréchet folgern. Die dargestellten Methoden funktionieren genauso auch in höherer Dimension, wie man in der Vorlesung "partielle Differentialgleichungen" sehen wird.

3.1 Sobolevräume

Es sei (a,b) ein offenes Intervall. Mit $C_c^1(a,b)$ bezeichnen wir den Raum der stetig differenzierbaren Funktionen mit kompaktem Träger, d.h. derer stetig differenzierbaren Funktionen f, für die es ein $\varepsilon > 0$ gibt (welches durchaus von f abhängen darf) sodass $f(t) = 0$ für $t \notin [a+\varepsilon, b-\varepsilon]$. Beachte, dass $C_c^1(a,b)$ dicht in $L^2(a,b)$ liegt, siehe Bemerkung 2.2.7.

Definition 3.1.1. Es sei $f \in L^2(a,b)$. Eine Funktion $g \in L^2(a,b)$ heißt *schwache Ableitung* von f, falls

$$\int_a^b f(x)\varphi'(x)\,dx = -\int_a^b g(x)\varphi(x)\,dx,$$

also $\langle f, \varphi' \rangle = -\langle g, \varphi \rangle$, für alle $\varphi \in C_c^1(a,b)$. Wenn f eine schwache Ableitung besitzt, so nennen wir f *schwach differenzierbar*.

Beispiel 3.1.2. Ist (a,b) ein beschränktes Intervall und $f \in C^1[a,b]$, so ist f' eine schwache Ableitung von f. In der Tat gilt mit partieller Integration

$$\int_a^b f(x)\varphi'(x)\,dx = [f(x)\varphi(x)|_a^b - \int_a^b f'(x)\varphi(x)\,dx = -\int_a^b f'(x)\varphi(x)\,dx,$$

da ja $\varphi(a) = \varphi(b) = 0$ ist.

A priori ist nicht klar, ob eine schwach differenzierbare Funktion nicht auch mehrere schwache Ableitungen besitzen kann. Dies ist allerdings nicht der Fall.

Lemma 3.1.3. *Ist $f \in L^2(a,b)$ schwach differenzierbar, und sind g_1, g_2 schwache Ableitungen von f, so ist $g_1 = g_2$.*

Beweis. Für $\varphi \in C_c^1(a,b)$ ist $\langle g_1 - g_2, \varphi \rangle = -\langle f, \varphi' \rangle + \langle f, \varphi' \rangle = 0$. Da $C_c^1(a,b)$ dicht in $L^2(a,b)$ liegt, folgt $g_1 = g_2$. □

Definition 3.1.4. Der *Sobolevraum* $H^1(a,b)$ ist der Vektorraum (!) der schwach differenzierbaren Funktionen $f \in L^2(a,b)$. Die schwache Ableitung von f wird mit f' bezeichnet.

Satz 3.1.5. *$H^1(a,b)$ ist ein Hilbertraum bezüglich des Skalarprodukts*

$$\langle f, g \rangle_{H^1(a,b)} = \langle f, g \rangle_{L^2(a,b)} + \langle f', g' \rangle_{L^2(a,b)}.$$

Beweis. Das kartesische Produkt $H := L^2(a,b) \times L^2(a,b)$ ist ein Hilbertraum bezüglich des Skalarprodukts

$$\langle (f_1, f_2), (g_1, g_2) \rangle_H := \langle f_1, g_1 \rangle_{L^2(a,b)} + \langle f_2, g_2 \rangle_{L^2(a,b)},$$

wie leicht zu sehen ist. Offensichtlich ist die Abbildung $\Phi : H^1(a,b) \to H$, gegeben durch $\Phi(f) = (f, f')$ isometrisch. Wir zeigen, dass das Bild von Φ abgeschlossen, also nach Lemma 1.1.8 selbst ein Hilbertraum, ist. Dann folgt die Vollständigkeit von $H^1(a,b)$.

Sei also (f_n, f_n') eine Folge im Bild von Φ, sodass $f_n \to f$ in $L^2(a,b)$ und $f_n' \to g$ in $L^2(a,b)$. Wir haben zu zeigen, dass f schwach differenzierbar ist mit $f' = g$. Sei hierzu $\varphi \in C_c^1(a,b)$. Wegen der Stetigkeit des Skalarproduktes in $L^2(a,b)$, folgt

$$\langle f, \varphi' \rangle = \lim_{n \to \infty} \langle f_n, \varphi' \rangle = \lim_{n \to \infty} -\langle f_n', \varphi \rangle = -\langle g, \varphi \rangle.$$

Dies zeigt, dass f schwach differenzierbar ist mit schwacher Ableitung g. □

Bemerkung 3.1.6. Der Beweis von Satz 3.1.5 zeigt insbesondere, dass der Operator $D : L^2(a,b) \to L^2(a,b)$, gegeben durch $\mathrm{def}(D) = H^1(a,b)$, $Df = f'$ abgeschlossen ist.

Es stellt sich die Frage, inwieweit die schwache Ableitungen die gleichen (oder zumindest ähnliche) Eigenschaften wie die gewöhnliche Ableitung hat. Wir haben

Lemma 3.1.7. *Ist $f \in H^1(a,b)$ mit $f' = 0$, so ist f konstant.*

Beweis. Wähle eine Funktion $u \in C_c^1((a,b)$ mit $\int_a^b u = 1$. Ist dann $\varphi \in C_c^1(a,b)$ so ist

$$\psi(t) = \varphi(t) - u(t) \int_a^b \varphi(s)\, ds$$

ein Element von $C_c^1(a,b)$ mit $\int_a^b \psi(t)\, dt = 0$. Genauer gilt sogar $\int_a^x \psi(t)\, dt = 0$ für x in einer Umgebung von b. Somit ist auch $\xi(x) := \int_a^x \psi(t)\, dt$ ein Element von $C_c^1(a,b)$. Ist also $f' = 0$, so folgt

$$\begin{aligned} 0 = \langle f', \xi \rangle = -\langle f, \psi \rangle &= -\int_a^b f(x) \Big[\varphi(x) - u(x) \int_a^b \varphi(t)\, dt \Big]\, dx \\ &= -\int_a^b f(x)\varphi(x)\, dx + \Big(\int_a^b \varphi(t)\, dt \Big) \Big(\int_a^b f(x) u(x)\, dx \Big) \\ &= \int_a^b (c - f(x))\varphi(x)\, dx, \end{aligned}$$

wobei wir $c := \int_a^b f(x) u(x)\, dx$ geschrieben haben. Weil $\varphi \in C_c^1(a,b)$ beliebig war, folgt aus der Dichtheit dieser Funktionen, dass $c - f = 0$ ist, also $f = c$. □

Wir beschreiben nun eine Klasse von Beispielen von Elementen in $H^1(a,b)$.

Lemma 3.1.8. *Es sei $g \in L^2(a,b)$ und $c \in [a,b]$. Dann ist die Funktion f, definiert durch*

$$f(t) := \int_c^t g(s)\,ds,$$

schwach differenzierbar und $f' = g$.

Beweis. Es sei $\varphi \in C_c^1(a,b)$. Mit dem Satz von Fubini folgt

$$\begin{aligned}
\int_a^b f(x)\varphi'(x)\,dx &= \int_a^b \varphi'(x) \int_c^x g(r)\,dr\,dx \\
&= -\int_a^c \int_x^c \varphi'(x)g(r)\,drdx + \int_c^b \int_c^x \varphi'(x)g(r)\,drdx \\
&= -\int_a^c \int_a^r \varphi'(x)g(r)\,drdx + \int_c^b g(r) \int_r^b \varphi'(x)\,dx\,dr \\
&= -\int_a^c \varphi(r)g(r)\,dr - \int_c^b g(r)\varphi(r)\,dr = -\int_a^b g(r)\varphi(r)\,dr
\end{aligned}$$ □

Satz 3.1.9. *Gegeben $f \in H^1(a,b)$ gibt es eine stetige Funktion $\tilde{f} : \overline{(a,b)} \to \mathbb{K}$ mit $f = \tilde{f}$ fast überall. Ferner ist*

$$\tilde{f}(t) - \tilde{f}(s) = \int_s^t f'(r)\,dr$$

Beweis. Definiere $g(t) := \int_a^t f'(s)\,ds$ für $t \in (a,b)$. Mit dominerter Konvergenz sieht man leicht, dass g eine stetige Funktion auf $\overline{(a,b)}$ ist. Wegen Lemma 3.1.8 ist $g \in H^1(a,b)$ mit $g' = f'$ und somit gibt es, wegen Lemma 3.1.7, eine Konstante C sodass $f = g + C =: \tilde{f}$ fast überall. Nach Definition gilt

$$\tilde{f}(t) - \tilde{f}(s) = \int_s^t f'(r)\,dr.$$ □

Bemerkung 3.1.10. Eine stetige Funktion $\tilde{f}$ mit $\tilde{f} = f$ fast überall nennt man einen *stetigen Repäsentanten* von f. Genauer ist ja f eigentlich eine Äquivalenzklasse $[f]$ und $f = \tilde{f}$ fast überall bedeutet gerade, dass $\tilde{f} \in [f]$.

