

# Vorlesungsmitschrift Funktionalanalysis 1

Vorlesung von Prof. Dr. Konrad Schmüdgen  
Universität Leipzig

Wintersemester 2008/2009

## Wichtige Vorbemerkungen

Bei diesem Skript handelt es sich um eine Vorlesungsmitschrift einiger Studenten. Die in der Vorlesung gebrachten Inhalte wurden erweitert und ergänzt; dennoch folgen die Inhalte weitestgehend den Konzeptionen und Ideen des Professors und sind deshalb in gewisser Weise sein „geistiges Eigentum“. Wir bitten, das bei der Verbreitung des Skriptes respektvoll zu berücksichtigen. Eigene Ansprüche (Copyright oder so etwas) melden wir nicht an. Die Erweiterung und Korrektur dieses Skriptes ist im Gegenteil ausdrücklich erwünscht.

Ein Hinweis an dieser Stelle: Skripte dienen dazu, behandelten Stoff nachschlagen oder unbekanntem Stoff erarbeiten zu können. Sie ersetzen weder eine Vorlesung, noch die Lektüre von Büchern oder den Besuch von Seminaren. Nur weil ein Sachverhalt im Skript steht, hat man diesen noch nicht im Kopf und kann ihn dementsprechend auch noch längst nicht anwenden. Wir selbst haben die Erfahrung gemacht, dass bloßes Mitlesen im Skript und das Vertrauen auf die dort stehenden Sätze und Formeln, ohne sie zu verinnerlichen, wenig Erkenntnisgewinn in der Mathematik mit sich bringt und auch fast unweigerlich zum Nichtbestehen einer eventuellen Klausur führt. Dass dieses Skript als Ergänzung zum Studium und nicht als (unzulänglicher) Ersatz zum Besuch und zur Nacharbeit einer Vorlesung gesehen wird, ist uns (und hoffentlich auch euch) ein wichtiges Anliegen. Und: Bitte nehmt diese Warnung ernst.

# Inhaltsverzeichnis

<b>1</b>	<b>Banachräume und Hilberträume</b>	<b>5</b>
1.1	Metrische Räume . . . . .	5
1.1.1	Topologische Grundbegriffe . . . . .	6
1.1.2	Vollständigkeit . . . . .	7
1.1.3	Kompaktheit . . . . .	8
1.2	Normierte lineare Räume . . . . .	9
1.2.1	Die $L^p$ -Räume . . . . .	11
1.2.2	Die Sobolevräume . . . . .	13
1.3	Hilberträume . . . . .	16
1.4	Die Rieszschen Sätze . . . . .	19
1.5	Orthonormalsysteme und Fourierentwicklung im Hilbertraum . . . . .	22
<b>2</b>	<b>Grundprinzipien der Funktionalanalysis</b>	<b>29</b>
2.1	Das Bairesche Kategoriethorem . . . . .	29
2.2	Stetige lineare Abbildungen auf normierten Vektorräumen . . . . .	31
2.2.1	Duale Räume . . . . .	36
2.3	Das Theorem von Banach-Steinhaus . . . . .	37
2.4	Offene Abbildung und abgeschlossener Graph . . . . .	39
2.5	Das Hahn-Banach-Theorem . . . . .	41
<b>3</b>	<b>Beschränkte lineare Operatoren im Hilbertraum</b>	<b>51</b>
3.1	Adjungierte, selbstadjungierte und unitäre Operatoren . . . . .	51
3.1.1	Adjungierte Operatoren . . . . .	51
3.1.2	Selbstadjungierte Operatoren . . . . .	53
3.1.3	Isometrische Operatoren und unitäre Operatoren . . . . .	54
3.2	Projektionsoperatoren . . . . .	57
3.3	Konvergenzarten im Hilbertraum . . . . .	59
3.4	Spektrum mit Resolvente . . . . .	61
3.4.1	Das Spektrum selbstadjungierter Operatoren . . . . .	66
3.5	Kompakte Operatoren im Hilbertraum . . . . .	68
3.5.1	Kompakte und vollstetige Operatoren in Banachräumen . . . . .	68
3.5.2	Kompakte Operatoren in Hilberträumen . . . . .	71
3.5.3	Der Hilbert-Schmidtsche Entwicklungssatz . . . . .	73

3.5.4	Die kanonische Form kompakter Operatoren . . . . .	74
3.5.5	Singuläre Zahlen und Schattenklassen . . . . .	76
3.5.6	Integraloperatoren mit quadratisch integrierbarem Kern . . . . .	76
<b>4</b>	<b>Das Spektraltheorem</b>	<b>79</b>
4.1	Spektralscharen . . . . .	80
4.1.1	Eigenschaften einer Spektralschar . . . . .	81
4.1.2	Spektralmaße . . . . .	82
4.1.3	Integration bezüglich einer Spektralschar . . . . .	83
4.2	Funktionalkalkül für beschränkte selbstadjungierte Operatoren . . . . .	86

# Kapitel 1

## Banachräume und Hilberträume

In diesem einleitenden Kapitel stellen wir grundlegende Ideen vor, die die Beschreibung von in der Funktionalanalysis wichtigen Räumen ermöglichen. Weiterhin werden wir auf Abbildungen zwischen diesen Räumen eingehen, auf Eigenschaften dieser Abbildungen und auf Arten, wie man sie darstellen kann. Der in diesem Kapitel vorgestellte Hilbertraum ist für unsere weiteren Untersuchungen von fundamentaler Bedeutung, denn die mathematische Struktur des Hilbertraums ermöglicht Beweise und Techniken, die auf Räumen mit weniger Struktur nicht zur Verfügung stehen. Einige pathologische Beispiele mögen davon Zeugnis bringen. Dennoch werden wir uns hier auf grundlegende Konzepte beschränken. In den späteren Kapiteln folgen dann tiefergehende Details.

Wir werden stets von Kenntnissen in Analysis und Algebra ausgehen, wie sie in Einführungsvorlesungen gelehrt werden. Aussagen wie beispielsweise die, dass eine stetige Funktion auf kompakten Mengen ihr Minimum und Maximum annimmt, werden wir nicht explizit erwähnen, wenn sie in Beweisen benötigt werden. Weiterhin setzen wir grundlegende Kenntnisse aus der Maßtheorie voraus, insbesondere sollte das Lebesgue-Integral bekannt sein.

### 1.1 Metrische Räume

Die folgenden Konzepte sind eventuell schon aus einführenden Vorlesungen bekannt. Wir werden hier nicht in tiefere topologische Konzepte eindringen können, diese kann man beispielsweise in einigen Büchern nachlesen. Eine grundlegende Idee, um auf beliebigen Mengen eine Art abstrakte Abstandsmessung zu ermöglichen, ist die der Metrik.

**Definition 1.1.1** (Metrischer Raum). Sei  $E$  eine Menge. Eine *Metrik* ist eine Abbildung  $d : E \times E \rightarrow \mathbb{R}$ , so dass für alle  $x, y, z \in E$  gilt:

$$(i) \quad d(x, y) \geq 0, \quad d(x, y) = 0 \Leftrightarrow x = y, \quad (\text{Definitheit})$$

$$(ii) \quad d(x, y) = d(y, x), \quad (\text{Symmetrie})$$

$$(iii) \quad d(x, z) \leq d(x, y) + d(y, z). \quad (\text{Dreiecksungleichung})$$

Die Menge  $E$  mit einer Metrik  $d$  heißt *metrischer Raum*.

Einige häufig gebrachte Beispiele sollen diesen Begriff der Metrik verdeutlichen. Insbesondere sind die bereits in der Analysis verwendeten Euklidischen Abstände Metriken. Es gibt jedoch auch kompliziertere Metriken, die beispielsweise über Integrale definiert sind. Je unanschaulicher die verwendeten Räume sind, desto weiter entfernt sich meist auch die verwendete Metrik von dem anschaulichen Begriff eines Abstandes.

**Beispiel 1.1.2.**

- (a) Sei  $E = \mathbb{R}^3$  und  $d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2}$  sei die *Euklidische Metrik* auf  $E$ . Dann ist  $(E, d)$  ein metrischer Raum.
- (b) Sei  $E = C[a, b]$  der Raum der stetigen Funktionen, wahlweise mit einer der Normen

$$d_1(f, g) = \max_{x \in [a, b]} |f(x) - g(x)|$$

oder

$$d_2(f, g) = \int_a^b |f(x) - g(x)| dx.$$

Dann ist sowohl  $(E, d_1)$  als auch  $(E, d_2)$  ein metrischer Raum.

- (c) Eine beliebige Menge mit der Metrik

$$d(x, y) = \begin{cases} 0 & \text{wenn } x = y \\ 1 & \text{wenn } x \neq y \end{cases}$$

ist ebenfalls ein metrischer Raum. Die Metrik  $d$  wird auch als *diskrete Metrik* bezeichnet. An diesem Beispiel sieht man, dass sich jede Menge zumindest mit der diskreten Metrik metrisieren lässt und es sich daher bei dem Begriff des metrischen Raumes um eine relativ schwache Definition handelt. Insbesondere werden wir viele nützliche Resultate nicht für metrische Räume beweisen können, sondern erst zusätzliche Strukturforderungen stellen müssen.

Auch aus bereits vorhandenen Metriken kann man neue gewinnen. Sei beispielsweise  $(E, d)$  ein metrischer Raum und  $P \in E$ . Dann kann man eine weitere Metrik  $d_1$  durch  $d_1(x, y) = d(x, P) + d(P, y)$  für  $x \neq y$  und  $d_1(x, y) = 0$  für  $x = y$  einführen, wobei  $x, y \in E$ . Diese Metrik wird auch die Metrik der französischen Eisenbahn (mit dem Punkt  $P$  als Paris) genannt, da die Zugverbindungen in Frankreich meistens über die Hauptstadt laufen.

### 1.1.1 Topologische Grundbegriffe

Um benötigte Begriffe einführen zu können, brauchen wir einige (wahrscheinlich bereits bekannte) Termini. Sie verallgemeinern die für die Metrik  $d(x, y) = |x - y|$  bekannten Begriffe aus der reellen Analysis auf metrische Räume. Es sei  $(E, d)$  ein metrischer Raum. Mit  $U_\varepsilon(x) = \{y \in E : d(x, y) \leq \varepsilon\}$  bezeichnen wir die  $\varepsilon$ -Umgebung von  $x$  für alle  $x \in E$  und  $\varepsilon > 0$ . Sei desweiteren  $M \subseteq E$ . Ein Punkt  $x \in M$  heißt *innerer Punkt* von  $M$  (im metrischen Raum  $(E, d)$ ), wenn es ein  $\varepsilon > 0$  gibt mit  $U_\varepsilon(x) \subseteq M$ .  $M$  heißt *offen*, wenn jeder Punkt von  $M$  ein innerer Punkt von  $M$  ist.

Weiter heißt eine Folge  $(x_n)_{n \in \mathbb{N}}$  von Elementen aus  $E$  *konvergent* in  $(E, d)$ , wenn es ein Element  $x \in E$  gibt mit  $\lim_{n \rightarrow \infty} d(x_n, x) = 0$ . Man kann zeigen, dass das sogenannte Grenzelement  $x$  eindeutig bestimmt ist durch die Folge  $(x_n)_{n \in \mathbb{N}}$ , falls diese konvergiert. Man verwendet die Schreibweisen  $x = \lim_{n \rightarrow \infty} x_n$  oder  $x_n \rightarrow x$  für  $n \rightarrow \infty$ .

Die Menge  $M$  heißt *abgeschlossen*, wenn das Grenzelement jeder im Raum  $(E, d)$  konvergenten Folge von Elementen aus  $M$  selbst zu  $M$  gehört. Der *Abschluß*  $\overline{M}$  von  $M$  ist die Menge aller Grenzelemente von im Raum  $(E, d)$  konvergenten Folgen  $(x_n)_{n \in \mathbb{N}}$  mit  $x_n \in M$  für alle  $n \in \mathbb{N}$ .

**Definition 1.1.3.** Eine Teilmenge  $M$  eines reellen Vektorraums heißt *konvex*, wenn aus  $x, y \in M$  und  $\lambda \in [0, 1]$  stets auch  $\lambda x + (1 - \lambda)y \in M$  folgt.

#### Beispiel 1.1.4.

- (a) Sei  $E = \mathbb{N}$  und  $d(x, y) = |x - y|$ , sowie  $M \subseteq \mathbb{N}$ . Sei weiter  $m \in M$ . Betrachte die Umgebung

$$U_{\frac{1}{2}}(m) = \{n \in E : |m - n| \leq \frac{1}{2}\} = \{m\} \subseteq M.$$

Deshalb ist  $m$  ein innerer Punkt und  $M$  ist offen. Die Begriffe wie Offenheit hängen also stark davon ab, welchen metrischen Raum wir betrachten. Wäre  $E = \mathbb{R}$  und  $M \subseteq \mathbb{N}$  gewesen, so wäre die Umgebung  $U_{\frac{1}{2}}(m)$  das Intervall  $[m - \frac{1}{2}, m + \frac{1}{2}]$  und  $m$  kein innerer Punkt gewesen.

- (b) Betrachte den metrischen Raum  $(E, d)$  mit  $E = (0, 1] \cup [2, +\infty)$  und der Metrik  $d(x, y) = |x - y|$ . Die Menge  $M = (0, 1]$  ist offen in  $(E, d)$ , obwohl sie in  $(\mathbb{R}, |\cdot|)$  nur halboffen wäre.

### 1.1.2 Vollständigkeit

Der Begriff der Vollständigkeit ist ebenfalls aus der Analysis bekannt. Eines der ersten und wichtigsten Resultate ist dort, dass die reellen Zahlen vollständig sind, dass also jede Cauchyfolge auch konvergiert, was bei den rationalen Zahlen nicht der Fall ist. Diese Begriffe kann man ohne weiteres auf metrische Räume verallgemeinern.

**Definition 1.1.5** (Cauchyfolge). Eine Folge  $(x_n)_{n \in \mathbb{N}}$  aus  $(E, d)$  heißt eine *Cauchyfolge*, wenn für alle  $\varepsilon > 0$  ein  $n_\varepsilon \in \mathbb{N}$  existiert, so dass für alle  $n, m \geq n_\varepsilon$  gilt  $d(x_n, x_m) < \varepsilon$ .

#### Beispiel 1.1.6.

- (a) Jede konvergente Folge ist eine Cauchyfolge, die Umkehrung gilt aber im Allgemeinen nicht.
- (b) Sei  $E = \mathbb{Q}$  und  $(x_n)_{n \in \mathbb{N}}$  eine Folge mit rationalen Einträgen, die in  $\mathbb{R}$  gegen  $\sqrt{2}$  konvergiert. Die Folge ist eine Cauchyfolge in  $(\mathbb{Q}, d)$ , konvergiert dort aber nicht, da  $\sqrt{2}$  nicht zu  $\mathbb{Q}$  gehört. Es soll noch einmal betont werden, dass der metrische Raum, auf den man sich bezieht, von großer Wichtigkeit ist.