Beachte, dass eine Funktion in $L^2(a,b)$ (bezüglich des Lebesguemaßes!) höchstens einen stetigen Repräsentanten besitzt, denn wenn zwei stetige Funktionen verschieden sind, so sind sie auf einer offenen Menge, also einer Menge positiven Maßes verschieden. Daher unterscheidet man häufig nicht zwischen einer Funktion in $H^1(a,b)$ und ihrem stetigen Repräsentanten und sagt eine Funktion $f \in H^1(a,b)$ *ist* stetig.

Korollar 3.1.11. *Für $a,b \in \mathbb{R}$ mit $a < b$ ist $C^1([a,b])$ dicht in $H^1(a,b)$.*

Beweis. Es sei $f \in H^1(a,b)$. Weil $C([a,b])$ dicht in $L^2(a,b)$ ist, gibt es eine Folge $g_n \in C([a,b])$ mit $g_n \to f'$ in $L^2(a,b)$. Es sei $f_n(x) := \tilde{f}(a) + \int_a^x g_n'(s)\,ds$. Dann ist $f_n \in C^1([a,b])$. Wir behaupten, dass $f_n \to f$ in $H^1(a,b)$. Nach Konstruktion ist $f_n' \to f'$ in $L^2(a,b)$, es bleibt also zu zeigen, dass $f_n \to f$ in $L^2(a,b)$.

Für $x \in (a,b)$ ist aber

$$|f_n(x) - f(x)| = \left| \int_a^x g_n(t) - f'(t)\,dt \right| \leq (b-a)^{\frac{1}{2}} \|g_n - f'\|_2,$$

sodass $\|f_n - f\|_2 \leq (b-a)^{\frac{1}{2}} \|g_n - f'\|_2 \to 0.$ □

Wir zeigen nun noch, dass für $-\infty < a < b < \infty$ die Abbildung $\iota : H^1(a,b) \to C([a,b])$, gegeben durch $\iota(f) = \tilde{f}$ stetig ist. Man sagt $H^1(a,b)$ ist *stetig in* $C([a,b])$ *eingebettet* und schreibt $H^1(a,b) \hookrightarrow C([a,b])$.

Satz 3.1.12. *Es seien $a, b \in \mathbb{R}$, mit $a < b$. Dann gibt es eine Konstante C so, dass*

$$\|\tilde{f}\|_\infty \leq C\|f\|_{H^1}.$$

Beweis. Wenn die Aussage falsch wäre, so gäbe es eine Folge $f_n \in H^1(a,b)$ mit $\|f_n\|_{H^1} \leq 1$ und gleichzeitig $\|\tilde{f}_n\|_\infty \geq n$. Wegen der Cauchy–Schwarz Ungleichung ist

$$\Big| \int_a^x f_n'(s)\, ds \Big| \leq \|\mathbb{1}_{(a,x)}\|_2 \|f_n'\|_2 \leq (b-a)^{\frac{1}{2}} < \infty.$$

Weil nach Satz 3.1.9 $\tilde{f}(t) = \tilde{f}(a) + \int_a^t f'(s)\, ds$, muss also $\tilde{f}_n(a)$ unbeschränkt sein. Andererseits ist aber $|\tilde{f}_n(x)| \geq |\tilde{f}_n(a)| - (b-a)^{\frac{1}{2}}$ für jedes $x \in (a,b)$ und daher müsste $\|f_n\|_2$ unbeschränkt sein, ein Widerspruch zur Beschränktheit in $H^1(a,b)$. □

Die Regel für die partielle Integration stand bei der Definition der schwachen Ableitung Pate. Dort hatten wir jedoch nur “Testfunktionen” mit kompaktem Träger zugelassen, sodass die Randterme beim partiellen Integrieren verschwinden. Dies war notwendig, um nicht eine L^2-Funktion in einem Punkt auswerten zu müssen. Da H^1-Funktionen jedoch stetig sind kann man sie in Punkten auswerten und es stellt sich die Frage, ob man partiell integrieren kann.

Proposition 3.1.13. *Es seien $a, b \in \mathbb{R}$ mit $a < b$ und $f, g \in H^1(a,b)$.*

(1) $fg \in H^1(a,b)$ und $(fg)' = fg' + f'g$.

(2) Es gilt die Regel der partiellen Integration

$$\int_a^b f(t)g'(t)\, dt = [fg]_a^b - \int_a^b f'(t)g(t)\, dt.$$

Beweis. (ÜA) □

Wir definieren nun noch einige Unterräume von $H^1(a,b)$.

Definition 3.1.14. Es seien $a, b \in \mathbb{R}$ mit $a < b$. Dann sei

$$H_0^1(a,b) := \{f \in H^1(a,b) : f(a) = f(b) = 0\}$$

und

$$H_{\text{per}}^1(a,b) := \{f \in H^1(a,b) : f(a) = f(b)\}.$$

Man beachte, dass $H_0^1(a,b)$ und $H_{\text{per}}^1(a,b)$ wegen der Stetigkeit der Einbettung von $H^1(a,b) \hookrightarrow C([a,b])$ abgeschlossene Teilräume von $H^1(a,b)$ sind und daher selbst Hilbertr̈aume sind.

Wir zeigen nun eine wichtige Ungleichung, die sogenannte *Poincaré-Ungleichung*. Aus dieser Ungleichung folgt, dass $\langle f, g\rangle_0 := \langle f', g'\rangle_{L^2}$ ein äquivalentes Skalarprodukt auf $H_0^1(a,b)$ definiert, was wichtige Konsequenzen beim Lösen von Randwertproblemen hat. Hierbei heißen zwei Skalarprodukte $\langle\,\cdot\,,\,\cdot\,\rangle_1$ und $\langle\,\cdot\,,\,\cdot\,\rangle_2$ *äquivalent*, wenn die assoziierten Normen äquivalent sind, d.h. wenn es Konstanten $c_1, c_2 > 0$ gibt mit $c_1\langle x, x\rangle_1 \leq \langle x, x\rangle_2 \leq c_2\langle x, x\rangle_2$ für alle $x \in H$.

Satz 3.1.15. *(Poincaré-Ungleichung)*
Es seien $a, b \in \mathbb{R}$ mit $a < b$. Dann gilt für $f \in H_0^1(a,b)$ stets

$$\int_a^b |f(x)|^2 \, dx \leq \frac{1}{2}(b-a)^2 \int_a^b |f'(x)|^2 \, dx.$$

Beweis. Nach Definition von $H_0^1(a,b)$ und wegen Satz 3.1.9 ist $f(x) = \int_a^x f'(s)\,ds$. Mit Cauchy–Schwarz folgt

$$|f(x)| \leq (x-a)^{\frac{1}{2}} \Big(\int_a^x |f'(s)|^2 \, ds \Big)^{\frac{1}{2}} \leq (x-a)^{\frac{1}{2}} \|f'\|_2.$$

Somit

$$\int_a^b |f(x)|^2 \, dx = \|f'\|_2^2 \int_a^b (x-a)\, dx = \frac{1}{2}(b-a)^2 \|f'\|_2^2.$$

□

Wir können auch höhere schwache Ableitungen betrachten.

Definition 3.1.16. Für $n \in \mathbb{N}$ definieren wir induktiv

$$H^{n+1}(a,b) := \{ f \in H^1(a,b) : f' \in H^n(a,b) \}$$

Weiter sei für $f \in H^{n+1}(a,b)$ die Funktion $f^{(n+1)}$ definiert als $(f^{(n)})'$. Die Elemente von $H^n(a,b)$ heißen *n-mal schwach differenzierbare Funktionen* und $f^{(n)}$ heißt die n-te schwache Ableitung.

Ähnlich wie im Fall $n = 1$ sieht man, dass $H^n(a,b)$ ein Hilbertraum bezüglich des Skalarprodukts

$$\langle f, g \rangle_{H^n} = \langle f, g \rangle_{L^2} + \sum_{k=1}^n \big\langle f^{(k)}, g^{(k)} \big\rangle_{L^2}$$

ist. Außerdem sieht man induktiv, dass $H^n(a,b) \hookrightarrow C^{n-1}([a,b])$, d.h. eine n-mal schwach differenzierbare Funktion ist $(n-1)$-mal im klassischen Sinn differenzierbar.

3.2 Randwertprobleme auf dem Intervall

Wir verwenden nun Hilbertraummethoden um Randwertprobleme der Form

$$\lambda u - u'' = f \tag{3.1}$$

zu lösen. Hierbei ist $f \in L^2(a,b)$ und es ist eine zweimal differenzierbare Funktion u gesucht, die diese Gleichung löst. Der Einfachheit halber arbeiten wir in disem Abschnitt stets über dem Skalarkörper $\mathbb{R}$.

Bei (3.1) handelt es sich um eine (lineare) Differentialgleichung zweiter Ordnung. Es ist aus der Vorlesung "Differentialgleichungen" bekannt, dass die Lösungsgesamtheit von (3.1) (sofern nicht leer) ein zweidimensionaler, affiner Raum ist. Um also die Eindeutigkeit der Lösung sicher zu stellen, muss man zusätzliche Forderungen stellen. In der Vorlesung "Differentialgleichungen" hat man meistens sogenannte Anfangswertprobleme betrachtet, bei denen man zum Startzeitpunkt a die Werte von u und u' vorschreibt.

Bei sogenannten *Randwertproblemen* schreibt man stattdessen Werte an *beiden* Rändern a und b vor. Wichtige Randbedingungen sind *Dirichlet Randbedingungen*

$$u(a) = u(b) = 0, \tag{3.2}$$

Neumann Randbedingungen

$$u'(a) = u'(b) = 0, \tag{3.3}$$

und *periodische Randbedingungen*

$$u(a) = u(b) \quad \text{und} \quad u'(a) = u'(b). \tag{3.4}$$

Die so entstehenden Randwertprobleme sind nicht etwa eine allgemeinere Form einer gewöhnlichen Differentialgleichung, sondern sonder eher eine eindimensionale Version der *partiellen Differentialgleichung*

$$\lambda u - \Delta u = f$$

wobei u eine Funktion von d Variablen ist und $\Delta = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_d^2}$ der Laplace Operator ist.

Die Technik die wir verwenden um (3.1) zu lösen ist anders als die in der Vorlesung "Differentialgleichungen" verwendeten Methoden. Zunächst verwenden wir einen anderen Lösungsbegriff. Wir suchen (zumindest a priori) keine *klassischen Lösungen* von (3.1) (d.h. Lösungen die C^2-funktionen sind) sondern *schwache Lösungen*, d.h. Lösungen die H^2-sind. Dies erlaubt es uns, Hilbertraummethoden zu verwenden.

Betrachten wir zunächst Gleichung (3.1) mit Dirichlet-Randbedingungen (3.2). Sei $u \in H^2(a,b)$ Lösung von (3.1). Wir multiplizieren u mit einer $H_0^1(a,b)$-Funktion v und integieren partiell und erhalten

$$\begin{aligned} \int_a^b f(t)v(t)\,dt &= \lambda \int_a^b u(t)v(t)\,dt - \int_a^b u''(t)v(t)\,dt \\ &= \lambda \int_a^b u(t)v(t)\,dt + \int_a^b u'(t)v'(t)\,dt - [u'v|_a^b \\ &= \lambda \int_a^b u(t)v(t)\,dt + \int_a^b u'(t)v'(t)\,dt, \end{aligned} \tag{3.5}$$

da ja $v(a) = v(b) = 0$ ist. Jede Lösung von (3.1) erfüllt also (3.5). Dies ist die sogenannte *variationelle Formulierung des Problems.*

Beachte, dass auf der rechten Seite von (3.5) fast das Skalarprodukt in $H_0^1(a,b)$ steht, genauer *ist* dies für $\lambda = 1$ das Skalarprodukt auf $H_0^1(a,b)$, für $\lambda \geq 0$ ist es ein äquivalentes Skalarprodukt.

Die linke Seite von (3.5) ist (als Funktion von v) ein beschränktes Funktional auf $H_0^1(a,b)$. Dies ermöglicht es uns, eine Lösung u von (3.5) aus dem Satz von Riesz-Fréchet zu erhalten. Genauer zeigen wir:

Satz 3.2.1. *Es seien $a, b \in \mathbb{R}$ mit $a < b$ und $\lambda \geq 0$. Für jede Funktion $f \in L^2(a,b)$ gibt es genau eine Funktion $u \in H^2(a,b) \cap H_0^1(a,b)$ mit $\lambda u - u'' = f$.*

Bemerkung 3.2.2. Beachte, dass die Dirichlet Randbedingungen erfüllt sind, da wir verlangen, dass u zusätzlich in H_0^1 liegt.

Beweis. Wir betrachten den Hilbertraum $V = H_0^1(a,b)$ mit dem Skalarprodukt $\langle u,v\rangle := \lambda\langle u,v\rangle_{L^2(a,b)} + \langle u',v'\rangle_{L^2(a,b)}$. Dieses Skalarprodukt ist äquivalent zum üblichen Skalarprodukt auf $H_0^1(a,b)$. Für $\lambda = 0$ folgt dies aus der Poincaré Ungleichung (Satz 3.1.15). Für $\lambda > 0$ ist dies klar.

Ist nun $f \in L^2(a,b)$, so ist die Abbildung $\varphi : V \to \mathbb{R}$, $v \mapsto \int_a^b \varphi(x)v(x)\,dx$ stetig. Es ist nämlich $|\varphi(v)| \le \|f\|_2\|v\|_2 \le \|f\|_2\|v\|_{H_0^1} \le c\|f\|_2\|v\|_V$. Nach dem Satz 2.3.9 von Riesz-Fréchet gibt es genau eine Funktion $u \in H_0^1(a,b)$ mit $\langle u,v\rangle = \varphi(v)$ für alle $v \in H_0^1(a,b)$, d.h. (3.5) gilt. Beachte insbesondere, dass für $v = \psi \in C_c^1(a,b)$

$$\int_a^b u'(x)\psi'(x)\,dx = -\int_a^b (\lambda u(x) - f(x))\psi(x)\,dx$$

gilt. Aus der Definition folgt, dass u' schwach differenzierbar ist mit schwacher Ableitung $\lambda u - f$. Somit ist $u \in H^2(a,b)$ und $u'' = \lambda u - f$. Dies zeigt, dass u die Gleichung (3.1) erfüllt. □

Wir zeigen nun das analoge Resultat für Neumann Randbedingungen. Beachte, dass wir in diesem Fall $\lambda > 0$ fordern müssen. In der Tat sind für $\lambda = 0$ sowohl $u \equiv 0$, als auch $u \equiv 1$ Lösungen der Gleichung (3.1) mit $f = 0$ und Neumann Randbedingungen.

Satz 3.2.3. *Seien $a, b \in \mathbb{R}$ mit $a < b$ und $\lambda > 0$. Dann gibt es für jedes $f \in L^2(a,b)$ genau eine Funktion u in*

$$\{v \in H^2(a,b) : v'(a) = v'(b) = 0\}$$

mit $\lambda u - u'' = f$.

Beweis. Für $\lambda > 0$ definiert $\langle u,v\rangle := \lambda\langle u,v\rangle_{L^2} + \langle u',v'\rangle_{L^2(a,b)}$ ein äquivalentes Skalarprodukt auf $H^1(a,b)$. Da $v \mapsto \int_a^b f(x)v(x)\,dx$ ein lineares Funktional auf $H^1(a,b)$ definiert, gibt es nach dem Satz von Riesz-Fréchet genau ein $u \in H^1$ mit $\langle u,v\rangle = \int_a^b fv$ für alle $v \in H^1(a,b)$. Insbesondere hat man für $v = \psi \in C_c^1(a,b)$

$$\int_a^b u'(x)\psi'(x)\,dx = -\int_a^b (\lambda u(x) - f(x))\psi(x)\,dx,$$

sodass $u \in H^2(a,b)$ mit $u'' = \lambda u - f$ ist. Es bleibt also nur noch zu zeigen, dass u die Neumann Randbedingungen $u'(a) = u'(b) = 0$ erfüllt. Betrachte hierzu $v(t) := (b-a)^{-1}(t-a)$. Mittels partieller Integration folgt

$$-\int_a^b u''v = -\int_a b(\lambda u - f)v = \int_a^b u'v' = [u'v|_a^b - \int_a^b u''v = u'(b) - \int_a^b u''v$$

und daher $u'(b) = 0$. Verwendet man stattdessen $v(t) := (b-a)^{-1}(b-t)$, so sieht man auf die gleiche Weise, dass $u'(a) = 0$ ist. □

Kapitel 4

Spektraltheorie

Der Satz über die Hauptachsentransformation aus der linearen Algebra besagt, dass eine symmetrische (bzw. hermitsche) Matrix unitär diagonalisierbar ist. Also hat der Vektorraum eine Orthogonalbasis aus Eigenvektoren der Matrix. In diesem Kapitel beweisen wir eine unendlichdimensionale Version dieses Satzes. Den Platz der symmetrischen/hermitschen Matrix nimmt nun ein selbstadjungierter Operator ein. Um allerdings die Existenz von Eigenwerten zu sichern, muss man sich auf eine besondere Klasse von Operatoren, die sogenannten *kompakten Operatoren* beschränken.