**Definition 1.1.7** (Vollständigkeit). Ein metrischer Raum  $(E, d)$  heißt *vollständig*, wenn jede Cauchyfolge aus  $(E, d)$  im metrischen Raum  $(E, d)$  konvergiert.

Die Definition von Vollständigkeit hängt sowohl von der Menge  $E$  als auch von der verwendeten Metrik  $d$  ab. Bezüglich einer anderen Metrik  $d'$  kann ein metrischer Raum  $(E, d')$  durchaus vollständig sein, auch wenn  $(E, d)$  nicht vollständig ist. Einige Beispiele mögen noch einmal verdeutlichen, dass längst nicht jeder metrische Raum die Vollständigkeitseigenschaft besitzt.

#### Beispiel 1.1.8.

- (a) Sei  $E$  eine beliebige Menge und  $d$  die diskrete Metrik.  $(E, d)$  ist immer vollständig, völlig unabhängig von der Beschaffenheit der Menge  $E$ .
- (b) Wie bereits erwähnt ist  $(\mathbb{R}, |\cdot|)$  vollständig.
- (c) Der metrische Raum  $(\mathbb{Q}, |\cdot|)$  ist hingegen nicht vollständig (siehe Beispiel 1.1.6).

**Definition 1.1.9** (Teilraum).  $(E, d)$  sei ein metrischer Raum,  $E_1$  eine Teilmenge von  $E$ . Die Einschränkung der Metrik  $d$  auf  $E_1$  ist eine Metrik auf  $E_1$ .  $E_1$  zusammen mit dieser eingeschränkten Metrik heißt *Teilraum* von  $(E, d)$ .

**Definition 1.1.10** (dicht).  $(E, d)$  sei ein metrischer Raum. Eine Teilmenge  $M$  von  $E$  heißt *dicht* in  $E$ , wenn es zu jedem Element  $x \in E$  eine Folge  $(x_n)_{n \in \mathbb{N}}$  von Elementen aus  $M$  gibt mit  $x = \lim_{n \rightarrow \infty} x_n$ .

Beispielsweise ist  $\mathbb{Q}$  dicht in  $\mathbb{R}$  bezüglich des Betrags. Oft dienen dichte Mengen dazu, einen Raum zu „approximieren“, denn in jeder noch so kleinen Umgebung eines Punktes aus  $E$  liegt stets ein Punkt aus  $M$ . Ohne Beweis geben wir den folgenden Satz, der garantiert, dass man jeden metrischen Raum dicht in einen vollständigen metrischen Raum einbetten kann.

**Satz 1.1.11** (Vollständige Erweiterung).  $(E_1, d_1)$  sei ein metrischer Raum. Dann existiert ein vollständiger metrischer Raum  $(E, d)$  mit folgenden Eigenschaften:

- (i)  $(E_1, d_1)$  ist ein Teilraum von  $(E, d)$ , so dass  $d_1$  die auf  $E_1$  eingeschränkte Metrik  $d$  ist, das heißt  $d(x, y) = d_1(x, y)$  für alle  $x, y \in E_1$ ,
- (ii)  $E_1$  ist dicht in  $E$ .

Der Raum  $(E, d)$  heißt dann die vollständige Erweiterung von  $(E_1, d_1)$ .

### 1.1.3 Kompaktheit

Sei  $(E, d)$  ein metrischer Raum und  $I$  eine Indexmenge. Eine offene Überdeckung ist eine Familie  $\{U_i, i \in I\}$  von offenen Teilmengen von  $E$ , so dass  $\bigcup_{i \in I} U_i = E$  ist, also die  $U_i$  den gesamten Raum „überdecken“.

**Satz 1.1.12.**  $(E, d)$  sei ein metrischer Raum und  $M$  eine nichtleere Teilmenge von  $E$ . Folgende Aussagen sind äquivalent:

- (i) Jede offene Überdeckung  $\{U_i, i \in I\}$  von  $M$  besitzt eine endliche Teilüberdeckung (Überdeckungskompaktheit)
- (ii) aus jeder Folge  $(x_n)_{n \in \mathbb{N}}$  von Elementen aus  $M$  kann man eine Teilfolge  $(x_{n_k})_{k \in \mathbb{N}}$  auswählen, die im Raum  $(E, d)$  gegen ein Element  $x \in M$  konvergiert. (Folgenkompaktheit)

Folgenkompaktheit besagt also, dass man aus jeder Folge eine konvergente Teilfolge auswählen kann. Vergleiche dies mit dem Satz von Bolzano-Weierstraß aus der reellen oder komplexen Analysis, der dies für beschränkte Folgen garantiert.

**Definition 1.1.13** (Kompaktheit).  $M$  heißt *kompakt*, wenn eine der äquivalenten Bedingungen (i) oder (ii) aus Satz 1.1.12 erfüllt ist.

**Bemerkung 1.1.14.** Sei  $M$  Teilmenge vom  $\mathbb{R}^n$  oder  $\mathbb{C}^n$  und  $d$  die euklidische Metrik. Dann ist  $M$  kompakt genau dann, wenn  $M$  abgeschlossen und beschränkt ist. Dies ist die Aussage des Satzes von Heine-Borel.

Im allgemeinen Fall gilt dieser Zusammenhang jedoch nicht. Beispielsweise ist eine Teilmenge eines durch die diskrete Metrik metrisierten Raumes stets abgeschlossen und beschränkt, aber nur dann kompakt, wenn sie endlich ist. Weiterhin gibt es in allen unendlichdimensionalen Vektorräumen abgeschlossene und beschränkte Mengen, die nicht kompakt sind.

Man kann noch weitere, schwächere Arten von Kompaktheit definieren.

**Definition 1.1.15.**  $(E, d)$  heißt

- (i) *lokalkompakt*, wenn zu jedem Punkt  $x \in E$  ein  $\varepsilon > 0$  existiert, so dass  $U_\varepsilon(x)$  kompakt ist.
- (ii)  *$\sigma$ -kompakt*, wenn es kompakte Teilmengen  $K_n$ ,  $n \in \mathbb{N}$ , von  $E$  gibt mit  $E = \bigcup_{n \in \mathbb{N}} K_n$ .

**Beispiel 1.1.16.** Sei  $E = \mathbb{R}^n$  oder  $E = \mathbb{C}^n$ .

- (a)  $E$  ist lokalkompakt, denn  $U_\varepsilon(x) = \{y \in E : d(x, y) \leq \varepsilon\}$  ist abgeschlossen und beschränkt und damit kompakt.
- (b)  $E$  ist  $\sigma$ -kompakt, denn  $E = \bigcup_{n=0}^{\infty} U_n(0)$  und  $U_n(0)$  ist kompakt.
- (c) Jede abgeschlossene Teilmenge des  $\mathbb{R}^n$  und des  $\mathbb{C}^n$  ist lokalkompakt.
- (d) Die Menge  $\mathbb{Q}$  ist nicht lokalkompakt, denn in jeder Umgebung jedes Punktes liegt eine Folge rationaler Zahlen, die in  $\mathbb{R}$  gegen eine irrationale Zahl konvergiert. Damit ist jede dieser Umgebungen nicht folgenkompakt.



## 1.2 Normierte lineare Räume

**Definition 1.2.1** (Norm). Sei  $E$  ein Vektorraum über  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$ . Eine *Norm* auf  $E$  ist eine Abbildung  $\|\cdot\| : E \rightarrow [0, \infty)$  mit folgenden Eigenschaften:

- (i)  $\|x\| = 0 \Rightarrow x = 0$ , (Definitheit)
- (ii)  $\|\lambda x\| = |\lambda| \|x\|$  für alle  $\lambda \in \mathbb{K}$  und  $x \in E$ , (Homogenität)
- (iii)  $\|x + y\| \leq \|x\| + \|y\|$  für alle  $x, y \in E$ . (Dreiecksungleichung)

Wenn nur (ii) und (iii) gelten, nennen wir  $\|\cdot\|$  eine *Halbnorm*. Ein Vektorraum mit Norm heißt *normierter linearer Raum*.

### Beispiel 1.2.2.

- (a)  $E = \mathbb{R}^2$  mit den jeweiligen Normen  $\|(x, y)\|_1 = \sqrt{x^2 + y^2}$ ,  $\|(x, y)\|_2 = 3|x| + |y|$  oder  $\|(x, y)\|_3 = \max(|x|, |y|) + 5|y|$  ist ein normierter linearer Raum, wie man leicht nachprüft.
- (b) Eine Halbnorm auf dem  $\mathbb{R}^2$  ist beispielsweise durch  $\|(x, y)\| = |x|$  definiert. Hier ist  $\|\cdot\|$  für die gesamte  $y$ -Achse Null, weshalb die Definitheit zur Norm fehlt.
- (c) Sei  $E$  der Vektorraum aller Funktionen von einer Menge  $X$  nach  $\mathbb{R}$ . Für jedes  $x \in X$  ist durch  $p_x(f) = |f(x)|$  mit  $f \in E$  eine Halbnorm auf  $E$  definiert.
- (d) Die gewöhnlichen Absolutbeträge auf  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$  sind eine Norm auf  $\mathbb{K}$  als  $\mathbb{K}$ -Vektorraum.

**Satz 1.2.3** (Induzierte Metrik).  $(E, \|\cdot\|)$  sei ein normierter linearer Raum. Wir definieren  $d(x, y) = \|x - y\|$  für alle  $x, y \in E$ .  $d$  ist dann eine Metrik auf  $E$  und  $(E, d)$  ist ein metrischer Raum.  $d$  heißt die von  $\|\cdot\|$  induzierte Metrik. □

Damit ist jeder normierte Raum ein metrischer Raum, was den Begriff des normierten Raumes stärker als den des metrischen Raumes macht. Die Umkehrung gilt nämlich nicht, wie man leicht am Beispiel der diskreten Metrik sieht, denn es gibt keine Norm, die sie induziert.

Durch die induzierte Metrik sind auf jedem normierten linearen Raum automatisch alle Begriffe definiert, die wir im Abschnitt 1.1 eingeführt hatten. Insbesondere wissen wir, was Konvergenz auf einem normierten linearen Raum ist und wie eine Cauchyfolge aussieht. Konvergenz  $x_n \rightarrow x$  bezüglich einer Norm ist nichts anderes, als dass  $\|x_n - x\|$  für  $n \rightarrow \infty$  im Sinne der Konvergenz von reellen Zahlen gegen 0 geht. Damit wissen wir auch, was Vollständigkeit bedeutet und kommen zu einer wichtigen Definition.

**Definition 1.2.4** (Banachraum). Ein vollständiger normierter linearer Raum heißt *Banachraum*.

Der Begriff des Banachraums ist fundamental für unsere weiteren Ergebnisse. Ohne die Vollständigkeit können fast keine strukturellen Aussagen getroffen werden, die für die Funktionalanalysis von Bedeutung sind. Auch wenn wir uns später auf noch stärker strukturierte Räume, die Hilberträume, zurückziehen werden, können doch schon viele starke Resultate für Banachräume gezeigt werden. Um den Begriff etwas auszuleuchten, bringen wir einige Beispiele für Banachräume und Beispiele für Räume, die keine Banachräume sind.

### Beispiel 1.2.5.

- (a) Sei  $E = \mathbb{R}^d$  oder  $E = \mathbb{C}^d$  mit  $\|x\| = \sqrt{|x_1|^2 + \dots + |x_d|^2}$ . Dann ist  $(E, \|\cdot\|)$  ein Banachraum. Wir werden in Kürze zeigen, dass jeder endlichdimensionale Vektorraum vollständig ist.

- (b) Sei  $E$  der Vektorraum aller stetigen (und damit beschränkten) Funktionen auf einem kompakten metrischen Raum  $(X, d)$  mit  $\|f\| = \sup_{x \in X} |f(x)|$ .  $(E, \|\cdot\|)$  ist ein Banachraum. In diesem Beispiel ist die Konvergenz einer Folge von Funktionen im normierten linearen Raum  $(E, \|\cdot\|)$  genau die gleichmäßige Konvergenz der Funktionenfolge auf der Menge  $E$ , denn diese ist ebenfalls über das Supremum definiert.
- (c) Sei  $E = \mathcal{C}^n[a, b]$  mit  $n \in \mathbb{N}_0$  und  $a, b \in \mathbb{R}$ ,  $a < b$ . Definiere  $\|f\|_n = \sum_{j=0}^n \sup_{x \in [a, b]} |f^{(j)}(x)|$ . Dann ist  $(E, \|\cdot\|_n)$  ein Banachraum. Mit einer Norm, die die Ableitungen nicht berücksichtigt, wäre  $(E, \|\cdot\|)$  nicht vollständig.
- (d) Sei  $E = \mathcal{C}[a, b]$  mit  $a, b \in \mathbb{R}$  und  $a < b$  sowie  $\|f\| = \int_a^b |f(t)| dt$ .  $(E, \|\cdot\|)$  ist zwar ein normierter linearer Raum, aber kein Banachraum. Man konstruiert ein Gegenbeispiel, indem man eine Folge von stetigen Funktionen auswählt, die gegen eine unstetige Funktion, beispielsweise eine Treppenfunktion konvergiert.

Auch wenn man auf vielen Vektorräumen die unterschiedlichsten Normen definieren kann, zerfallen diese doch meist in mehrere Äquivalenzklassen. Normen der gleichen Klasse sind dabei nicht wesentlich verschieden, sie definieren beispielsweise die gleiche Form der Konvergenz.

**Definition 1.2.6.** Zwei Normen  $\|\cdot\|_1$  und  $\|\cdot\|_2$  auf einem Vektorraum  $E$  heißen *äquivalent*, wenn es Konstanten  $c_1 > 0$  und  $c_2 > 0$  gibt mit

$$\|x\|_1 \leq c_1 \|x\|_2 \quad \text{und} \quad \|x\|_2 \leq c_2 \|x\|_1$$

für alle  $x \in E$ .

**Satz 1.2.7.** Zwei Normen  $\|\cdot\|_1$  und  $\|\cdot\|_2$  sind genau dann äquivalent, wenn die normierten linearen Räume  $(E, \|\cdot\|_1)$  und  $(E, \|\cdot\|_2)$  die gleiche Konvergenz definieren, das heißt wenn eine Folge  $(x_n)_{n \in \mathbb{N}}$  aus  $E$  in  $(E, \|\cdot\|_1)$  genau dann konvergiert, wenn sie in  $(E, \|\cdot\|_2)$  konvergiert.

Es folgt das verblüffende Resultat, dass es auf endlichdimensionalen normierten Vektorräumen keine wesentlich verschiedenen Normen gibt. Dies liegt vor allem daran, dass in endlichdimensionalen Räumen eine Menge genau dann kompakt ist, wenn sie abgeschlossen und beschränkt ist, siehe Bemerkung 1.1.14.