4.1 Spektrum und Resolvente

Definition 4.1.1. Es sei X ein normierter Raum, $T \in \mathscr{L}(X)$. Die *Resolventenmenge* von T ist die Menge

$$\rho(T) := \{\lambda \in \mathbb{K} : (\lambda - T) \text{ ist bijektiv und } (\lambda - T)^{-1} \in \mathscr{L}(X)\}.$$

Das Komplement $\sigma(T) := \mathbb{K} \setminus \rho(T)$ der Resolventenmenge heißt *Spektrum von* T. Für $\lambda \in \rho(T)$ heißt $R(\lambda, T) := (\lambda - T)^{-1}$ die *Resolvente in* λ, die Abbildung $\rho(T) \ni \lambda \mapsto R(\lambda, T)$ heißt *Resolventenabbildung*. Ist $\lambda \in \mathbb{K}$ so, dass es einen Vektor $x \neq 0$ gibt mit $Tx = \lambda x$, so heißt λ Eigenwert von T und x heißt *zugehöriger Eigenvektor*. Wir schreiben $\sigma_p(T)$ für die Menge aller Eigenwerte von T und nennen $\sigma_p(T)$ das *Punktspektrum von* T.

Bemerkung 4.1.2. Ist $S \in \mathscr{L}(X)$ bijektiv mit $S^{-1} \in \mathscr{L}(X)$, so sagen wir S ist *invertierbar*. Unter Verwendung des sogenannten *Satz vom abgeschlossenen Graphen* kann man zeigen, dass auf einem Banachraum (!) jeder bijektive Operator invertierbar ist.

Ist X endlichdimensional, so ist ein linearer Operator injektiv, genau dann, wenn er surjektiv ist, genau dann wenn er bijiektiv ist. In diesem Fall ist also $\lambda \in \rho(T)$ genau dann, wenn λ ein Eigenwert von T ist. Auf unendlichdimensionalen Räumen ist dies nicht notwendigerweise der Fall.

Beispiel 4.1.3. Es sei $H = \ell^2$ und R der *Rechtsshift*, gegeben durch

$$R(x_1, x_2, \ldots) = (0, x_1, x_2, \ldots).$$

Dann ist R injektiv, aber nicht surjektiv. Also ist $0 \in \sigma(R)$, aber 0 ist kein Eigenwert von R.

Wir zeigen nun, dass die Resolventenmenge offen ist. Hierzu verwenden wir die folgende operatorwertige Version der geometrischen Reihe.

Lemma 4.1.4. *(Neumann Reihe)*

Es sei X ein Banachraum und $T \in \mathscr{L}(X)$ mit $\|T\| < 1$. Dann ist $I - T$ invertierbar und $(I-T)^{-1} = \sum_{k=0}^\infty T^k$.

Beweis. Weil X vollständig ist, ist wegen Proposition 1.2.4 auch $\mathscr{L}(X)$ vollständig. Weil $\|T\| =: q < 1$, also $\|T^k\| \leq \|T\|^k \leq q^k$, ist $\sum_{k=0}^\infty \|T^k\| < \infty$. Somit ist die Reihe $\sum_{k=1}^n T^k$ absolutkonvergent, wegen Vollständigkeit also konvergent. Sei S der Grenzwert der Reihe. Es ist

$$(I-T)S = \lim_{n\to\infty}(I-T)\sum_{k=0}^n T^k = \lim_{n\to\infty} I - T^{k+1} = I$$

da $T^{k+1} \to 0$. Genauso sieht man, dass $S(I-T) = I$, es ist also $(I-T)$ invertierbar und $S = (I-T)^{-1}$. □

Satz 4.1.5. *Es sei X ein Banachraum und $T \in \mathscr{L}(X)$. Dann ist $\sigma(T)$ kompakt, genauer ist $|\lambda| \leq \|T\|$ für $\lambda \in \sigma(T)$.*

Beweis. Ist $|\lambda| > \|T\|$, so ist $\|\lambda^{-1}T\| < 1$, also ist wegen Lemma 4.1.4 $I - \lambda^{-1}T$ invertierbar mit Inverse $\sum_{k=0}^\infty \lambda^{-n}T^n$. Weil $(\lambda - T) = \lambda(I - \lambda^{-1}T)$ ist, ist auch $\lambda - T$ invertierbar (mit $(\lambda-T)^{-1} = \lambda^{-1}\sum_{k=0}^\infty \lambda^{-n}T^n$). Somit ist $|\lambda| \leq \|T\|$ für alle $\lambda \in \sigma(T)$. Insbesondere ist das Spektrum beschränkt.

Es bleibt zu zeigen, dass das Spektrum abgeschlossen ist, oder, äquivalent, dass die Resolventenmenge abgeschlossen ist. Sei hierzu $\lambda_0 \in \rho(T)$ und $\varepsilon := \|R(\lambda_0, T)\|^{-1} > 0$. Es ist

$$\lambda - T = \lambda_0 - T - (\lambda_0 - \lambda) = (\lambda_0 - T)\Big[I - (\lambda_0-\lambda)(\lambda_0 - T)^{-1}\Big] =: (\lambda_0 - T)(I - R)$$

Weil $(\lambda_0 - T)$ invertierbar ist, ist $\lambda - T$ also genau dann invertierbar, wenn $I - R$ invertierbar ist. Ist $|\lambda - \lambda_0| < \varepsilon$, so ist $\|R\| < 1$ und daher $I - R$ invertierbar nach Lemma 4.1.4. Somit ist $B(\lambda_0, \varepsilon) \subset \rho(T)$, die Resolventenmenge also offen. □

Bemerkung 4.1.6. Mit mehr Informationen über den Dualraum von $\mathscr{L}(X)$ und etwas Funktionentheorie kann man beweisen, dass im Fall $\mathbb{K} = \mathbb{C}$ stets $\sigma(T) \neq \emptyset$. Im Falle dass $\mathbb{K} = \mathbb{R}$ ist kann jedoch das Spektrum durchaus leer sein, wie man bereits an der Matrix

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

die Minimalpolynom $1 + X^2$ hat, sehen kann.

Wir beschreiben nun das Spektrum des adjungierten Operators im Hilbertraum.

Lemma 4.1.7. *Ist H ein Hilbertraum und $T \in \mathscr{L}(H)$, so ist $\sigma(T^*) = \overline{\sigma(T)}$. Weiterhin ist für $\lambda \in \rho(T)$ stets $R(\bar\lambda, T^*) = R(\lambda, T)^*$.*

Beweis. Für $S, T \in \mathscr{L}(H)$ ist $\langle TSx, y\rangle = \langle Sx, T^*y\rangle = \langle x, S^*T^*y\rangle$ für alle $x, y \in H$. Somit ist $(TS)^* = S^*T^*$.

Ist nun $\lambda \in \rho(T)$, so ist $I = (\lambda - T)R(\lambda, T) = R(\lambda, T)(\lambda - T)$ und daher

$$I = I^* = \big[(\lambda - R)R(\lambda, T)\big]^* = R(\lambda, T)^*(\bar\lambda - T^*) = (\bar\lambda - T^*)R(\lambda, T)^*.$$

Dies zeigt, dass $\bar\lambda \in \rho(T^*)$ und $R(\bar\lambda, T^*) = R(\lambda, T)^*$. Wendet man das gleiche Argument auf T^* an und verwendet, dass $T^{**} = T$ ist, so ergibt sich $\lambda \in \rho(T)$ genau dann, wenn $\bar\lambda \in \rho(T^*)$. □

Wir diskutieren nun noch die Beziehung zwischen Spektrum und numerischem Wertebereich. Es gilt

Proposition 4.1.8. *Ist H ein Hilbertraum und $T \in \mathscr{L}(H)$, so gilt $\sigma(T) \subset \overline{\Theta(T)}$. Insbesondere ist das Spektrum eines selbstadjungierten Operators reell.*

Beweis. Ist $\lambda \notin \overline{\Theta(T)}$, so ist $d := \inf\{|\lambda - \langle Tx, x\rangle| : \|x\| = 1\} > 0$. Für x mit $\|x\| = 1$ gilt also

$$0 < d \leq |\lambda - \langle Tx, x\rangle| = |\langle(\lambda - T)x, x\rangle| \leq \|(\lambda - T)x\|.$$

Dies zeigt, dass $\lambda - T$ injektiv ist und dass die Umkehrabbildung $(\lambda - T)^{-1} : \mathrm{rg}(\lambda - T) \to H$ beschränkt ist. Somit ist $\lambda - T$ ein Isomorphismus zwischen H und seinem Bild, insbesondere ist $\mathrm{rg}(\lambda - T)$ abgeschlossen. Angenommen $\mathrm{rg}(\lambda - T) \neq H$. Dann gibt es ein $x_0 \in \mathrm{rg}(\lambda - T)^\perp$ mit $\|x_0\| = 1$. Insbesondere ist $x_0 \perp (\lambda - T)x_0$, also

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

also $\lambda = \langle Tx_0, x_0\rangle$. Es müsste also $\lambda \in \Theta(T)$ sein, ein Widerspruch. Insgesamt haben wir gezeigt, dass $\lambda \in \rho(T)$. □

4.2 Kompakte Operatoren

Definition 4.2.1. Es seien X und Y normierte Räume. Eine lineare Abbildung $T : X \to Y$ heißt kompakt, falls für jede beschränkte Folge x_n die Folge Tx_n eine konvergente Teilfolge besitzt. Wir schreiben $\mathscr{K}(X, Y)$ für die Menge aller kompakten Operatoren von X nach Y.

Bemerkung 4.2.2. Bekanntlich heißt eine Teilmenge K eines metrischen Raumes (M, d) kompakt, falls jede Folge in K eine (in K!) konvergente Teilfolge besitzt. Eine Menge F heißt relativkompakt, falls ihr Abschluss kompakt ist. Somit ist ein linearer Operator genau dann kompakt, wenn er beschränkte Mengen auf relativkompakte Mengen abbildet.

Proposition 4.2.3. *Es seien X, Y, Z und V normierte Räume.*

(1) $\mathscr{K}(X, Y)$ ist ein abgeschlossener Teilraum von $\mathscr{L}(X, Y)$. Insbesondere ist $\mathscr{K}(X)$ ein Banachraum iezüglich $\|\cdot\|_{\mathrm{op}}$.

(2) Ist $T \in \mathscr{K}(X, Y)$, $R \in \mathscr{L}(V, X)$ und $S \in \mathscr{L}(Y, Z)$, so ist $STR \in \mathscr{K}(V, Z)$. Insbesondere ist für $X = Y$ $\mathscr{K}(X) := \mathscr{K}(X, X)$ ein Ideal *(im Sinne der Algebra) in $\mathscr{L}(X)$.*

Beweis. (1) Zunächst, ist jeder kompakte Operator T beschränkt. Ansonsten gäbe es nämlich eine Folge x_n mit $\|x_n\| = 1$ und $\|Tx_n\| \geq n$. Aber dann kann Tx_n keine konvergente Teilfolge haben.

Seien nun $T, S \in \mathscr{K}(X, Y)$, $\lambda \in \mathbb{K}$ und eine beschränkte Folge x_n gegeben. Für eine Teilfolge x_{n_k} konvergiert dann Tx_{n_k}, etwa gegen y. Geht man falls nötig zu einer weiteren Teilfolge $x_{n_{k_l}}$ über, so konvergiert auch $Sx_{n_{k_l}}$, etwa gegen z. Dann konvergiert aber $(T + \lambda S)x_{n_{k_l}}$ gegen $y + \lambda z$. Folglich ist $T + \lambda S$ kompakt.

Um die Abgeschlossenheit zu zeigen, verwenden wir ein Diagonalfolgenargument. Es sei T_n eine Folge in $\mathscr{K}(X, Y)$ mit $T_n \to T \in \mathscr{L}(X, Y)$. Ist eine beschränkte Folge x_n in X gegeben, so gibt es – weil T_1 kompakt ist – eine Teilfolge $(x_{n(1,k)})_{k\in\mathbb{N}}$ sodass $T_1 x_{n(1,k)}$ konvergiert, etwa gegen y_1. Weil wiederum T_2 kompakt ist gibt es eine Teilfolge $(x_{n(2,k)})_{k\in\mathbb{N}}$ von $(x_{n(1,k)})$, derart, dass auch $T_2 x_{n(2,k)}$ konvergiert, etwa gegen y_2. Wir iterieren dies und erhalten für $j \in \mathbb{N}$ Teilfolgen $x_{n(j,k)}$ derart, dass $x_{n(j,k)}$ eine Teilfolge von $x_{n(j-1,k)}$ ist und $T_j x_{n(j,k)}$ gegen ein Element y_j von Y konvergiert.

Wir setzen nun $x_{n_k} := x_{n(k,k)}$. Dann ist x_{n_k} eine Teilfolge der ursprünglichen Folge. Außerdem ist $(x_{n_k})_{k \geq j}$ eine Teilfolge von $x_{n(j,k)}$, sodass $T_j x_{n_k} \to y_j$ für *alle* $j \in \mathbb{N}$. Wir behaupten, dass (y_j) eine Cauchyfolge in Y ist. Es ist nämlich (oBdA $\|x_n\| \leq 1$)

$$\|y_n - y_m\| \leq \limsup_{k\to\infty} \big(\|y_n - T_n x_{n_k}\| + \|T_n - T_m\| + \|T_n x_{n_k} - y_m\|\big) \leq \|T_n - T_m\|.$$

Sei nun $y_j \to y$. Dann zeigt die Abschätzung

$$\begin{aligned}\|Tx_{n_k} - y\| &\leq \|Tx_{n_k} - T_j x_{n_k}\| + \|T_j x_{n_k} - y_j\| + \|y_j - y\| \\ &\leq \|T - T_j\| + \|T_j x_{n_k} - y_j\| + \|y_j - y\|,\end{aligned}$$

dass $Tx_{n_k} \to y$ für $k \to \infty$, es ist also T kompakt.

(2) Mit (x_n) ist auch Rx_n eine beschränkte Folge, weil ja R beschränkt ist. Weil T kompakt ist, besitzt TRx_n also eine konvergente Teilfolge, welche vom stetigen operator S wiederum auf eine konvergente Folge abgebildet wird. □

Bemerkung 4.2.4. Jeder Operator mit endlichdimensionalem Bild (engl. *finite rank operator*) ist ein kompakter Operator. Wegen Proposition 4.2.3 ist somit auch jeder Grenzwert einer Folge von Operatoren mit endlichdimensionalem Bild ein kompakter Operator. 1936 fragte Stanisław Mazur ob umgekehrt jeder kompakte Operator Grenzwert einer Folge von Operatoren mit endlichdimensionalem Bild ist. Als Preis für denjenigen, der dieses Problem löst, setzte er eine lebende Gans aus. Per Enflo hat 1972 gezeigt, dass es durchaus Banachräume gibt, in denen dies nicht der Fall ist (und hat dafür auch eine Gans erhalten).

Wir können nun ein einfaches Beispiel eines kompakten Operators geben.

Beispiel 4.2.5. Wir betrachten den Banachraum ℓ^p. Ist $\mathbf{m} \in c_0$, so ist der zugehörige Multiplikationsoperator M kompakt. Ist nämlich $\mathbf{m} \in c_{00}$ – dem Raum der Folgen, die schließlich Null sind – so hat der zugehörige Multiplikationsoperator endlichdimensionales Bild, ist also insbesondere kompakt. Ist nun ein allgemeines $\mathbf{m} = (m_k)_{k\in\mathbb{N}}$ gegeben, so setze $\mathbf{m}_n := (m_1, \ldots, m_n, 0, 0, \ldots) \in c_{00}$. Bezeichnet man die zugehörigen Multiplikationsoperatoren mit M respektive M_n, so ist $\|M_n - M\| = \sup_{k>n} |m_k| \to 0$. Aus Proposition 4.2.3 folgt, dass M als Grenzwert von kompakten Operatoren kompakt ist.

Beispiel 4.2.6. Ist $\dim H = \infty$, so ist die Identität I *nicht* kompakt. In diesem Fall gibt es nämlich ein abzählbares Orthonormalsystem $(e_n)_{n\in\mathbb{N}}$. Die Folge e_n ist beschränkt. Allerdings hat diese Folge keine konvergente Teilfolge. Wenn nämlich $e_{n_k} \to x$ konvergiert, so müsste $\langle x, e_m\rangle = \lim \langle e_{n_k}, e_m\rangle = 0$ für alle $m \in \mathbb{N}$ sein, also $x = 0$ da (e_n) eine Orthonormalbasis seines Aufspanns ist. Andererseits ist aber $\|e_{n_k}\| \equiv 0 \not\to 0$, sodass e_{n_k} nicht gegen 0 konvergieren kann.