**Satz 1.2.8.** Auf einem endlichdimensionalen normierten linearen Raum sind alle Normen äquivalent.

*Beweis.* Sei  $E$  ein endlichdimensionaler Vektorraum. O.B.d.A. sei  $E = \mathbb{R}^d$  oder  $E = \mathbb{C}^d$ . Sei weiter  $\|\cdot\|$  eine beliebige Norm auf  $E$ .  $\|x\|_1 = |x_1| + \dots + |x_d|$ , die sogenannte 1-Norm, ist ebenfalls eine Norm auf  $E$ . Da die Äquivalenz von Normen eine Äquivalenzrelation ist, reicht es zu zeigen, dass  $\|\cdot\|$  und  $\|\cdot\|_1$  äquivalent sind.

Sei  $x = (x_1, \dots, x_d) \in E$  und  $e_k = (0, \dots, 0, 1, 0, \dots, 0)$  sei der  $k$ -te Einheitsvektor mit einem Eintrag an der  $k$ -ten Stelle. Es gilt

$$\|x\| = \left\| \sum_{k=1}^d x_k e_k \right\| \leq \sum_{k=1}^d \|x_k e_k\| = \sum_{k=1}^d |x_k| \|e_k\| \leq \|x\|_1 \cdot c,$$

wobei  $c := \max_{k=1, \dots, d} \|e_k\|$ .

Sei  $S = \{x \in E \mid \|x\|_1 = 1\}$  die Einheitssphäre in  $E$ .  $S$  ist kompakt, da  $S$  abgeschlossen und beschränkt ist.  $\|\cdot\|$  ist eine stetige Funktion auf  $S$ , denn aus  $x_k \rightarrow x$  in  $(E, \|\cdot\|_1)$  folgt

$$\| \|x_k\| - \|x\| \| \leq \|x_k - x\| \leq c \|x_k - x\|_1 \rightarrow 0,$$

also  $\|x_k\| \rightarrow \|x\|$  für  $k \rightarrow \infty$ . Damit besitzt  $\|\cdot\|$  ein Minimum  $m$  auf  $S$ ,  $m = \min_{x \in S} \|x\|$ . Es gilt  $m > 0$ , denn sonst gäbe es ein  $x \in S$  mit  $\|x\| = 0$ . Für  $x \neq 0$  ist

$$\|x\| = \left\| \|x\|_1 \frac{x}{\|x\|_1} \right\| = \|x\|_1 \underbrace{\left\| \frac{x}{\|x\|_1} \right\|}_{\in S} \geq \|x\|_1 \cdot m$$

und für  $x = 0$  stimmt die Relation trivialerweise. □

Da in der 1-Norm Folgen genau dann konvergieren, wenn sie koordinatenweise konvergieren, gilt folgendes Korollar.

**Korollar 1.2.9.** In jedem endlichdimensionalen normierten linearen Raum ist die Konvergenz von Folgen gleich der koordinatenweisen Konvergenz der Folge bezüglich einer beliebigen festen Basis des Raumes.

**Korollar 1.2.10.** Jeder endlichdimensionale normierte lineare Raum ist ein Banachraum.

*Beweis.* Sei o.B.d.A.  $E = \mathbb{R}^d$ , denn wir können ansonsten einen Isomorphismus nach  $\mathbb{R}^d$  konstruieren. Die Norm ist dann zur euklidischen Norm äquivalent, für die das Cauchy-Kriterium gilt.  $\square$

Damit haben wir einen Satz gezeigt, der uns auf vielen Räumen angenehme Eigenschaften sichert. Die meisten unserer Untersuchungen werden sich jedoch auf unendlichdimensionalen Räumen abspielen, für die die Banachraumeigenschaft für viele Normen nicht mehr gegeben sein muss. Oft wird man eine geeignete Norm suchen, die die Struktur des Raumes gut repräsentiert und die Vollständigkeit garantiert.

### 1.2.1 Die $L^p$ -Räume

Es sei  $X$  ein lokalkompakter,  $\sigma$ -kompakter topologischer (metrischer) Raum, zum Beispiel kann  $X$  eine offene oder abgeschlossene Teilmenge des  $\mathbb{R}^n$  sein.  $\mu$  sei ein positives Borelmaß auf  $X$ . Für ein Borelmaß muss insbesondere gelten:

$$\mu(K) < \infty \text{ für jedes Kompaktum } K \subseteq X.$$

Beispielsweise ist das Lebesguemaß  $\lambda$  ein Borelmaß. Wir wollen im folgenden nicht zwischen einem Maß  $\mu$  und seiner Vervollständigung unterscheiden. Sei  $1 \leq p < \infty$ . Wir definieren

$$\mathcal{L}^p := \left\{ f \mid f : X \rightarrow \mathbb{K} \text{ messbar, } \int |f(x)|^p \mu(dx) < \infty \right\}$$

sowie

$$\mathcal{L}^\infty := \{ f \mid f : x \rightarrow \mathbb{K} \text{ messbar, } \exists N \subseteq X : \mu(N) = 0 \exists M > 0 : |f(x)| \leq M \forall x \in X \setminus N \}.$$

$\mathcal{L}^\infty$  ist die Menge der wesentlich beschränkten Funktionen, also die Menge aller Funktionen, die bis auf einer Nullmenge beschränkt sind. Für  $f \in \mathcal{L}^p$  sei

$$\|f\|_p = \left( \int_X |f|^p d\mu \right)^{1/p}$$

und für  $f \in \mathcal{L}^\infty$  sei

$$\|f\|_\infty = \inf \left\{ \sup_{x \in X \setminus N} |f(x)| \mid N \subseteq X, \mu(N) = 0 \right\}.$$

Dies sind jedoch keine Normen, denn das Integral einer Funktion, die  $\mu$ -fast überall den Wert 0 annimmt, ist auch 0, ohne dass die Funktion selbst die Nullfunktion sein muss. Sei deshalb

$$\mathcal{N}_\mu := \{ f \mid f : X \rightarrow \mathbb{K}, f = 0 \text{ } \mu\text{-fast überall} \}.$$

Dann gilt  $\int f d\mu = 0$  für alle  $f \in \mathcal{N}_\mu$ . Insbesondere ist  $\|f\|_p = 0$  für diese  $f$ . Definiere weiter für  $f, g \in \mathcal{L}^p$ ,  $1 \leq p \leq \infty$ , folgende Relation:

$$f \sim g : \Leftrightarrow f - g \in \mathcal{N}_\mu,$$

das heißt  $f \sim g$  genau dann, wenn  $f = g$   $\mu$ -fast überall.  $\sim$  ist eine Äquivalenzrelation und die Äquivalenzklassen sind

$$[f] = f + \mathcal{N}_\mu,$$

auf denen  $\|\cdot\|_p$  wie folgt definiert wird:

$$\|[f]\|_p = \|f\|_p.$$

Diese Definition ist sinnvoll, da  $\|f + g\|_p = \|f\|_p$  für  $g \in \mathcal{N}_\mu$  gilt und somit  $\|f_1\|_p = \|f_2\|_p$  ist, wenn  $f_1 \sim f_2$ . Die Definition ist also repräsentantenunabhängig, der Ausdruck damit wohldefiniert.

**Satz 1.2.11** (Hölder-Ungleichung). Sei  $1 < p < \infty$  und  $q \in \mathbb{R}$  so gewählt, dass  $\frac{1}{p} + \frac{1}{q} = 1$ . Für  $f \in \mathcal{L}^p$  und  $g \in \mathcal{L}^q$  gilt

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

**Satz 1.2.12** (Minkowski-Ungleichung). Seien  $f, g \in \mathcal{L}^p$  für  $1 \leq p \leq \infty$ . Es gilt

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

Aus der Minkowski-Ungleichung folgt, dass  $\mathcal{L}^p$  ein Vektorraum ist. Außerdem ist  $\|\cdot\|_p$  eine Norm auf dem Vektorraum  $L^p$ , der wie folgt definiert ist:

$$L^p = L^p(X, \mu) = \{[f] \mid f \in \mathcal{L}^p\}.$$

Damit folgt, dass  $(L^p, \|\cdot\|_p)$  ein normierter linearer Raum ist. Man schreibt fast immer  $f$  anstelle von  $[f]$ . Ohne Beweis geben wir den folgenden wichtigen Satz.

**Satz 1.2.13** (Riesz-Fischer-Theorem).  $L^p(X, \mu)$  ist für  $p \in [1, \infty]$  vollständig, also ein Banachraum.

Der Konvergenzbegriff im Banachraum  $L^p(X, \mu)$  für  $1 \leq p < \infty$  ist ein wenig anders als vielleicht von der reellen Konvergenz aus Analysis gewohnt. Sei  $f_n, f \in \mathcal{L}^p(X, \mu)$  für  $n \in \mathbb{N}$ . Dann gilt  $\lim_{n \rightarrow \infty} f_n = f$  in  $L^p(X, \mu)$  genau dann, wenn

$$\lim_{n \rightarrow \infty} \int_X |f_n(x) - f(x)|^p \mu(dx) = 0.$$

Diese Konvergenz nennt man auch *Konvergenz im  $p$ -ten Mittel*.

**Beispiel 1.2.14.** Sei  $X = \mathbb{N}$  und  $\mu$  das Zählmaß auf  $X$ , das heißt  $\mu(M) = \#M$ . Die Funktionen auf  $X$  sind Folgen, also  $f = (f_n)_{n \in \mathbb{N}}$ . Das Integral berechnet sich dann nach  $\int f d\mu = \sum_{n=1}^{\infty} f_n$ . Also ist  $f$   $\mu$ -integrierbar genau dann, wenn  $\sum_{n=1}^{\infty} |f_n| < \infty$ . Es gibt außerdem nur tatsächliche Gleichheit von Funktionen, denn  $f = g$   $\mu$ -fast überall gilt genau dann, wenn  $f_n = g_n$  für alle  $n \in \mathbb{N}$ . Ähnliche Definitionen kann man für  $1 < p < \infty$  vornehmen. In diesem abzählbaren Spezialfall schreiben wir  $L^p(\mathbb{N}, \mu) = \ell^p(\mathbb{N}) = \ell^p$ . Es gilt dann für  $1 \leq p < \infty$ :

$$\ell^p = \left\{ (f_n)_{n \in \mathbb{N}} \mid f_n \in \mathbb{C}, \sum_{n=1}^{\infty} |f_n|^p < \infty \right\},$$

sowie

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

Zum Beispiel ist  $(\frac{1}{n}) \notin \ell^1$ , aber  $(\frac{1}{n}) \in \ell^2$ . Für  $p = \infty$  gilt

$$\ell^\infty = \{(f_n)_{n \in \mathbb{N}} \mid (f_n) \text{ ist beschränkt}\}$$

und  $\|(f_n)\|_\infty = \sup_{n \in \mathbb{N}} |f_n|$ .

Ein weiteres Beispiel ist der Raum

$$c_0 = \left\{ (f_n) \mid \lim_{n \rightarrow \infty} f_n = 0 \right\}$$

aller Nullfolgen mit der Norm  $\|f\| = \sup_{n \in \mathbb{N}} |f_n|$ .  $(c_0, \|\cdot\|)$  ist ein Banachraum. Ebenfalls kann man den Tupelraum mit  $X = \{1, \dots, d\}$  betrachten.

### 1.2.2 Die Sobolevräume

Es sei  $\Omega$  eine offene Teilmenge des  $\mathbb{R}^n$ . Um die Sobolevräume zu verstehen, müssen wir zuerst den Begriff der verallgemeinerten Ableitung einführen. Es sei

$$\mathcal{D}(\Omega) = \{f \in C^\infty(\mathbb{R}^d) \mid \exists K_f \subset\subset \Omega \text{ mit } f(x) = 0 \forall x \notin K_f\},$$

wobei wir  $K_f \subset\subset \Omega$  für  $K_f$  kompakt in  $\Omega$  enthalten schreiben. Das bedeutet, dass es eine kompakte Menge  $V$  gibt mit  $\overline{K_f} \subseteq V \subset \Omega$ .  $\mathcal{D}$  ist der Raum aller glatten (unendlich oft differenzierbaren) Funktionen mit kompaktem Träger. Er wird manchmal auch mit  $C_0^\infty$  oder  $\mathcal{C}_0^\infty$  bezeichnet. Der Träger einer Funktion  $f$  ist dabei definiert als

$$\text{supp } f = \{x \in \Omega \mid f(x) \neq 0\}.$$

Weiter definieren wir den Raum

$$L_{\text{loc}}^p(\Omega) = \left\{ f \mid f : \Omega \rightarrow \mathbb{K} \text{ messbar, } \int_K |f(x)|^p \lambda(dx) < \infty \forall K \subset\subset \Omega \right\},$$

wobei  $\lambda$  hier und im folgenden das Lebesgue-Maß auf  $\mathbb{R}^d$  bezeichnet.  $L_{\text{loc}}^p$  ist der Raum aller lokal  $L^p$ -integrierbaren Funktionen. Wir führen eine in der Mathematik häufig gebrauchte und sehr nützliche Schreibweise ein. Sei  $\alpha \in \mathbb{N}_0^d$ . Wir nennen  $\alpha$  einen Multiindex und vereinbaren  $|\alpha| = \alpha_1 + \dots + \alpha_d$ . Weiterhin setzen wir  $D^\alpha = D^{\alpha_1} D^{\alpha_2} \dots D^{\alpha_d}$ , wobei  $D^i = \frac{\partial}{\partial x_i}$  die Ableitung nach der  $i$ -ten Komponente sei. Ist nun  $f \in C^k(\overline{\Omega})$  und  $\varphi \in \mathcal{D}(\Omega)$ , dann gilt

$$\int_{\Omega} f(D^\alpha \varphi) \lambda(dx) = (-1)^{|\alpha|} \int_{\Omega} (D^\alpha f) \varphi \lambda(dx)$$

für  $\alpha \in \mathbb{N}_0^d$  mit  $|\alpha| \leq k$ .

**Definition 1.2.15.** Sei  $f \in L_{\text{loc}}^1(\Omega)$  und  $\alpha \in \mathbb{N}_0^d$ . Eine Funktion  $h \in L_{\text{loc}}^1(\Omega)$  heißt  $\alpha$ -te verallgemeinerte Ableitung von  $f$ , wenn

$$\int_{\Omega} f(x)(D^\alpha \varphi)(x) \lambda(dx) = (-1)^{|\alpha|} \int_{\Omega} h(x)\varphi(x) \lambda(dx)$$

für alle Funktionen  $\varphi \in \mathcal{D}(\Omega)$ .