4.3 Der Spektralsatz für kompakte, selbstadjungierte Operatoren

Wir zeigen zunächst, dass kompakte, selbstadjungierte Operatoren auf einem Hilbertraum einen Eigenwert besitzen.

Lemma 4.3.1. *Es sei H ein Hilbertraum, $T \in \mathscr{K}(H)$ selbstadjungiert. Dann liegt $\sigma(T)$ im Intervall $[-\|T\|, \|T\|]$ und $\|T\|$ oder $-\|T\|$ ist ein Eigenwert von T.*

Beweis. Weil das Spektrum eines selbsadjungierten Operators T reell ist (Proposition 4.1.8) und in $\overline{B(0, \|T\|)}$ liegt (Satz 4.1.5) ist die erste Aussage klar. Weil T selbstadjungiert ist, folgt aus Proposition 2.4.4(2), dass es eine Folge x_n mit $\|x_n\| \leq 1$ gibt, mit $|\langle Tx_n, x_n\rangle| \to \|T\|$. Wir nehmen ohne Einschränkung an, dass $\langle Tx_n, x_n\rangle \to \|T\| =: \lambda > 0$.

Weil T kompakt ist, konvergiert – nach Übergang zu einer Teilfolge – Tx_n, etwa gegen y. Nun ist aber

$$\begin{aligned}\limsup_{n\to\infty} \|Tx_n - \lambda x_n\|^2 &\leq \limsup_{n\to\infty} \left(\|Tx_n\|^2 + \lambda^2\|x_n\| - 2\lambda\langle Tx_n, x_n\rangle\right)\\ &\leq \lambda^2 + \lambda^2 - 2\lambda^2 = 0.\end{aligned}$$

Somit konvergiert auch λx_n und also auch x_n, etwa gegen x. Nun folgt aber $\lambda x = \lim \lambda Tx_n = Tx$ und somit ist λ Eigenwert von T mit Eigenvektor x (Beachte: Wäre $x = 0$, so müsste $\langle Tx_n, x_n\rangle \to \langle y, 0\rangle = 0$ konvergieren, also $\lambda = 0$ sein). □

Satz 4.3.2. *(Spektralsatz)*

Es sei H ein Hilbertraum und T ein kompakter, selbstadjungierter Operator auf H. Dann ist $\sigma(T) \setminus \{0\} = \sigma_p(T)$. Weiter gibt es eine (möglicherweise abbrechende) Folge $\lambda_n \to 0$ und ein Orthonormalsystem e_n, derart, dass

$$Tx = \sum_n \lambda_n \langle x, e_n\rangle e_n.$$

Insbesondere ist $Te_n = \lambda e_n$.

Beweis. Nach Lemma 4.3.1 ist $\lambda_1 = \|T\|$ oder $\lambda_1 = -\|T\|$ ein Eigenwert. Es sei e_1 mit $\|e_1\| = 1$ ein zugehöriger Eigenvektor. Sei nun $H = \operatorname{span}\{e_1\} \oplus R_1$, wobei $R_1 := \operatorname{span}\{e_1\}^\perp$ ist. Dann ist R_1 ein abgeschlossener Teilraum von H, $TR_1 \subset R_1$ (denn ist $y \perp e_1$, so ist $\langle e_1, Ty\rangle = \langle Te_1, y\rangle = \lambda_1\langle e_1, y\rangle = 0$) und die Einschränkung T_1 von T auf R_1 ist kompakt (denn $T_1 = P_{R_1}T\iota$, wo $\iota : R_1 \to H$ die kanonische Einbettung ist und P_{R_1} die orthogonale Projektion auf H_1 ist) und selbstadjungiert (das ist klar).

Somit kann man obiges Argument iterieren und erhählt eine normierte Folge e_n von normierten Eigenvektoren von T zum Eigenwert λ_n. Hierbei ist $H_n := \operatorname{span}\{e_1, \ldots, e_n\}$ und $R_n := H_n^\perp$, sodass TR_n invariant lässt und die Einschränkung T_n auf R_n wiederum kompakt und selbstadjungiert ist. Dann ist e_{n+1} der Eigenvektor zum Eigenwert $\pm\|T_n\|$, der laut Lemma 4.3.1 existiert. Beachte, dass $\|T_{n+1}\| \leq \|T_n\|$, also ist die Folge $|\lambda_n|$ fallend. Beachte weiter, dass $e_{n+1} \perp e_1, \ldots e_n$, nach Konstruktion.

Wir setzen nun $H_\infty := \overline{\operatorname{span}}\{e_n : n \in \mathbb{N}\}$ und $R_\infty = H_\infty^\perp$. Man sieht wie oben, dass T den Rest R_∞ invariant lässt. Wir schreiben T_∞ für die Einschränkung von T auf R_∞ und P_∞ für die Projektion auf R_∞. Dann ist $H = H_\infty \oplus_2 R_\infty$ und (e_n) ist eine Orthonormalbasis von H_∞. Es folgt

$$Tx = \sum_n \lambda_n\langle x, e_n\rangle e_n + T_\infty P_\infty x.$$

Es bleibt also nur zu zeigen, dass $T_\infty = 0$ ist. Es ist aber nach Konstruktion $\|T_\infty\| \leq \|T_n\| = |\lambda_n|$. Es genügt also zu zeigen, dass $\lambda_n \to 0$. Angenommen, dies wäre falsch. Dann gäbe es eine Teilfolge λ_{n_k} die gegen ein $\mu \neq 0$ konvergiert. Weil aber T kompakt ist, müsste in diesem Fall $Te_{k_n} = \lambda_{n_k}e_{k_n}$ eine konvergente Teilfolge besitzen. Aber dann müsste auch e_{k_n} eine konvergente Teilfolge besitzen, was absurd ist (weil $\|e_{k_n} - e_{k_m}\| = \sqrt{2}$ für $n \neq m$). □

Beispiel 4.3.3. Es sei $E = \ell^2$ und $\mathbf{m} = (m_n) \in c_0$ reell und M der assoziierte Mulitplikationsoperator. Dann ist M kompakt (siehe Beispiel 4.2.5) und selbstadjungiert, $\sigma(M) = \{m_n : n \in \mathbb{N}\} \cup \{0\}$ und die Orthonormalbasis aus Eigenwerten ist gerade die kanonische Basis $\mathbf{e}_n = (\delta_{kn})_{k\in\mathbb{N}}$.

Man kann den Spektralsatz auch so verstehen, dass man "im Wesentlichen" immer in dieser Situation eines Multiplikationsoperators ist.

Definition 4.3.4. Es seien H_1, H_2 unitäre Räume und $T_j \in \mathscr{L}(H_j)$ für $j = 1, 2$. Die Operatoren heißen *unitär äquivalent*, falls es einen unitären Operator $U : H_1 \to H_2$ gibt mit $T_2 = UT_1U^{-1}$. Mit anderen Worten, falls das Diagramm

$$\begin{array}{ccc} H_1 & \xrightarrow{T_1} & H_1 \\ {\scriptstyle U}\big\downarrow & & \big\uparrow{\scriptstyle U^{-1}} \\ H_2 & \xrightarrow[T_2]{} & H_2 \end{array}$$

kommutiert.

Korollar 4.3.5. *(Spektralsatz, Multiplikationsoperator Version)*
Es sei H ein separabler Hilbertraum und T ein kompakter, selbstadjungierter Operator. Dann ist T unitär äquivalent zu einem Multiplikationsoperator auf ℓ^2.

Beweis. Es sei (e_n) das Orthonormalsystem aus Satz 4.3.2, also $Te_n = \lambda_n e_n$. Indem wir (e_n) um eine Orthonormalbasis von $\ker T$ erweitern, können wir annehmen, dass (e_n) eine Orthonormalbasis von H ist. Es sei $U : H \to \ell^2$ die zugehörige Koordinatenabbildung, also $Ux = (\langle x, e_n\rangle)_{n\in\mathbb{N}}$. Dann ist UTU^{-1} der Multiplikationsoperator zur Folge (λ_n). $\square$

4.4 Kompakte Integraloperatoren

Wir hatten bereits bemerkt, dass Integraloperatoren $K \in \mathscr{L}(L^2(a, b))$, definiert durch

$$Kf(x) := \int_a^b k(x, y)f(y)\, dy, \tag{4.1}$$

für ein $k \in L^2((a, b)^2)$, in Anwendungen häufig auftreten und ihr Studium einer der Hauptgründe für die Entwicklung der Funktionalanalysis waren. Genauer will man sogenannte *Fredholmsche Integralgleichungen* der Form

$$\lambda u(x) - \int_a^b k(x, y)u(y)\, dy = f(x)$$

lösen. Diejenigen λ für die diese Gleichung *wohlgestellt* (im Sinne von Hadamard) ist sind genau diejenigen λ, die in der Resolventenmenge $\rho(K)$ liegen. Hierbei heißt die Gleichung wohlgestellt, falls für jedes $f \in L^2(a, b)$ eine eindeutige Lösung $u \in L^2(a, b)$ (eindeutige Lösbarkeit) und falls $f_n \to f$, so konvergieren auch die zugehörigen Lösungen $u_n \to u$ (stetige Abhängigkeit von den Daten). Das Spektrum von K besteht also genau aus denjenigen λ, für die die Gleichung nicht wohlgestellt ist.