Man kann zeigen (Lemma von DuBois-Reymond), dass  $h(x)$  als Funktion  $\lambda$ -fast überall eindeutig bestimmt ist. Man schreibt

$$h(x) = (D^\alpha f)(x),$$

auch wenn die gewöhnliche Ableitung nicht existiert. Falls  $f \in C^m(\Omega)$ , dann existiert die verallgemeinerte Ableitung  $D^\alpha f$  für jedes  $\alpha$  mit  $|\alpha| \leq m$  und ist gleich der gewöhnlichen Ableitung. Man schreibt für  $D^\alpha f$  im verallgemeinerten Sinn auch oft  $f^{(\alpha)}$ .

**Bemerkung 1.2.16.** Ist  $f \in L^p(\Omega, \lambda)$  mit  $1 \leq p \leq \infty$ , dann ist auch  $f \in L_{\text{loc}}^1(\Omega)$ .

Wenn  $\Omega$  eine offene oder abgeschlossene Teilmenge des  $\mathbb{R}^d$  ist, dann sei von nun an  $L^p(\Omega)$  der  $L^p(\Omega, \lambda)$  bezüglich des Lebesgue-Maßes. Mit dem Begriff der verallgemeinerten Ableitung zur Verfügung können wir nun die Sobolevräume definieren. Sie spielen unter anderem eine herausragende Rolle in der Theorie der partiellen Differentialgleichungen.

**Definition 1.2.17** (Sobolevraum). Sei  $m \in \mathbb{N}_0$  und  $p \in [1, \infty]$ .  $W_p^m(\Omega)$  sei die Menge aller Lebesguemessbaren Funktionen  $f$  auf  $\Omega$ , für die die verallgemeinerte Ableitung  $f^{(\alpha)}$  existiert und in  $L^p(\Omega)$  ist, vorausgesetzt dass  $\alpha \in \mathbb{N}_0^d$  mit  $|\alpha| \leq m$ . Definiere weiterhin

$$\|f\|_{m,p} = \sum_{|\alpha| \leq m} \left\| f^{(\alpha)} \right\|_p.$$

Der Raum  $(W_p^m(\Omega), \|\cdot\|_{m,p})$  heißt *Sobolevraum*.

**Satz 1.2.18.**  $(W_p^m(\Omega), \|\cdot\|_{m,p})$  ist ein Banachraum.

*Beweis.* Sei  $(f_n)$  eine Cauchyfolge aus  $(W_p^m(\Omega), \|\cdot\|_{m,p})$  und sei  $1 \leq p < \infty$ . Dann gilt

$$\|f_n - f_k\|_{m,p} = \sum_{|\alpha| \leq m} \left\| f_n^{(\alpha)} - f_k^{(\alpha)} \right\|_{m,p},$$

also ist  $(f_n^{(\alpha)})$  eine Cauchyfolge in  $L^p(\Omega)$ . Da  $L^p(\Omega)$  nach dem Satz von Riesz-Fischer 1.2.13 vollständig ist, existiert ein  $g_\alpha \in L^p(\Omega)$  mit

$$f_n^{(\alpha)} \rightarrow g_\alpha \text{ in } L^p(\Omega) \text{ für } n \rightarrow \infty.$$

Für  $\alpha = 0$  gilt  $f_n^{(0)} = f_n \rightarrow g_0$ . Sei  $f := g_0$ . Wir zeigen nun, dass  $g_\alpha$  die  $\alpha$ -te verallgemeinerte Ableitung von  $f$  ist. Nach Definition gilt

$$\int_{\Omega} f_n D^\alpha \varphi \lambda(dx) = (-1)^{|\alpha|} \int_{\Omega} f_n^{(\alpha)} \varphi \lambda(dx).$$

Mit  $K = \text{supp } D^\alpha \varphi$  schätzen wir ab:

$$\begin{aligned} \left| \int_{\Omega} f_n D^\alpha \varphi \, d\lambda - \int_{\Omega} f D^\alpha \varphi \, d\lambda \right| &\leq \int_{\Omega} |f_n - f| \underbrace{|D^\alpha \varphi|}_{\leq M, \text{ da } \varphi \in \mathcal{D}(\Omega)} \, d\lambda \\ &\leq M \int_K 1 \cdot |f_n - f| \, d\lambda \leq M \left( \int_K |f_n - f|^p \, d\lambda \right)^{1/p} \left( \int_K 1^q \, d\lambda \right)^{1/q} \\ &\leq M \|f_n - f\|_{L^p(\Omega)} \cdot |K|^{1/q} \rightarrow 0 \end{aligned}$$

für  $n \rightarrow \infty$ . Dabei wurde die Höldersche Ungleichung (Satz 1.2.11) benutzt. Analog folgt

$$\left| \int_{\Omega} f_n^{(\alpha)} \varphi \, d\lambda - \int_{\Omega} g_\alpha \varphi \, d\lambda \right| \rightarrow 0$$

für  $n \rightarrow \infty$ . Ein Grenzübergang ergibt insgesamt

$$\int_{\Omega} f D^\alpha \varphi \, d\lambda = (-1)^{|\alpha|} \int_{\Omega} g_\alpha \varphi \, d\lambda.$$

Es gilt  $g_\alpha \in L_{\text{loc}}^1(\Omega)$ , denn

$$\int_{K'} 1 \cdot |g_\alpha| \, d\lambda \leq \left( \int_{K'} |g_\alpha|^p \, d\lambda \right)^{1/p} \left( \int_{K'} 1^q \, d\lambda \right)^{1/q} \leq \|g_\alpha\|_{L^p(\Omega)} |K'|^{1/q} < \infty$$

für jedes Kompaktum  $K' \subset \Omega$ . Damit ist  $g_\alpha$  die  $\alpha$ -te verallgemeinerte Ableitung von  $f$ . Also gilt

$$\|f_n - f\|_{m,p} = \sum_{|\alpha| \leq m} \left\| f_n^{(\alpha)} - f^{(\alpha)} \right\|_p = \sum_{|\alpha| \leq m} \left\| f_n^{(\alpha)} - g_\alpha \right\|_p \rightarrow 0$$

für  $n \rightarrow \infty$ , das heißt  $f_n \rightarrow f$  in  $W_p^m(\Omega)$ . □

**Definition 1.2.19.** Wir definieren  $W_p^{\circ m}$  als den Abschluss der Teilmenge  $\mathcal{D}(\Omega)$  im Banachraum  $W_p^m$ .  $W_p^{\circ m}$  ist dann ebenfalls ein Banachraum.

Eine Funktion  $f \in W_p^m(\Omega)$  liegt in  $W_p^{\circ m}(\Omega)$ , wenn es eine Folge  $\varphi_n \in \mathcal{D}(\Omega)$  gibt mit  $\varphi_n \rightarrow f$  in  $(W_p^m(\Omega), \|\cdot\|_{m,p})$ . Die Sobolevräume haben folgende Eigenschaften.

**Bemerkung 1.2.20.**

- (i) Wenn  $\Omega$  beschränkt ist, dann gilt  $\mathcal{C}^m(\overline{\Omega}) \subseteq W_p^m(\Omega)$ . Im allgemeinen ist  $\mathcal{C}^m(\overline{\Omega})$  nicht dicht in  $W_p^m(\Omega)$ . Für einen hinreichend glatten Rand  $\partial\Omega$  jedoch ist  $\mathcal{C}^m(\Omega)$  dicht in  $W_p^m(\Omega)$ . Für solche Mengen kann man  $W_p^m(\Omega)$  als Vervollständigung des normierten linearen Raumes  $(\mathcal{C}^m(\overline{\Omega}), \|\cdot\|_{m,p})$  definieren.
- (ii)  $\mathcal{C}^\infty(\Omega) \cap W_p^m(\Omega)$  ist immer dicht in  $W_p^m(\Omega)$ .
- (iii) Wir können die Sobolevräume im Falle  $\Omega = \mathbb{R}^d$  mit Hilfe von Fouriertransformationen charakterisieren. Sei  $f \in L^1(\mathbb{R}^d)$ . Es sei weiter

$$\hat{f}(t) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{-i\langle t,x \rangle} f(x) \lambda(dx)$$

die Fouriertransformierte von  $f$ . Dabei sei  $\langle \cdot, \cdot \rangle$  das übliche Skalarprodukt auf dem  $\mathbb{R}^d$ .  $\hat{f}$  existiert für alle  $t \in \mathbb{R}^d$ , denn  $f \in L^1(\mathbb{R}^d)$  und es gilt

$$|\hat{f}(t)| \leq \frac{1}{(2\pi)^{d/2}} \|f\|_{L^1(\mathbb{R}^d)}.$$

wegen  $|e^{ix}| = 1$  für alle  $x \in \mathbb{R}$ . Für  $f \in L^2(\mathbb{R}^d)$  sei

$$\hat{f}(t) = \frac{1}{(2\pi)^{d/2}} \lim_{k \rightarrow \infty} \int_{\|x\| < k} e^{-i\langle t,x \rangle} f(x) \lambda(dx).$$

Dann gilt für  $m \in \mathbb{N}$ :

$$W_2^m(\mathbb{R}^d) = \left\{ f \in L^2(\mathbb{R}^d) \mid \hat{f}(t)(1 + \|t\|^2)^{m/2} \in L^2(\mathbb{R}^d) \right\}.$$

Für  $m \in \mathbb{R} \setminus \mathbb{N}$  nimmt man dies als Definition und nennt den entstehenden Raum einen *Sobolev-Slobodecki-Raum*.

Etwas genauer wollen wir uns nun mit den Sobolevräumen  $W_2^m(a, b)$  beschäftigen, wobei  $a < b$  reelle Zahlen sind.

**Definition 1.2.21.** Eine Funktion  $f(x)$  auf  $[a, b]$  heißt *absolutstetig*, wenn es eine Funktion  $g \in L^1([a, b])$  gibt mit

$$f(x) = f(a) + \int_a^x g(t) \lambda(dt)$$

für alle  $x \in [a, b]$ .

Man kann zeigen, dass  $f$   $\lambda$ -fast überall differenzierbar ist mit  $f' = g$   $\lambda$ -fast überall, wenn  $f$  absolutstetig mit  $g$  ist. Wir definieren  $AC[a, b]$  als die Menge aller absolutstetigen Funktionen auf  $[a, b]$ . Es gilt  $\mathcal{C}[a, b] \supseteq AC[a, b] \supseteq \mathcal{C}^1[a, b]$ . Wir bringen nun ohne Beweis einige für die Theorie der Sobolevräume wichtige Sätze.

**Satz 1.2.22.**

$$W_2^m(a, b) = \left\{ f \in \mathcal{C}[a, b] \mid f, \dots, f^{(m-2)} \text{ stetig differenzierbar, } f^{(m-1)} \in AC[a, b], f^{(m)} \in L^2([a, b]) \right\}.$$

**Bemerkung 1.2.23.** Aus diesem Satz folgt beispielsweise

- (i)  $W_2^1(a, b) = \{ f \in AC[a, b] \mid f^{(1)} \in L^2([a, b]) \}$ ,
- (ii)  $W_2^2(a, b) = \{ f \in \mathcal{C}^1([a, b]) \mid f' \in AC([a, b]), f'' \in L^2([a, b]) \}$ ,
- (iii)  $W_2^0(a, b) = \{ f \in W_2^1(a, b) \mid f(a) = f(b) = 0 \}$ .

**Satz 1.2.24** (Sobolev-Lemma). Seien  $d, n, p \in \mathbb{N}_0$  mit  $n > p + \frac{1}{2}$ . Sei weiter  $f \in L^1_{loc}(\Omega)$  mit

$$\frac{\partial^k}{\partial x_i^k} f \in L^2_{loc}(\Omega), \quad i = 1, \dots, d, \quad k = 0, 1, \dots, n.$$

Dann existiert eine Funktion  $f_0 \in C^p(\Omega)$  mit  $f(x) = f_0(x)$   $\lambda$ -fast überall.

**Korollar 1.2.25.** Wenn  $f \in W^p_2(\Omega)$  mit  $n > p + \frac{d}{2}$ , dann ist  $f \in C^p(\Omega)$ . Das bedeutet insbesondere  $W^p_2(\Omega) \subseteq C^p(\Omega)$ , falls  $n > p + \frac{d}{2}$ .

**Lemma 1.2.26** (Weyl-Lemma). Sei  $g \in C^\infty(\Omega)$  und  $u \in L^1_{loc}(\Omega)$ . Es gelte

$$\int_{\Omega} u(\Delta\varphi) \mu(dx) = \int_{\Omega} g\varphi \mu(dx) \quad (1.2.1)$$

für alle  $\varphi \in \mathcal{D}(\Omega)$ . Dann existiert eine Funktion  $u_1 \in C^\infty(\Omega)$  mit  $\Delta u_1 = g$  auf  $\Omega$  und  $u = u_1$   $\mu$ -fast überall.

**Bemerkung 1.2.27.** (i) Gleichung (1.2.1) heißt: Es gilt  $\Delta u = g$ , wobei  $\Delta u$  im Sinne der verallgemeinerten Ableitung definiert ist. Wenn  $\Delta u = g$  im verallgemeinerten Sinne und  $g \in C^\infty$ , folgt daraus  $u \in C^\infty$ .

(ii) Obiges Lemma ist ein Spezialfall von Sätzen über elliptische Regularität. Aus der Theorie der partiellen Differentialgleichungen folgen für Funktionen  $u$ , die Gleichungen wie  $\Delta u = 0$  oder  $u_t - \Delta u = 0$  erfüllen, sehr angenehme Eigenschaften wie Analytizität.

(iii) Für  $g = 0$  ergibt sich: Wenn  $\Delta u = 0$  im Sinne der verallgemeinerten Ableitung, dann ist  $u \in C^\infty$  und  $u$  ist im verallgemeinerten Sinn eine harmonische Funktion, für die beispielsweise das Maximumsprinzip für harmonische Funktionen gilt.

## 1.3 Hilberträume

Im folgenden sei  $E$  ein Vektorraum über  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$ .

**Definition 1.3.1** (Skalarprodukt). Ein Skalarprodukt auf  $E$  ist eine Abbildung  $\langle \cdot, \cdot \rangle : E \times E \rightarrow \mathbb{K}$  mit folgenden Eigenschaften:

(i)  $\langle \lambda_1 x_1 + \lambda_2 x_2, y \rangle = \lambda_1 \langle x_1, y \rangle + \lambda_2 \langle x_2, y \rangle$  für alle  $x_1, x_2, y \in E$  und für alle  $\lambda_1, \lambda_2 \in \mathbb{K}$ . (Linearität in der ersten Komponente)

(ii)  $\langle x, y \rangle = \overline{\langle y, x \rangle}$  für alle  $x, y \in E$ . (Hermitesität)

(iii)  $\langle x, x \rangle \geq 0$  für alle  $x \in E$  und aus  $\langle x, x \rangle = 0$  folgt  $x = 0$ . (positive Definitheit)

Aus den Eigenschaften (i) und (ii) folgt insbesondere die *Sesquilinearität* des Skalarprodukts. Das bedeutet  $\langle x + z, y \rangle = \langle x, y \rangle + \langle z, y \rangle$ ,  $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$  sowie  $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle = \langle x, \bar{\lambda} y \rangle$ .

**Definition 1.3.2** (Unitärer Raum). Ein Vektorraum  $E$  mit Skalarprodukt heißt *unitärer Raum* oder auch *Prähilbertraum*.

Wir werden nun die aus der Analysis bekannte Cauchy-Schwarz-Ungleichung für das Skalarprodukt auf allgemeinen unitären Räumen formulieren. Dabei wird lediglich die Sesquilineareigenschaft des Skalarprodukts benötigt, die Ungleichung gilt also auch für beliebige andere Sesquilinearformen.

**Satz 1.3.3** (Cauchy-Schwarz-Ungleichung). Sei  $E$  ein unitärer Raum. Wir definieren  $\|x\| = \sqrt{\langle x, x \rangle}$  für  $x \in E$ . Für alle  $x, y \in E$  gilt dann die Cauchy-Schwarz-Ungleichung:

$$|\langle x, y \rangle| \leq \|x\| \|y\|. \quad (1.3.1)$$



*Beweis.* Da  $\langle x, y \rangle = \overline{\langle y, x \rangle}$ , gibt es ein  $\alpha \in \mathbb{C}$ ,  $|\alpha| = 1$  mit  $\alpha \langle y, x \rangle = |\langle x, y \rangle|$ .

Für  $\lambda \in \mathbb{R}$  gilt:

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

Letzteren Ausdruck erkennt man als quadratische Gleichung mit Unbekannter  $\lambda$ , also folgt aus der Ungleichung, dass

$$\frac{4 |\langle x, y \rangle|^2}{4 \|y\|^4} - \frac{\|x\|^2}{\|y\|^2} \leq 0$$

sein muss und hieraus folgt die Behauptung.  $\square$

In der Cauchy-Schwarz-Ungleichung haben wir  $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$  definiert und dabei das Symbol für eine Norm benutzt. Dass es sich tatsächlich um eine Norm handelt, zeigt das folgende Korollar.

**Korollar 1.3.4.** Die durch  $\|x\| := \sqrt{\langle x, x \rangle}$  für alle  $x \in E$  definierte Norm ist eine Norm auf dem Vektorraum  $E$ .

*Beweis.* Wir weisen die Normeigenschaften einzeln nach. Zuerst die Homogenität:

$$\|\lambda x\| = \sqrt{\langle \lambda x, \lambda x \rangle} = \sqrt{\lambda \overline{\lambda} \langle x, x \rangle} = |\lambda| \|x\|.$$

Als nächstes beweisen wir die Dreiecksungleichung. Dafür benötigen wir die eben bewiesene Cauchy-Schwarz-Ungleichung (1.3.1), die wir wegen  $c + \bar{c} = 2 \operatorname{Re}(c) \leq 2|c|$  für eine komplexe Zahl  $c$  wie folgt anwenden:

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

Durch Wurzelziehen folgt die Dreiecksungleichung. Als letztes fehlt noch die Definitheit. Sei  $\|x\| = 0$ . Dann folgt  $\langle x, x \rangle = 0$ . Nach dem Axiom aus Definition 1.3.1 (iii) folgt  $x = 0$ . Also ist durch  $\|x\| := \sqrt{\langle x, x \rangle}$  tatsächlich eine Norm auf  $E$  gegeben.  $\square$

**Definition 1.3.5** (Hilbertraum). Ein unitärer Raum  $(E, \langle \cdot, \cdot \rangle)$ , der bezüglich seiner induzierten Norm  $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$  vollständig ist, heißt *Hilbertraum*.

Wir werden ab sofort jeden Hilbertraum als normierten Raum auffassen, indem wir die induzierte Norm verwenden.

**Definition 1.3.6** (orthogonal). Sei  $(E, \langle \cdot, \cdot \rangle)$  ein unitärer Raum und  $x, y \in E$ . Wir schreiben  $x \perp y$ , wenn  $\langle x, y \rangle = 0$ . Man sagt dann, dass  $x$  und  $y$  *orthogonal* seien oder senkrecht aufeinander stehen.

Folgende, leicht nachzurechnende Gleichungen zeigen den Zusammenhang zwischen Skalarprodukt und zugehöriger Norm.  $(E, \langle \cdot, \cdot \rangle)$  sei ein unitärer Raum und  $\|x\| = \sqrt{\langle x, x \rangle}$  für alle  $x \in E$ . Dann gilt die sogenannte *Polarisationsidentität*:

$$\langle x, y \rangle = \frac{1}{4} (\|x + y\|^2 - \|x - y\|^2) \tag{1.3.2}$$

im Fall  $\mathbb{K} = \mathbb{R}$ , beziehungsweise

$$\langle x, y \rangle = \frac{1}{4} (\|x + y\|^2 - \|x - y\|^2 + i \|x + iy\|^2 - i \|x - iy\|^2) \tag{1.3.3}$$

für den Fall  $\mathbb{K} = \mathbb{C}$  eines komplexen unitären Raumes. Umgekehrt funktioniert dies nicht so einfach, das heißt wir bekommen nicht aus jeder Norm automatisch ein Skalarprodukt. Man kann jedoch eine notwendige und hinreichende Bedingung dafür angeben, dass eine Norm auf einem Vektorraum ein eindeutig bestimmtes Skalarprodukt induziert.

**Satz 1.3.7** (Satz von John von Neumann).  $\|\cdot\|$  sei eine Norm auf dem Vektorraum  $E$  über  $K = \mathbb{R}$  oder  $K = \mathbb{C}$ . Die Norm  $\|\cdot\|$  wird von einem Skalarprodukt  $\langle \cdot, \cdot \rangle$  erzeugt genau dann, wenn die Parallelogrammidentität

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2) \quad (1.3.4)$$

für  $x, y \in E$  gilt. Ist diese erfüllt, dann ist das Skalarprodukt durch die Norm eindeutig bestimmt und durch die entsprechende Formel in Gleichung (1.3.2) oder (1.3.3) gegeben.

*Beweis.* Wir beweisen nur den reellen Fall, also  $\mathbb{K} = \mathbb{R}$ . Wenn  $\|x\| = \sqrt{\langle x, x \rangle}$  für  $x \in E$ , wenn also die Norm durch ein Skalarprodukt gegeben ist, dann folgt die Parallelogrammidentität (1.3.4) durch

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

Sei umgekehrt die Parallelogrammidentität erfüllt. Wir weisen die Skalarprodukteigenschaften der durch (1.3.2) definierten Abbildung nach. Zunächst gilt

$$\begin{aligned} \|x_1 + x_2 + y\|^2 &= \frac{1}{2} \|x_1 + x_2 + y\|^2 + \frac{1}{2} \|x_1 + x_2 + y\|^2 \\ &= \frac{1}{2} (2\|x_1 + y\|^2 + 2\|x_2\|^2 - \|x_1 - x_2 + y\|^2 + 2\|x_2 + y\|^2 + 2\|x_1\|^2 - \|-x_1 + x_2 + y\|^2) \\ &= \|x_1 + y\|^2 + \|x_2 + y\|^2 + \|x_1\|^2 + \|x_2\|^2 - \frac{1}{2} (\|x_1 - x_2 + y\|^2 + \|-x_1 + x_2 + y\|^2) \end{aligned}$$

Ersetze in obiger Gleichung  $y$  durch  $-y$ . Damit haben wir ebenfalls

$$\|x_1 + x_2 - y\|^2 = \|x_1 - y\|^2 + \|x_2 - y\|^2 + \|x_1\|^2 + \|x_2\|^2 - \frac{1}{2} (\|x_1 - x_2 - y\|^2 + \|-x_1 + x_2 - y\|^2)$$

Nach Gleichung (1.3.2) gilt durch Einsetzen der beiden gezeigten Gleichungen:

$$\begin{aligned} \langle x_1 + x_2, y \rangle &= \frac{1}{4} (\|x_1 + x_2 + y\|^2 - \|x_1 + x_2 - y\|^2) \\ &= \frac{1}{2} (\|x_1 + y\|^2 - \|x_1 - y\|^2 + \|x_2 + y\|^2 - \|x_2 - y\|^2) = \langle x_1, y \rangle + \langle x_2, y \rangle, \end{aligned}$$

also folgt hieraus die Additivität in der 1. Komponente. Damit haben wir insbesondere

$$\langle 2x, y \rangle = \langle x, y \rangle + \langle x, y \rangle = 2\langle x, y \rangle$$

und mittels Induktion folgt  $\langle nx, y \rangle = n\langle x, y \rangle$  für  $n \in \mathbb{N}$ , also die Homogenität für natürliche Zahlen. Sei nun  $\lambda = \frac{m}{n} \in \mathbb{Q}$ . Es ergibt sich

$$n\langle \lambda x, y \rangle = \langle n \frac{m}{n} x, y \rangle = \langle mx, y \rangle = m\langle x, y \rangle$$

und hieraus folgt  $\langle \lambda x, y \rangle = \frac{m}{n} \langle x, y \rangle$ . Wegen der Stetigkeit der Norm und (1.3.2) sind die Funktionen  $\langle \lambda x, y \rangle$  und  $\lambda \langle x, y \rangle$  stetig. Für  $\lambda \in \mathbb{Q}$  sind beide Funktionen sogar gleich und damit folgt  $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle$  für alle  $\lambda \in \mathbb{R}$  durch Approximation von reellen durch rationale Zahlen. Also haben wir die Linearität in der ersten Komponente gezeigt. Die anderen Eigenschaften, Definitheit und Symmetrie, sind leicht nachzurechnen.  $\square$

**Beispiel 1.3.8** (Hilberträume).

- (a)  $(E, \langle \cdot, \cdot \rangle)$  mit  $E = \mathbb{C}^n$  und  $\langle x, y \rangle = \sum_{i=1}^n x_i \overline{y_i}$  ist ein Hilbertraum.
- (b) Sei  $E = L^2(X, \mu)$  mit einem Maßraum  $(X, \mu)$ . Definiere  $\langle f, g \rangle = \int_X f(x) \overline{g(x)} \mu(dx)$ . Damit sind  $L^2[a, b]$  und  $\ell^2$  (für  $X = \mathbb{N}$ ) Hilberträume. Man nennt  $\ell^2$  den *Hilbertschen Folgenraum*.
- (c) Die Sobolev-Räume  $W_2^m(\Omega)$ ,  $m \in \mathbb{N}$ , mit dem Skalarprodukt

$$\langle f, g \rangle = \sum_{|\alpha| \leq m} \langle f^{(\alpha)}, g^{(\alpha)} \rangle_{L^2(\Omega, \mu)}$$

sind Hilberträume. Die zugehörige Norm ist  $\|f\| = \sqrt{\langle f, f \rangle}$ . Wir hatten weiter oben bereits eine andere Norm definiert durch  $\|f\|_{m,2} = \sum_{|\alpha| \leq m} \|f^{(\alpha)}\|_{L^2(\Omega, \mu)}$ . Diese beiden Normen,  $\|\cdot\|$  und  $\|\cdot\|_{m,2}$ , sind äquivalent.

## 1.4 Die Rieszschen Sätze

Zuerst beschäftigen wir uns mit der direkten orthogonalen Summe von Hilberträumen. Seien  $(H_1, \langle \cdot, \cdot \rangle_1)$  und  $(H_2, \langle \cdot, \cdot \rangle_2)$  zwei Hilberträume und sei  $H$  die direkte Summe der linearen Räume  $H_1$  und  $H_2$ , also  $H = H_1 + H_2 = \{(x, y) \mid x \in H_1, y \in H_2\}$ . Wir definieren

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

für  $x_1, y_1 \in H_1$  und  $x_2, y_2 \in H_2$ . Dann ist  $\langle \cdot, \cdot \rangle$  ein Skalarprodukt auf  $H$  und  $(H, \langle \cdot, \cdot \rangle)$  ein Hilbertraum. Bevor wir mit dieser Konstruktion fortfahren können, benötigen wir zunächst eine formale Definition von Isomorphismen für Hilberträume.

**Definition 1.4.1.** Ein *Isomorphismus* eines Hilbertraums  $(G, \langle \cdot, \cdot \rangle_G)$  auf einen Hilbertraum  $(K, \langle \cdot, \cdot \rangle_K)$  ist eine bijektive lineare Abbildung  $f$  von  $G$  auf  $K$ , für die

$$\langle f(x), f(y) \rangle_K = \langle x, y \rangle_G$$

für alle  $x, y \in G$  gilt.

Sei  $\tilde{H}_1 = \{(x_1, 0) \mid x_1 \in H_1\}$  und  $\tilde{H}_2 = \{(0, x_2) \mid x_2 \in H_2\}$ . Dann sind  $\tilde{H}_1$  und  $\tilde{H}_2$  lineare Teilmengen von  $H$ . Wir definieren

$$f_1(x_1) = (x_1, 0), \quad f_2(x_2) = (0, x_2), \quad x_1 \in H_1, x_2 \in H_2.$$

$f_1$  ist ein Isomorphismus von  $H_1$  auf  $\tilde{H}_1$  und  $f_2$  ein Isomorphismus von  $H_2$  auf  $\tilde{H}_2$ . Wir identifizieren nun aufgrund dieser Isomorphismen kurzerhand  $x_1$  mit  $(x_1, 0)$  und  $x_2$  mit  $(0, x_2)$ , wenn  $x_1 \in H_1$  und  $x_2 \in H_2$ . Mit diesen Vorbereitungen wollen wir nun orthogonale Summen von Hilberträumen definieren. Sei dazu  $(x_1, x_2) \in H$ . Dann ist nach den vorangegangenen Überlegungen  $(x_1, x_2) = (x_1, 0) + (0, x_2) = x_1 + x_2$ . Für diese beliebig gewählten  $x_1$  und  $x_2$  gilt dann

$$\langle (x_1, 0), (0, x_2) \rangle_H = \langle x_1, 0 \rangle_{H_1} + \langle 0, x_2 \rangle_{H_2} = 0,$$

das heißt alle Vektoren aus  $H_1$  sind orthogonal zu allen Vektoren aus  $H_2$ . Man sagt, dass die linearen Teilräume  $H_1$  und  $H_2$  orthogonal sind und schreibt  $H_1 \perp H_2$ .  $H$  heißt dann die *direkte orthogonale Summe* der Hilberträume  $H_1$  und  $H_2$ . Man schreibt dafür

$$H = H_1 \oplus H_2.$$

Nun definieren wir die abzählbare direkte orthogonale Summe von Hilberträumen. Für  $n \in \mathbb{N}$  sei jeweils  $(H_n, \langle \cdot, \cdot \rangle_n)$  ein Hilbertraum. Setze

$$H := \left\{ (x_n)_{n \in \mathbb{N}} \mid \sum_{n=1}^{\infty} \|x_n\|_{H_n}^2 < \infty, x_n \in H_n \text{ für } n \in \mathbb{N} \right\}.$$

Aus der Cauchy-Schwarz-Ungleichung (Satz 1.3.3) folgt, dass  $H$  ein Vektorraum ist. Wir definieren weiter

$$\langle (x_n), (y_n) \rangle_H = \sum_{n=1}^{\infty} \langle x_n, y_n \rangle_{H_n},$$

wobei  $(x_n), (y_n) \in H$ . Man zeigt, dass  $(H, \langle \cdot, \cdot \rangle)$  ein Hilbertraum ist und schreibt

$$H = \bigoplus_{n=1}^{\infty} H_n.$$

$H$  heißt die direkte orthogonale Summe der Hilberträume  $H_n$  für  $n \in \mathbb{N}$ .