Ist K kompakt und selbstadjungiert, so folgt aus dem Spektralsatz 4.3.2, dass $\sigma(K) = \sigma_p(K)$ höchstens bzählbar ist und 0 der einzig Mögliche Häufungspunkt von $\sigma(K)$ ist. Es stellt sich also die Frage, wann K kompakt und selbstadjungiert ist.

Lemma 4.4.1. *Ist $k \in L^2((a,b)^2)$ und K durch (4.1) gegeben, so ist der adjungierte Operator K^* gegeben durch*

$$K^*f(y) = \int_a^b \overline{k(x,y)} f(x)\, dx.$$

Insbesondere ist K genau dann selbstadjungiert, wenn $k(x,y) = \overline{k(y,x)}$ für fast alle (x,y) ist.

Beweis. Es seien $f, g \in L^2(a,b)$. Für den Moment sei der zum Kern $m(x,y) := \overline{k(y,x)}$ gehörende Kernoperator mit M bezeichnet. Mit dem Satz von Fubini folgt

$$\begin{aligned}
\langle Kf, g\rangle &= \int_a^b \int_a^b k(x,y) f(y)\, dy \overline{g(x)}\, dx \\
&= \int_a^b f(y) \int_a^b f(y) \overline{\overline{k(x,y)} g(x)}\, dx dy \\
&= \int_a^b f(y) Mg(y)\, dy = \langle f, Mg\rangle.
\end{aligned}$$ □

Es bleibt also die Frage zu klären, wann K kompakt ist, was wiederum zur Frage führt, wie kompakte Teilmengen von $L^2(a,b)$ aussehen. Zunächst beschreiben wir relativkompakte Teilmengen von $C([a,b])$.

Satz 4.4.2. *(Arzelà–Ascoli)*
Es sei M eine Teilmenge von $C([a,b])$ mit folgenden Eigenschaften:

(i) *M ist punktweise beschänkt, d.h. $\sup_{f\in M} |f(x)| < \infty$ für alle $x \in [a,b]$.*

(ii) *M ist gleichgradig stetig, d.h. zu $\varepsilon > 0$ gibt es ein $\delta > 0$ derart, dass für alle $f \in M$ und $x, y \in I$ mit $|x-y| \le \delta$ stets $|f(x) - f(y)| \le \varepsilon$ ist.*

Dann ist M relativkompakt.

Beweis. Es sei f_n eine Folge in M. Wir müssen zeigen, dass f_n eine (gleichmäßig) konvergente Teilfolge besitzt. Es sei $\{q_m : m \in \mathbb{N}\} = \mathbb{Q} \cap [a,b]$ eine Aufzählung der rationalen Zahlen in $[a,b]$. Weil für $m \in \mathbb{N}$ die Folge $(f_n(q_m))_{n\in\mathbb{N}}$ (wegen (i)!) beschränkt ist, besitzt diese Folge eine konvergente Teilfolge $f_{n(m,k)}(q_m))_{k\in\mathbb{N}}$. Mittels eines Diagonalfolgenarguments, siehe den Beweis von Proposition 4.2.3, finden wir eine Teilfolge $(f_{n_k})_{k\in\mathbb{N}}$ derart, dass $f_{n_k}(q_m)$ für alle $m \in \mathbb{N}$ konvergiert. Wir schreiben $f(q_m) := \lim_{k\to\infty} f_{n_k}(q_m)$.

Auf diese Weise ist eine Funktion $f : [a,b] \cap \mathbb{Q} \to \mathbb{K}$ definiert. Mittels Grenzübergang folgt sofort aus (ii), dass es zu $\varepsilon > 0$ ein $\delta > 0$ gibt, mit $|f(q) - f(r)| \le \varepsilon$ wenn nur $|q - r| \le \delta$. Hieraus folgt, dass sich f zu einer stetigen Funktion auf $[a,b]$ fortsetzen lässt. Ist nämlich $x \in [a,b]$, so gibt es eine Folge r_n in $[a,b] \cap \mathbb{Q}$ mit $r_n \to x$. Wegen obiger Eigenschaft ist $f(r_n)$ eine Cauchyfolge, also konvergent, etwa gegen $f(x)$. Die auf diese Art und Weise gewonnene Funktion f ist stetig. Um das zu sehen wählt man zu gegebenem ε den Parameter δ wie oben und r_n und s_n Folgen rationaler Zahlen mit$r_n \to x$ und $s_n \to y$ wobei $|x - y| \le \varepsilon/2$, so ist $|r_n - s_n| \le \varepsilon$ für genügend große n. Dann gilt

$$|f(x) - f(y)| \le \limsup_{n\to\infty} \big(|f(x) - f(r_n)| + |f(r_n) - f(s_n)| + |f(s_n) - f(y)|\big) \le \varepsilon$$

Dies zeigt, dass f in der Tat stetig ist.

Es bleibt zu zeigen, dass $f_k := f_{n_k} \to f$ bezüglich $\|\cdot\|_\infty$. Sei hierzu $\varepsilon > 0$ gegeben. Wir finden ein $\delta > 0$ sodass für x, y mit $|x - y| \le \delta$ stets $|f(x) - f(y)| \le \varepsilon$ und $|f_k(x) - f_k(y)| \le \varepsilon$

für alle $k \in \mathbb{N}$ gilt. Weil $[a,b]$ kompakt ist und $[a,b] \subset \bigcup B(q_m, \delta)$, gibt es rationale Zahlen $r_1, \ldots, r_n$ mit $[a,b] \subset B(r_1, \delta) \cup \ldots \cup B(r_n, \delta)$. Wähle nun k_0 so groß, dass $|f_k(r_j) - f(r_j)| \leq \varepsilon$ für $j = 1, \ldots, n$ und $k \geq k_0$. Ist nun $x \in [a,b]$ so findet sich ein j mit $|x - r_j| \leq \delta$ und daher

$$|f_k(x) - f(x)| \leq |f_k(x) - f_k(r_j)| + |f_k(r_j) - f(r_j)| + |f(r_j) - f(x)| \leq 2\varepsilon$$

für alle $k \geq k_0$. □

Beispiel 4.4.3. Ist $M \subset C([a,b])$ eine beschränkte Menge stetig differenzierbarer Funktionen mit $\sup_{f \in M} \|f'\|_\infty =: C < \infty$, so ist M präkompakt. Ist nämlich $\varepsilon > 0$ gegeben, so setze $\delta := (1 + C)^{-1}\varepsilon$. Sind dann $x, y \in [a,b]$ mit $|x - y| \leq \delta$ gegeben, so ist wegen des Mittelwertsatzes $f(x) - f(y) = f'(\xi)(x - y)$ für ein ξ zwischen x und y. Daher

$$|f(x) - f(y)| = |f'(\xi)||x - y| \leq C\delta \leq \varepsilon.$$

Dies zeigt, dass M gleichgradig stetig ist. Die punktweise Beschränktheit ist klar.

Beispiel 4.4.4. Für $a, b \in \mathbb{R}$ mit $a < b$ ist die Einbettung $\iota : H^1(a,b) \to C([a,b])$, $\iota(f) = \tilde{f}$ kompakt. Es ist zu zeigen, dass ι beschränkte Mengen auf präkompakte Mengen abbildet. Ist $M \subset H^1(a,b)$ beschränkt, etwa durch C, so ist $\iota(M)$ punktweise beschränkt, was sofort aus Satz 3.1.12 folgt. Weiterhin ist

$$|\iota(f)(t) - \iota(f)(s)| = \left| \int_s^t f'(r)\, dr \right| \leq \|f'\|_2 |t - s|^{\frac{1}{2}} \leq C|t - s|^{\frac{1}{2}}$$

was die gleichgradige Stetigkeit von $\iota(M)$ zeigt. Die Behauptung folgt somit aus dem Satz von Arzelà–Ascoli.

Beispiel 4.4.5. Es seien $a, b \in \mathbb{R}$ mit $a < b$. Ist $k \in C([a,b]^2)$, so ist der Kernoperator T_k, gegeben durch

$$T_k f(x) := \int_a^b k(x,y) f(y)\, dy$$

ein kompakter Operator von $L^1(a,b)$ nach $C([a,b])$. Insbesondere definiert T_k einen kompakten Operator auf $L^2(a,b)$.