**Beispiel 1.4.2.** Ist  $H_1 = L^2([a, b], \lambda)$  und  $H_2 = L^2([b, c], \lambda)$ , dann ist  $H = H_1 \oplus H_2 \cong L^2([a, c], \lambda)$ , wie man sich leicht überlegt.

**Definition 1.4.3.** Ein abgeschlossener linearer Teilraum eines Hilbertraumes heißt *Unterraum*.

Im folgenden sei  $(H, \langle \cdot, \cdot \rangle)$  ein Hilbertraum.

**Definition 1.4.4.**  $H_1$  sei ein linearer Teilraum von  $H$ . Es sei

$$H_1^\perp = \{y \in H \mid y \perp x \text{ für alle } x \in H_1\}.$$

$H_1^\perp$  heißt *orthogonales Komplement* von  $H_1$  in  $H$ .

Wir werden zeigen, dass ein Hilbertraum sich immer aus einem Unterraum und dessen orthogonalem Komplement zusammensetzen lässt. Dies ist die Aussage des 1. Satzes von Riesz. Bevor wir diesen beweisen können, benötigen wir jedoch folgendes Lemma.

**Lemma 1.4.5.** Sei  $H_1$  eine abgeschlossene konvexe Menge in  $H$ . Für  $x \in H$  sei

$$\rho(x) = \inf_{y \in H_1} \|x - y\|.$$

Dann existiert ein eindeutig bestimmtes Element  $x_1 \in H_1$  mit  $\rho(x) = \|x - x_1\|$ .

*Beweis.* (Existenz): Es existiert eine Folge  $(y_n)_{n \in \mathbb{N}}$  mit  $y_n \in H_1$ , so dass  $\rho(x) = \lim_{n \rightarrow \infty} \|x - y_n\|$ .  $(y_n)$  ist dann eine Cauchyfolge, denn nach der Parallelogrammidentität 1.3.4 gilt

$$\begin{aligned} \|y_n - y_m\|^2 &= \|y_n - x + x - y_m\|^2 = 2\|y_n - x\|^2 + 2\|x - y_m\|^2 - \|2x - y_n - y_m\|^2 \\ &= 2\|y_n - x\|^2 + 2\|x - y_m\|^2 - 4\left\|x - \frac{y_n + y_m}{2}\right\|^2 \\ &\leq 2\|y_n - x\|^2 + 2\|y_m - x\|^2 - 4\rho(x)^2 \longrightarrow 0 \end{aligned}$$

für  $m, n \rightarrow \infty$ . Die letzte Ungleichung erhält man, da  $\frac{y_n + y_m}{2} \in H_1$  aufgrund der Konvexität von  $H_1$  gilt. Wegen der Vollständigkeit von  $H$  konvergiert  $(y_n)$ , das heißt es existiert ein  $x_1 \in H$  mit  $y_n \rightarrow x_1$ . Da  $H_1$  abgeschlossen ist, gilt sogar  $x_1 \in H_1$ . Es folgt

$$\rho(x) = \lim_{n \rightarrow \infty} \|x - y_n\| = \left\| \lim_{n \rightarrow \infty} (x - y_n) \right\| = \|x - x_1\|$$

aufgrund der Stetigkeit von  $\|\cdot\|$ .

(Eindeutigkeit): Angenommen,  $x_1, x'_1 \in H_1$  erfüllen  $\rho(x) = \|x - x_1\| = \|x - x'_1\|$ . Dann haben wir

$$\begin{aligned} 0 \leq \|x_1 - x'_1\|^2 &= \|x_1 - x + x - x'_1\|^2 = 2\|x_1 - x\|^2 + 2\|x - x'_1\|^2 - 4\left\|x - \frac{x_1 + x'_1}{2}\right\|^2 \\ &\leq 2\|x_1 - x\|^2 + 2\|x - x'_1\|^2 - 4\rho(x)^2 = 0 \end{aligned}$$

und damit gilt  $x_1 = x'_1$ . □

**Satz 1.4.6** (1. Satz von Riesz).  $H_1$  sei ein Unterraum eines Hilbertraums  $H$ . Dann ist  $H_1^\perp$  auch ein Unterraum von  $H$  und es gilt  $H = H_1 \oplus H_1^\perp$ . Das heißt, dass zu jedem  $x \in H$  eindeutig bestimmte Elemente  $x_1 \in H_1$  und  $x_2 \in H_1^\perp$  existieren, so dass  $x = x_1 + x_2$  gilt.

*Beweis.* Sei  $x \in H$ . Da  $H_1$  ein linearer Teilraum von  $H$  und damit konvex ist, können wir  $x_1 \in H_1$  nach Lemma 1.4.5 wählen, so dass  $\rho(x) = \|x - x_1\|$  ist. Dann ist

$$\|x - x_1\| = \rho(x) = \inf_{y \in H_1} \|x - y\| \leq \|x - x_1 - ty_1\|$$

für alle  $t \in \mathbb{C}$  und festes  $y_1 \in H_1$ . Nach einer Übungsaufgabe gilt deshalb  $\langle x - x_1, y_1 \rangle = 0$ , also  $x - x_1 =: x_2 \in H_1^\perp$ . Es folgt  $x = x_1 + x_2$  mit  $x_1 \in H_1$  und  $x_2 \in H_1^\perp$ . Angenommen,  $x = x_1 + x_2 = x'_1 + x'_2$  mit  $x'_1 \in H_1$  und  $x'_2 \in H_1^\perp$ . Dann gilt  $\langle z, z \rangle = 0$  für  $z := x_1 - x'_1 = x_2 - x'_2$ , da  $x_1 - x'_1 \in H_1$  und  $x_2 - x'_2 \in H_1^\perp$ . Daraus folgt  $z = 0$ . □

Für  $x \in H$  sei  $x = x_1 + x_2$  die Zerlegung nach dem 1. Rieszschen Satz. Wir definieren die *Projektionsoperatoren*  $P_{H_1}$  und  $P_{H_1^\perp}$  durch

$$P_{H_1}(x) = x_1, \quad P_{H_1^\perp}(x) = x_2.$$

Auf Projektionsoperatoren kommen wir in Abschnitt 3.2 zurück. Im folgenden kommen wir zu der Definition, die der Funktionalanalysis in ihren frühen Tagen ihren Namen gegeben hat. Die Untersuchung von Funktionalen, das heißt Abbildungen von normierten Vektorräumen in ihre Grundkörper, wird in den folgenden Abschnitten eine grundlegende Rolle spielen. Sei also  $(E, \|\cdot\|)$  ein normierter linearer Raum über  $\mathbb{K} = \mathbb{C}$  oder  $\mathbb{K} = \mathbb{R}$ .

**Definition 1.4.7.** Ein *lineares Funktional* auf  $(E, \|\cdot\|)$  ist eine lineare Abbildung  $\mathcal{F} : E \rightarrow \mathbb{K}$ , das heißt

$$\mathcal{F}(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 \mathcal{F}(x_1) + \lambda_2 \mathcal{F}(x_2)$$

für  $\lambda_1, \lambda_2 \in \mathbb{K}$  und  $x_1, x_2 \in E$ .  $\mathcal{F}$  heißt *stetig*, wenn aus  $x_n \rightarrow x$  in  $(E, \|\cdot\|)$  stets  $\mathcal{F}(x_n) \rightarrow \mathcal{F}(x)$  folgt.

$E'$  sei die Menge aller stetigen linearen Funktionale auf  $(E, \|\cdot\|)$ . Dann ist  $E'$  ein  $\mathbb{K}$ -Vektorraum. Wir wollen uns damit beschäftigen, wie man zu einem Hilbertraum  $(H, \langle \cdot, \cdot \rangle)$  den Raum  $H'$  bestimmt. Sei also  $(H, \langle \cdot, \cdot \rangle)$  ein Hilbertraum und  $y \in H$ . Wir setzen  $\mathcal{F}_y(x) := \langle x, y \rangle$ . Dann ist  $\mathcal{F}_y$  linear, denn  $\mathcal{F}_y(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 \langle x_1, y \rangle + \lambda_2 \langle x_2, y \rangle = \lambda_1 \mathcal{F}_y(x_1) + \lambda_2 \mathcal{F}_y(x_2)$ . Dies wollen wir präzisieren.

**Satz 1.4.8** (2. Satz von Riesz). Zu jedem stetigen linearen Funktional  $\mathcal{F}$  auf  $H$  existiert genau ein  $y \in H$ , so dass  $\mathcal{F}(x) = \mathcal{F}_y(x)$  für alle  $x \in H$ .

*Beweis.* (Existenz): Sei  $H_1$  der Nullraum von  $\mathcal{F}$ , also  $H_1 = \{x \in H \mid \mathcal{F}(x) = 0\}$ . Da  $\mathcal{F}$  linear ist, ist  $H_1$  ein linearer Teilraum von  $H$ . Außerdem ist  $H_1$  abgeschlossen, da  $H_1 = \mathcal{F}^{-1}(\{0\})$  das Urbild einer abgeschlossenen Menge ist. Nach dem ersten Satz von Riesz (Satz 1.4.6) ist  $H = H_1 \oplus H_1^\perp$ . Wir machen

eine Fallunterscheidung. Sei zuerst angenommen, dass  $H_1^\perp = \{0\}$ , das heißt  $\mathcal{F}(x) = 0$  für alle  $x \in H$ . Setze dann  $y = 0$ , denn dann genügt  $\mathcal{F}_y(x) = \langle x, 0 \rangle = 0$  der Bedingung. Ist hingegen  $H_1^\perp \neq \{0\}$ , dann sei  $u \in H_1^\perp$ ,  $u \neq 0$ . Da  $H_1 \cap H_1^\perp = \{0\}$ , gilt  $\mathcal{F}(u) \neq 0$ . Also ist

$$\mathcal{F}\left(x - \frac{\mathcal{F}(x)}{\mathcal{F}(u)}u\right) = \mathcal{F}(x) - \frac{\mathcal{F}(x)}{\mathcal{F}(u)}\mathcal{F}(u) = 0$$

und damit  $x - \frac{\mathcal{F}(x)}{\mathcal{F}(u)}u \in H_1$ . Da  $u \in H_1^\perp$ , gilt

$$\left\langle x - \frac{\mathcal{F}(x)}{\mathcal{F}(u)}u, u \right\rangle = 0,$$

was äquivalent ist zu  $\langle x, u \rangle = \frac{\mathcal{F}(x)}{\mathcal{F}(u)}\langle u, u \rangle$ . Setze nun  $y := \frac{\mathcal{F}(u)}{\langle u, u \rangle}u$ . Dann gilt für alle  $x \in H$ :

$$\mathcal{F}(x) = \frac{\mathcal{F}(u)}{\langle u, u \rangle}\langle x, u \rangle = \left\langle x, \frac{\mathcal{F}(u)}{\langle u, u \rangle}u \right\rangle = \mathcal{F}_y(x).$$

(Eindeutigkeit): Seien  $y_1, y_2 \in H$ , so dass  $\mathcal{F}(x) = \mathcal{F}_{y_1}(x) = \mathcal{F}_{y_2}(x)$  für alle  $x \in H$ . Dann folgt  $\langle x, y_1 \rangle = \langle x, y_2 \rangle$  und damit  $\langle x, y_1 - y_2 \rangle = 0$  für alle  $x \in H$ , insbesondere für  $x = y_1 - y_2$ . Also gilt  $\langle y_1 - y_2, y_1 - y_2 \rangle = \|y_1 - y_2\|^2 = 0$ , was  $y_1 = y_2$  impliziert.  $\square$

### Beispiel 1.4.9.

- (a) Sei  $(H, \langle \cdot, \cdot \rangle) = L^2(X, \mu)$  mit dem Skalarprodukt  $\langle f, g \rangle = \int_X f(x)\overline{g(x)}\mu(dx)$ . Dann gibt es für alle  $\mathcal{F} \in H'$  ein  $g \in L^2(X, \mu)$  mit  $\mathcal{F}(f) = \int_X f(x)\overline{g(x)}\mu(dx)$ .
- (b) Sei  $(H, \langle \cdot, \cdot \rangle) = \ell^2$  und  $\mathcal{F} \in H'$ . Dann gibt es ein  $(y_n)_{n \in \mathbb{N}} \in H$  mit  $\mathcal{F}((x_n)) = \sum_{k=1}^{\infty} x_k \overline{y_k}$ .

## 1.5 Orthonormalsysteme und Fourierreentwicklung im Hilbertraum

Im  $\mathbb{R}^n$  gibt es die Standard- oder kanonische Basis  $x_k = (0, \dots, 0, 1, 0, \dots, 0)$  mit jeweils einem Eintrag an der  $k$ -ten Stelle. Für alle  $y \in \mathbb{R}^n$  gibt es dann  $\alpha_1, \dots, \alpha_n$  mit  $y = \alpha_1 x_1 + \dots + \alpha_n x_n$  und man kann die Koeffizienten durch  $\alpha_n = \langle x_n, y \rangle$  bestimmen. Die Darstellung von Elementen durch eine Linearkombination von normierten, orthogonalen Basisvektoren wollen wir im folgenden auch auf uneigentliche Linearkombinationen, also unendliche Summen, verallgemeinern.

**Definition 1.5.1** (Orthogonalsystem, Orthonormalsystem). Sei  $(H, \langle \cdot, \cdot \rangle)$  ein Hilbertraum. Eine Menge  $O = \{x_i \mid i \in I\}$  von Elementen aus  $H$  zu einer Indexmenge  $I$  heißt *Orthogonalsystem*, wenn  $\langle x_i, x_j \rangle = 0$  für  $i \neq j$  gilt.  $O$  heißt *Orthonormalsystem*, wenn  $\langle x_i, x_j \rangle = \delta_{ij}$  für alle  $i, j \in I$  gilt, wenn die  $x_i$  also normiert sind.