In der Tat ist T_k offensichtlich beschränkt mit $\|T_k\|_{\mathrm{op}} \leq \|k\|_\infty$, woraus insbesondere folgt, dass T_k beschränkte Mengen auf (punktweise) beschränkte Mengen abbildet. Wir zeigen nocht, dass T_k beschränkte Mengen auf gleichgradig stetige Mengen abbildet, dann folgt die Behauptung aus dem Satz von Arzelà–Asoli.

Sei also $\varepsilon > 0$ gegeben. Aufgrund der Gleichstetigkeit von k auf dem Kompaktum $[a,b]^2$, gibt es ein $\delta > 0$ mit $|k(x,z) - k(y,z)| \leq \varepsilon$ für $x, y, z \in [a,b]$ mit $|x - y| \leq \varepsilon$. Für solche x, y gilt also

$$|T_k f(x) - T_k f(y)| \leq \int_a^b |k(x,z) - k(y,z)||f(z)|\, dz \leq \|f\|_1 \varepsilon,$$

was zeigt dass T_k beschränkte Mengen auf gleichgradig stetige abbildet.

Bezeichnen wir den Operator von $L^1(a,b)$ nach $C([a,b])$ für den Moment mit $\tilde{T}_k$ und den Operator auf $L^2(a,b)$ mit T_k, so ist $T_k = \iota_2 T \iota_1$, wobei ι_1 die stetige (!) Einbettung von $L^2(a,b)$ nach $L^1(a,b)$ ist und ι_2 die stetige Einbettung von $C([a,b])$ nach $L^2(a,b)$. Somit folgt die Kompaktheit von T_k auf L^2 aus der Idealeigenschaft kompakter Operatoren.

4.5* Das Spektrum des Dirichlet–Laplace Operators

Wir betrachten den sogenannten *Dirichlet–Laplace Operator* Δ_D, gegeben durch $\operatorname{def}(D) = H^2(0,1) \cap H_0^1(0,1)$ und $\Delta_D u = u''$. Dies ist kein beschränkter Operator, man kann allerdings zeigen, dass Δ_D abgeschlossen ist. Es folgt aus Satz 3.2.1, dass die Gleichung $-\Delta_D u = f$ für jedes $f \in L^2(0,1)$ genau eine Lösung besitzt. Wir wollen nun diese Lösung berechnen. Es liegt nahe, hierzu f zweimal zu integrieren. Allerdings müssen wir auf die Randbedingunen achten.

Ist $-u'' = f$ so ist

$$\begin{aligned}
u(x) &= -\int_0^x \int_0^t f(s)\,ds\,dt + c_0 x + c_1 \\
&= -\Big[t \int_0^t f(s)\,ds\Big|_{t=0}^x + \int_0^x s f(s)\,ds + c_0 x + c_1 \\
&= \int_0^x (s-x) f(s)\,ds + c_0 x + c_1.
\end{aligned}$$

Weil $u(0) = 0$ sein soll, muss $c_1 = 0$ sein. Weil $u(1) = 0$ sein soll, muss

$$c_0 = \int_0^1 (1-s) f(s)\,ds$$

sein. Daher ist

$$\begin{aligned}
u(x) &= \int_0^x (s-x) f(s)\,ds + x \int_0^1 (1-s) f(s)\,ds \\
&= \int_0^x (s-x) f(s)\,ds + \int_0^x x(1-s) f(s)\,ds + \int_x^1 x(1-s) f(s)\,ds \\
&= \int_0^x s(1-x) f(s)\,ds + \int_x^1 x(1-s) f(s)\,ds \\
&= \int_0^1 G(x,s) f(s)\,ds
\end{aligned}$$

wobei

$$G(x,y) := \begin{cases} y(1-x) & \text{falls } y \le x \\ x(1-y) & \text{falls } y > x. \end{cases}$$

Wir haben gezeigt, dass die Umkehrabbildung $-\Delta_D^{-1}$ ein Integraloperator mit Kern G ist. Der Kern G heißt *Greenfunktion* des Problems $-u'' = f$. Beachte, dass G stetig und symmetrisch ist. Somit ist $-\Delta_D^{-1}$ selbstadjungiert und kompakt. Insbesondere besteht das Spektrum von $-\Delta_D^{-1}$ mit Ausnahme der 0 aus Eigenwerten. Man sieht leicht, dass $\lambda - \Delta_D$ für $\lambda \neq 0$ genau dann bijektiv ist mit $(\lambda - \Delta_D)^{-1} \in \mathscr{L}(L^2(0,1))$, wenn $\lambda^{-1} - \Delta_D^{-1}$ invertierbar ist mit beschränkter Inverse, also genau dann, wenn $\lambda^{-1} \in \rho(\Delta_D^{-1})$. Ist weiterhin $f \in L^2(0,1)$ und $\lambda \in \sigma(\Delta_D^{-1})$ mit $\Delta_D^{-1} f = \lambda f$, so ist $f \in \operatorname{def}(\Delta_D)$ und $\Delta_D f = \lambda^{-1} f$. Insbesondere folgt aus dem Spektralsatz

Satz 4.5.1. *Das Spektrum des Dirichlet–Laplace Operators, also das Komplement der Menge*

$$\rho(\Delta_D) := \{\lambda \in \mathbb{C} : \lambda - \Delta_D \textit{ ist bijektiv und } (\lambda - \Delta_D)^{-1} \in \mathscr{L}(L^2(0,1))\},$$

besteht aus einer Folge λ_n mit $\lambda_n \uparrow \infty$. Weiterhin ist jedes λ_n ein Eigenwert von Δ_D und die zugehörigen (normierten) Eigenvektoren bilden eine Orthonormalbasis von $L^2(0,1)$.

Satz 4.5.1 gibt uns ein detailliertes Wissen über den Dirichlet–Laplaceoperator. Dieses Wissen erlaubt es uns, auch verwandte Probleme zu studieren. Als Beispiel diskutieren wir – informell – die *Wärmeleitungsgleichung*

$$\begin{cases} \dfrac{\partial}{\partial t}u(t,x) = \dfrac{\partial^2}{\partial x^2}u(t,x) \quad (t,x) \in (0,\infty)\times(0,1) \\ \quad u(t,0) = u(t,1) = 0 \\ \quad u(0,x) = f(x). \end{cases}$$

Gesucht ist also eine Funktion u von *zwei* Variablen t und x, sodass die partielle Differentialgleichung $\partial_t u = \partial_{xx} u$ erfüllt ist. Weiterhin verlangen wir, dass in der x-Variable Dirichlet Randbedingungen erfüllt sind. Zu guter letzt ist ein Anfangswert $u(0,x) = f(x)$ mit $f \in L^2(0,1)$ gegeben. Mit etwas gutem Willen können wir also die Wärmeleitungsgleichung zu einer *gewöhnlichen* Differentialgleichung

$$\begin{cases} u'(t) = \Delta_D u(t) \\ u(0) = f \end{cases}$$

auf dem Hilbertraum $L^2(0,1)$ umschreiben. Hier ist die Funktion u eine Funktion von *einer* Variablen, allerdings mit Werten in $L^2(0,1)$ (sodass sich hier die zweite Variable versteckt). Sei nun (e_k) die Orthonormalbasis von $L^2(0,1)$ aus Eigenwerten von Δ_D, also $\Delta_D e_k = \lambda_k e_k$. Wenn wir nun das Skalarprodukt obiger Gleichung mit e_k bilden, so folgt

$$\frac{d}{dt}\langle u(t), e_k\rangle = \langle u'(t), e_k\rangle = \langle \Delta_D u(t), e_k\rangle = \langle u(t), \Delta_D e_k\rangle = \lambda_k \langle u(t), e_k\rangle.$$

dies ist eine gewöhnliche Differentialgleichung für die skalarwertige Funktiong_k, gegeben durch $g_k(t) := \langle u(t), e_k\rangle$. Ist der Anfangswert $u(0) = e_k$, also $g_k(0) = 1$, so folgt $g_k(t) = e^{\lambda_k t}$. Entwickelt man einen beliebigen Vektor $f \in L^2(0,1)$ in der Basis $f = \sum \langle f, e_k\rangle e_k$, so folgt

$$u(t) = \sum_{k=1}^{\infty} e^{\lambda_k t}\langle f, e_k\rangle e_k.$$

Beachte, dass $\lambda_k < 0$ für alle k, sodass diese Reihe in der Tat konvergiert. Somit haben wir, zumindest informell, die Wärmeleitungsgleichung gelöst. Mehr Details werden in der Vorlesung “Partielle Differentialgleichungen” gegeben.