### Beispiel 1.5.2.

- (a) Sei  $(H, \langle \cdot, \cdot \rangle) = \ell^2$ . Die kanonischen Folgen  $x_k = (0, \dots, 0, 1, 0, \dots)$  mit einem Eintrag an der  $k$ -ten Stelle bilden ein Orthonormalsystem.
- (b) Sei  $(H, \langle \cdot, \cdot \rangle) = L^2[-\pi, \pi]$ . Die Menge  $\{1, \sin nx, \cos nx \mid n \in \mathbb{N}\}$  bildet ein Orthogonalsystem. Es gilt die Beziehung

$$\langle \cos nx, \cos nx \rangle = \langle \sin nx, \sin nx \rangle = \pi.$$

Normieren wir die Vektoren, dann bildet die Menge  $\left\{ \frac{1}{\sqrt{2\pi}}, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin nx}{\sqrt{\pi}} \mid n \in \mathbb{N} \right\}$  ein Orthonormalsystem.

**Lemma 1.5.3** (Satz des Pythagoras). Sei  $H$  ein Hilbertraum und  $x_1, \dots, x_n$  ein Orthogonalsystem. Dann gilt

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

*Beweis.* Der Beweis ist einfach.

$$\begin{aligned} \|x_1 + \dots + x_n\|^2 &= \langle x_1 + \dots + x_n, x_1 + \dots + x_n \rangle = \langle x_1, x_1 \rangle + \dots + \langle x_1, x_n \rangle + \langle x_2, x_1 \rangle + \dots + \langle x_n, x_n \rangle \\ &= \langle x_1, x_1 \rangle + \dots + \langle x_n, x_n \rangle = \|x_1\|^2 + \dots + \|x_n\|^2, \end{aligned}$$

denn wegen der Orthonormalität ist  $\langle x_i, x_j \rangle = \delta_{ij}$ . □

**Satz 1.5.4.** Sei  $x_1, x_2, \dots$  ein Orthogonalsystem im Hilbertraum  $H$ . Dann gilt

$$\sum_{k=1}^{\infty} x_k \text{ konvergiert in } H \iff \sum_{k=1}^{\infty} \|x_k\| < \infty,$$

das heißt die Partialsummenfolge  $s_m := \sum_{k=1}^m x_k$  konvergiert genau dann bezüglich  $\|\cdot\|$ , wenn die Reihe der Normen konvergiert.

*Beweis.* Es gilt für  $m > n$  nach dem Satz des Pythagoras (Lemma 1.5.3):

$$\|s_m - s_n\|^2 = \left\| \sum_{i=n+1}^m x_i \right\|^2 = \sum_{i=n+1}^m \|x_i\|^2 = \tilde{s}_m - \tilde{s}_n,$$

wobei wir  $\tilde{s}_m = \sum_{i=1}^m \|x_i\|^2$  gesetzt haben. Also ist  $(s_k)_{k \in \mathbb{N}}$  eine Cauchyfolge in  $H$  genau dann, wenn  $(\tilde{s}_k)_{k \in \mathbb{N}}$  eine Cauchyfolge in  $\mathbb{R}$  ist. □

Wir kommen nun zur Fourierentwicklung von Elementen in einem Hilbertraum  $H$ . Sei dafür  $\{x_n \mid n \in \mathbb{N}\}$  ein Orthonormalsystem in  $H$ .

**Satz 1.5.5** (Besselsche Ungleichung). Für jedes  $x \in H$  gilt

$$\sum_{k=1}^{\infty} |\langle x, x_k \rangle|^2 \leq \|x\|^2. \tag{1.5.1}$$

*Beweis.* Sei  $n \in \mathbb{N}$  und  $y_n := x - \sum_{k=1}^n \langle x, x_k \rangle x_k$ . Es gilt für  $m \leq n$ :

$$\langle y_n, x_m \rangle = \langle x, x_m \rangle - \sum_{k=1}^n \langle x, x_k \rangle \langle x_k, x_m \rangle = \langle x, x_m \rangle - \langle x, x_m \rangle = 0.$$

Damit ist also  $\{y_n, \langle x, x_1 \rangle x_1, \dots, \langle x, x_n \rangle x_n\}$  ein Orthogonalsystem. Nach dem Satz des Pythagoras (Lemma 1.5.3) gilt

$$\|x\|^2 = \left\| y_n + \sum_{k=1}^n \langle x, x_k \rangle x_k \right\|^2 = \|y_n\|^2 + \sum_{k=1}^n |\langle x, x_k \rangle|^2 \|x_k\|^2 \geq \sum_{k=1}^n |\langle x, x_k \rangle|^2$$

wegen  $\|x_k\| = 1$  für alle  $k \in \mathbb{N}$ . Nun folgt der Grenzübergang  $n \rightarrow \infty$  und damit die Aussage. □

**Korollar 1.5.6.** Für alle  $x \in H$  konvergiert

$$\sum_{k=1}^{\infty} \langle x, x_k \rangle x_k$$

in  $H$ .

*Beweis.* Nach Satz 1.5.4 ist die Konvergenz der Reihe in  $H$  äquivalent zur Konvergenz von

$$\sum_{k=1}^{\infty} |\langle x, x_k \rangle|^2 \|x_k\|^2 \leq \|x\|^2,$$

wobei die Besselsche Ungleichung (1.5.1) benutzt wurde.  $\square$

**Satz 1.5.7.** Für ein Orthonormalsystem  $\{x_n \mid n \in \mathbb{N}\}$  sind folgende Aussagen äquivalent:

- (i)  $x = \sum_{k=1}^{\infty} \langle x, x_k \rangle x_k$  für alle  $x \in H$ .
- (ii) Aus  $\langle z, x_k \rangle = 0$  für alle  $k \in \mathbb{N}$  folgt  $z = 0$ .
- (iii)  $\|x\|^2 = \sum_{k=1}^{\infty} |\langle x, x_k \rangle|^2$  für alle  $x \in H$ . (Parsevalsche Gleichung)
- (iv)  $\langle x, y \rangle = \sum_{k=1}^{\infty} \langle x, x_k \rangle \langle x_k, y \rangle$  für alle  $x, y \in H$ .

*Beweis.* (i)  $\Rightarrow$  (iv): Es gilt

$$\langle x, y \rangle = \left\langle \sum_{k=1}^{\infty} \langle x, x_k \rangle x_k, \sum_{l=1}^{\infty} \langle y, x_l \rangle x_l \right\rangle = \sum_{k,l=1}^{\infty} \langle x, x_k \rangle \overline{\langle y, x_l \rangle} \langle x_k, x_l \rangle = \sum_{k=1}^{\infty} \langle x, x_k \rangle \langle x_k, y \rangle.$$

(iv)  $\Rightarrow$  (iii): Setze  $y = x$ .

(iii)  $\Rightarrow$  (ii): Sei  $\langle z, x_k \rangle = 0$  für alle  $k \in \mathbb{N}$ . Es gilt  $\|z\|^2 = \sum_{k=1}^{\infty} |\langle z, x_k \rangle|^2 = 0$ , also  $z = 0$ .

(ii)  $\Rightarrow$  (i): Sei  $x \in H$ . Setze dann  $y := \sum_{k=1}^{\infty} \langle x, x_k \rangle x_k$  und  $z = x - y$ . Dann gilt für  $l \in \mathbb{N}$ :

$$\langle z, x_l \rangle = \langle x - y, x_l \rangle = \left\langle x - \sum_{k=1}^{\infty} \langle x, x_k \rangle x_k, x_l \right\rangle = \langle x, x_l \rangle - \sum_{k=1}^{\infty} \langle x, x_k \rangle \langle x_k, x_l \rangle = 0$$

und damit  $x = y$ .  $\square$

**Definition 1.5.8** (Vollständigkeit). Ein Orthonormalsystem mit einer dieser vier äquivalenten Eigenschaften heißt *vollständiges normiertes Orthogonalsystem*, *vollständiges Orthonormalsystem* oder auch *Orthonormalbasis* von  $H$ .

**Beispiel 1.5.9.** Sei  $(H, \langle \cdot, \cdot \rangle) = \ell^2$ . Die kanonischen Folgen  $x_n = (0, \dots, 0, 1, 0, \dots)$  mit einem Eintrag an der  $n$ -ten Stelle bilden ein Orthonormalsystem  $O = \{x_n \mid n \in \mathbb{N}\}$ . Sei  $z \in \ell^2$  mit  $z = (z_n)_{n \in \mathbb{N}} \perp x_k$  für alle  $k \in \mathbb{N}$ . Dann gilt  $\langle z, x_k \rangle = z_k = 0$  für alle  $k$ , also  $z = 0$ . Damit ist  $O$  vollständig nach Satz 1.5.7 (ii).

Wir beschäftigen uns nun mit der Existenz von Orthogonalbasen in Hilberträumen. Wir werden sehen, dass es ein einfaches Kriterium gibt, das äquivalent zur Existenz einer Orthogonalbasis ist. Dazu zunächst eine Definition und ein paar Beispiele.

**Definition 1.5.10** (separabel). Ein unitärer Raum  $(E, \langle \cdot, \cdot \rangle)$  heißt *separabel*, wenn es eine abzählbare Teilmenge  $M$  von  $E$  gibt, die dicht in  $E$  liegt.

**Beispiel 1.5.11.**

- (a) Sei  $E = \mathbb{R}^n$ . Die Menge  $M = \{(r_1, \dots, r_n) \mid r_j \in \mathbb{Q}, j = 1, \dots, n\}$  ist abzählbar, da  $\mathbb{Q}$  abzählbar ist. Dann liegt  $M$  dicht in  $\mathbb{R}^n$ , denn  $\mathbb{Q}$  liegt dicht in  $\mathbb{R}$  und damit gilt

$$\|(x_1, \dots, x_n) - (y_1, \dots, y_n)\|^2 = \sum_{j=1}^n \underbrace{|x_j - y_j|}_{\leq \varepsilon^2 2^{-j}} \leq \varepsilon^2,$$

also folgt  $\|x - y\| \leq \varepsilon$  für  $(x_1, \dots, x_n) \in \mathbb{R}^n$  beliebig und entsprechend gewählte  $r_1, \dots, r_n \in \mathbb{Q}$ , was immer möglich ist, da  $\mathbb{Q}$  dicht in  $\mathbb{R}$  liegt. Daraus folgt, dass der  $\mathbb{R}^n$  separabel ist.



- (b)  $E = \mathbb{C}^n$  ist separabel.
- (c)  $E = L^2[a, b]$  bezüglich des Lebesgue-Maßes ist separabel. Beim Beweis dieser Aussage kann man ähnlich wie in Beispiel (a) durch Approximation der stetigen Funktionen mit Treppenfunktionen, die auf  $\mathbb{Q}$  definiert werden, vorgehen. Wir beweisen diese Tatsache jedoch mit dem *Satz von Weierstraß*, welcher besagt, dass zu jeder stetigen Funktion  $f(x)$  auf dem Intervall  $[a, b]$  eine Folge  $(p_n)_{n \in \mathbb{N}}$  von Polynomen  $p_n(x)$  existiert, die gleichmäßig auf  $[a, b]$  gegen  $f(x)$  konvergieren. Mit  $\mathbb{Q}[x][i]$  bezeichnen wir die Menge der Polynome der Form

$$\sum_{k=0}^n (r_k + is_k)x^k, n \in \mathbb{N}_0, r_k, s_k \in \mathbb{Q},$$

also alle Polynome mit rationalen und rational-imaginären Koeffizienten. Ebenso beschreibt  $\mathbb{C}[x]$  die Menge der Polynome mit komplexen Koeffizienten. Unser Ziel ist es nun, zu zeigen, dass  $\mathbb{Q}[x][i]$  dicht in  $L^2[a, b]$  liegt. Da  $\mathbb{Q}[x][i]$  abzählbar ist, würde somit die Separabilität von  $L^2[a, b]$  folgen. Sei  $\|f\| = \sup_{x \in [a, b]} |f(x)|$  die Norm auf den Polynomringen. Dann liegt  $\mathbb{Q}[x][i]$  bezüglich  $\|\cdot\|$  dicht in  $\mathbb{C}[x]$ ,  $\mathbb{C}[x]$  ist wieder  $\|\cdot\|$ -dicht in  $\mathcal{C}[a, b]$  nach dem Satz von Weierstraß. Aus  $p_n \rightarrow f$  für  $n \rightarrow \infty$  auf  $[a, b]$  folgt, dass  $p_n \rightarrow f$  in  $(L^2[a, b], \|\cdot\|_2)$ , denn

$$\lim_{n \rightarrow \infty} \|p_n - f\|_2^2 = \lim_{n \rightarrow \infty} \int_a^b \underbrace{|p_n(x) - f(x)|^2}_{\rightarrow 0 \text{ auf } [a, b]} \lambda(dx) = \int_a^b \lim_{n \rightarrow \infty} |p_n(x) - f(x)|^2 \lambda(dx) = 0,$$

da  $p_n$  auf  $[a, b]$  gleichmäßig gegen  $f$  konvergiert. Daraus folgt nun, dass  $\mathbb{Q}[x][i]$  dicht in  $L^2[a, b]$  ist und somit erhält man die behauptete Aussage.

- (d) Sei  $\Omega$  eine offene oder abgeschlossene Teilmenge im  $\mathbb{R}^n$ . Dann ist  $L^2(\Omega, \lambda_n)$  separabel.
- (e) Auch der  $\ell^2(\mathbb{N})$  ist separabel.
- (f) Wenn  $X$  eine überabzählbare Menge ist und  $\mu$  das Zählmaß auf  $X$ , dann ist  $\ell^2(X, \mu)$  nicht separabel.

Der folgende Satz zeigt, dass die Separabilität eine relativ starke Eigenschaft von Hilberträumen ist, denn sie garantiert die Existenz einer Orthonormalbasis. Umgekehrt kann eine solche Basis nicht existieren, ohne dass der Raum separabel ist.

Um den Beweis elegant führen zu können, benötigen wir noch ein wichtiges Hilfsmittel, das *Orthogonalisierungsverfahren nach Gram und Schmidt*. Wir möchten ein Orthonormalsystem mit beliebig vielen Vektoren für einen unitären Raum  $(E, \langle \cdot, \cdot \rangle)$  konstruieren. Sei dazu zuerst  $M = \{y_n, n \in \mathbb{N}\}$  eine abzählbare Teilmenge von  $E$ . Wir konstruieren nun eine Folge  $(r_n)_{n \in \mathbb{N}}$ , indem wir zuerst  $r_1$  als die kleinste natürliche Zahl wählen, so dass  $y_{r_1} \neq 0$  ist. Wir setzen dann

$$x_1 = \frac{y_{r_1}}{\|y_{r_1}\|},$$

damit  $\|x_1\| = 1$  gilt. Dann ist die lineare Hülle von  $\{x_1\}$  identisch mit der von  $\{y_1, \dots, y_{r_1}\}$ , denn alle Vektoren  $y_i$  mit  $i < r_1$  waren ja Null.

**Satz 1.5.12.** Sei  $(H, \langle \cdot, \cdot \rangle)$  ein Hilbertraum.  $H$  besitzt eine Orthonormalbasis aus endlich oder abzählbar vielen Elementen genau dann, wenn  $H$  separabel ist.

*Beweis.* Wir zeigen zuerst die Implikation „ $\Rightarrow$ “. Sei dazu  $(x_k)_{k \in N}$  mit  $N = \{1, \dots, n\}$  oder  $N = \mathbb{N}$  eine Orthonormalbasis. Nach Satz 1.5.7 (i) gibt es für alle  $x \in H$  die Darstellung  $x = \sum_{k \in N} \langle x, x_k \rangle x_k$ . Die Menge

$$M = \left\{ \sum_{k=1}^n (r_k + is_k)x_k \text{ für } n \in \mathbb{N}, r_k, s_k \in \mathbb{Q} \right\},$$

also die Menge aller endlichen Linearkombinationen von Basisvektoren mit rationalen und imaginär-rationalen Koeffizienten, liegt daher dicht in  $H$ . Da  $M$  abzählbar ist, folgt diese Richtung des Beweises.

” $\Leftarrow$ “: Sei  $H$  separabel. Dann existiert eine abzählbare Teilmenge  $M = \{y_n, n \in \mathbb{N}\}$  die dicht in  $H$  liegt. Nach dem Orthogonalisierungsverfahren von Gram und Schmidt konstruieren wir ein Orthonormalsystem  $\{x_k \mid k \in \mathbb{N}\}$  mit  $\text{span} \{x_k \mid k \in \mathbb{N}\} = \text{span} M$ . Wir werden unter Benutzung von Satz 1.5.7 (ii) zeigen, dass dieses Orthonormalsystem vollständig ist. Sei  $z \in H$  mit  $\langle z, x_k \rangle = 0$  für alle  $k \in \mathbb{N}$ . Dann ist  $\langle z, y \rangle = 0$  für alle  $y \in \text{span} M$ . Da  $M$  dicht in  $H$  liegt folgt, dass  $\langle z, x \rangle = 0$  für alle  $x \in \overline{M} = H$  ist. Damit muss aber  $z = 0$  sein und das Orthonormalsystem ist vollständig.  $\square$

**Korollar 1.5.13.** Sei  $H$  ein Hilbertraum.

- (i) Wenn  $H$  als Vektorraum mit  $d = \dim H$  endlichdimensional ist, dann ist  $H$  zum Hilbertraum  $\mathbb{R}^d$  beziehungsweise  $\mathbb{C}^d$  mit dem üblichen Skalarprodukt isomorph.
- (ii) Ist  $H$  als Vektorraum unendlichdimensional und separabel, dann ist  $H$  zum Hilbertschen Folgenraum  $\ell^2(\mathbb{N})$  isomorph.

*Beweis.* Wir beweisen nur die zweite Aussage. Nach Satz 1.5.12 besitzt  $H$  eine Orthonormalbasis aus höchstens abzählbar vielen Elementen, etwa  $\{x_k \mid k \in \mathbb{N}\}$ . Wir definieren eine Abbildung  $f : H \rightarrow \ell^2(\mathbb{N})$  durch  $f(x) = (\langle x, x_k \rangle)_{k \in \mathbb{N}}$ . Da  $\sum_{k=1}^{\infty} |\langle x, x_k \rangle|^2 = \|x\|^2$  aufgrund der Vollständigkeit von  $H$  ist, ist  $f(x) \in \ell^2(\mathbb{N})$ . Für  $x, y \in H$  gilt

$$\langle f(x), f(y) \rangle = (\langle x, x_k \rangle)_k, (\langle y, x_k \rangle)_k = \sum_{k=1}^{\infty} \langle x, x_k \rangle \overline{\langle y, x_k \rangle} = \sum_{k=1}^{\infty} \langle x, x_k \rangle \langle x_k, y \rangle = \langle x, y \rangle$$

nach Satz 1.5.7 (iv).  $f$  erhält also das Skalarprodukt. Weiterhin ist  $f$  auch injektiv, denn aus  $f(x) = 0$ , also  $\langle x, x_k \rangle = 0$  für alle  $k \in \mathbb{N}$ , folgt  $\|x\|^2 = \sum_{k=1}^{\infty} |\langle x, x_k \rangle|^2 = 0$  nach Satz 1.5.7 (iii). Daraus folgt dann, dass  $x = 0$ . Indem wir die Surjektivität von  $f$  zeigen, haben wir bestätigt, dass  $f$  einen Hilbertraumisomorphismus induziert. Sei  $(a_k)_{k \in \mathbb{N}} \in \ell^2(\mathbb{N})$ . Wir setzen  $x = \sum_{k=1}^{\infty} a_k x_k$ . Nach Satz 1.5.4 konvergiert diese Reihe im Hilbertraum  $H$ . Damit haben wir

$$\langle x, x_l \rangle = \sum_{k=1}^{\infty} \langle a_k x_k, x_l \rangle = \sum_{k=1}^{\infty} a_k \underbrace{\langle x_k, x_l \rangle}_{\delta_{kl}} = a_k,$$

das heißt  $f(x) = (\langle x, x_k \rangle) = (a_k)$  und  $f$  ist surjektiv.  $\square$

**Bemerkung 1.5.14.** Sei  $\Omega$  eine offene oder abgeschlossene Teilmenge des  $\mathbb{R}^n$ . Dann ist  $L^2(\Omega, \lambda)$  separabel und unendlichdimensional, also nach Korollar 1.5.13 (ii) isomorph zum  $\ell^2(\mathbb{N})$ .

Nun wenden wir die eben erarbeiteten Ergebnisse auf gewöhnliche Fourierreihen an.

**Satz 1.5.15.** Die trigonometrischen Systeme  $\left\{ \frac{1}{\sqrt{2\pi}}, \frac{\sin nx}{\sqrt{\pi}}, \frac{\cos mx}{\sqrt{\pi}} \text{ für } n, m \in \mathbb{N} \right\}$  und  $\left\{ \frac{e^{inx}}{\sqrt{2\pi}} \text{ für } n \in \mathbb{Z} \right\}$  sind im Hilbertraum  $L^2[-\pi, \pi]$  vollständig.

Der Satz beruht im Wesentlichen auf dem Satz von Fejer, welcher besagt, dass es zu jeder stetigen Funktion  $f$  auf dem Intervall  $[-\pi, \pi]$  mit  $f(-\pi) = f(\pi)$  eine Folge von trigonometrischen Polynomen  $\{q_n, n \in \mathbb{N}\}$  der Form  $q(x) = a_0 + \sum_{k=1}^n (a_k \sin kx + b_k \cos kx)$  gibt, wobei  $a_i, b_i \in \mathbb{C}$ , welche für  $n \rightarrow \infty$  auf dem Intervall  $[-\pi, \pi]$  gleichmäßig gegen  $f$  konvergiert.

*Beweis.* Sei  $z \in L^2[-\pi, \pi]$  mit  $\langle z, q \rangle = 0$  für alle trigonometrischen Polynome  $q$ . Der Abschluss der trigonometrischen Polynome in  $L^2[-\pi, \pi]$  enthält nach dem Satz von Fejer alle stetigen Funktionen auf  $[-\pi, \pi]$  mit  $f(-\pi) = f(\pi)$ . Diese liegen dicht in  $L^2[-\pi, \pi]$ . Daraus folgt, dass  $z \perp L^2[-\pi, \pi]$ , also  $z = 0$ . Jedes dieser beiden Systeme ist vollständig nach Satz 1.5.7 (ii).  $\square$

Trigonometrische Fouriersysteme sind also vollständig in  $L^2[-\pi, \pi]$ , das heißt

$$\left\{ \frac{e^{inx}}{\sqrt{2\pi}}, n \in \mathbb{Z} \right\}$$

ist ein vollständiges Orthonormalsystem in  $L^2[-\pi, \pi]$ . Alle Aussagen von Satz 1.5.7 gelten also insbesondere für dieses System. Sei  $f \in L^2[-\pi, \pi]$  und  $\sum_{k=-\infty}^{\infty} c_k \frac{e^{ikx}}{\sqrt{2\pi}}$  sei die Fourierreihe von  $f$ , wobei

$$c_k = \left\langle f, \frac{e^{ikx}}{\sqrt{2\pi}} \right\rangle = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} f(x) e^{-ikx} \mu(dx), \quad k \in \mathbb{Z}.$$

Die Fourierreihe von  $f$  konvergiert gegen  $f$  in  $L^2[-\pi, \pi]$  nach Satz 1.5.7, das heißt

$$\lim_{n \rightarrow \infty} \left\| f - \sum_{k=-n}^n c_k \frac{e^{ikx}}{\sqrt{2\pi}} \right\|^2 = \lim_{n \rightarrow \infty} \int \left| f(x) - \sum_{k=-n}^n c_k \frac{e^{ikx}}{\sqrt{2\pi}} \right|^2 \mu(dx) = 0.$$

Daraus kann man natürlich keine punktweise Konvergenz oder Konvergenz fast überall folgern, sondern es handelt sich um einen neuen Konvergenzbegriff. Es gibt jedoch Kriterien, die die punktweise oder sogar gleichmäßige Konvergenz von Fourierreihen unter stärkeren Voraussetzungen an  $f$  garantieren.



# Kapitel 2

## Grundprinzipien der Funktionalanalysis

### 2.1 Das Bairesche Kategoriethorem

Sei  $(E, d)$  ein metrischer Raum und sei  $K_\varepsilon(x_0) = \{y \in E \mid d(x_0, y) < \varepsilon\}$  die Kugel mit Radius  $\varepsilon$  um  $x_0$ . Wir erinnern zunächst an die Definition einer dichten Teilmenge eines metrischen Raumes, siehe auch Definition 1.1.10, und führen einen dazu konträren Begriff ein.

**Definition 2.1.1.**

- (i) Seien  $M, K \subseteq E$ .  $K$  heißt *dicht* in  $M$ , wenn zu jedem  $x \in M$  eine Folge  $(x_n)_{n \in \mathbb{N}}$  mit  $x_n \in K$  und  $\lim_{n \rightarrow \infty} x_n = x$  existiert. Äquivalent dazu ist  $M \subseteq \overline{K}$ .
- (ii) Eine Menge  $M \subseteq E$  heißt *nirgends dicht* in  $E$ , wenn  $M$  in keiner Umgebung eines Punktes dicht ist. Dies bedeutet, dass  $K_\varepsilon(x_0) \not\subseteq \overline{M}$  ist für alle  $x_0 \in M$  und alle  $\varepsilon > 0$ .

**Definition 2.1.2.** Eine Menge  $M \subseteq E$  heißt von *1. Kategorie*, wenn  $M$  abzählbare Vereinigung von nirgends dichten Mengen ist. Eine Menge, die nicht von 1. Kategorie ist, heißt von *2. Kategorie*.

**Beispiel 2.1.3.**

- (a) Sei  $E = \mathbb{R}$  mit Metrik  $d(x, y) = |x - y|$ . Die Punktmenge  $\{x_0\}$  ist nirgends dicht, denn  $\overline{\{x_0\}} = \{x_0\}$  enthält kein Intervall  $(a, b)$  und somit auch keine  $\varepsilon$ -Umgebung. Damit sind die Punktfolgen sämtlich nirgends dicht. Für jede Wahl von  $(x_n)_{n \in \mathbb{N}}$  in  $\mathbb{R}$  ist also  $M = \bigcup_{n=1}^{\infty} \{x_n\}$  von 1. Kategorie. Wegen der Überabzählbarkeit von  $\mathbb{R}$  gilt  $M \neq \mathbb{R}$  für jede Wahl der  $x_n$ , da wir höchstens abzählbar viele verschiedene reelle Zahlenerfasst haben.
- (b) Sei  $E = \mathbb{N}$  mit der Metrik  $d(x, y) = |x - y|$ . Hier ist  $\{n\}$  nicht nirgends dicht, weil zum Beispiel  $\overline{\{n\}} = K_{\frac{1}{2}}(n) = \{n\}$  gilt und damit haben wir eine Umgebung gefunden, nämlich  $K_{\frac{1}{2}}(n)$ , in der  $\{n\}$  dicht ist. Hier sieht man wieder einmal die große Wichtigkeit des ausgewählten metrischen Raumes, auf den man sich bezieht.
- (c) Sei  $E = \mathbb{Q}$ ,  $d(x, y) = |x - y|$ . Dann gilt  $E = \bigcup_{n=1}^{\infty} \{r_n\}$ , wobei  $\mathbb{Q} = \{r_i \mid i \in \mathbb{N}\}$  eine Aufzählung der abzählbaren Menge  $\mathbb{Q}$  ist.  $\{r_n\}$  ist nirgends dicht, denn in jeder noch so kleinen Umgebung eines Punktes liegen unendlich viele weitere rationale Zahlen, und der gesamte Raum  $\mathbb{Q}$  ist damit von 1. Kategorie.

Im letzten Beispiel sahen wir, dass auch der gesamte metrische Raum von 1. Kategorie sein kann. Für den betrachteten Raum  $\mathbb{Q}$  gilt jedoch, dass er nicht vollständig ist. Dass dies sogar zwingend folgt, besagt der folgende Satz, das Bairesche Kategoriethorem.

**Satz 2.1.4** (Bairesches Kategoriethorem). *Jeder nichtleere vollständige metrische Raum  $(E, d)$  ist von 2. Kategorie.*

Anders ausgedrückt: Sei  $(E, d)$  ein vollständiger metrischer Raum. Wenn  $E = \bigcup_{n=1}^{\infty} A_n$ , dann gibt es ein  $k \in \mathbb{N}$ , so dass  $A_k$  nicht nirgends dicht ist, das heißt, dass  $\overline{A_k}$  einen inneren Punkt enthält. Es gibt also ein  $x_0 \in E$  und ein  $\varepsilon > 0$ , so dass  $K_{\varepsilon}(x_0) \subseteq \overline{A_k}$ .

*Beweis.* Angenommen,  $E$  sei von 1. Kategorie, das heißt  $E = \bigcup_{n=1}^{\infty} A_n$  mit  $A_n$  nirgends dicht in  $E$  für alle  $n \in \mathbb{N}$ . Insbesondere ist  $A_1$  nirgends dicht in  $E$ . Also gibt es ein  $x_1 \in E$  mit  $x_1 \notin \overline{A_1}$ . Dann existiert ein  $\varepsilon_1 \in (0, 1)$  mit  $K_{\varepsilon_1}(x_1) \cap A_1 = \emptyset$ .  $A_2$  ist ebenfalls nirgends dicht in  $E$ .  $K_{\varepsilon_1}(x_1) \not\subseteq \overline{A_2}$ , also existiert  $x_2 \in K_{\varepsilon_1}(x_1)$  mit  $x_2 \notin \overline{A_2}$ , das heißt es gibt ein  $\varepsilon_2 \in (0, \frac{1}{2})$  mit der Eigenschaft  $K_{\varepsilon_2}(x_2) \cap A_2 = \emptyset$  und  $\overline{K_{\varepsilon_2}(x_2)} \subseteq K_{\varepsilon_1}(x_1)$ . Fortsetzung des Verfahrens ergibt Folgen  $(x_n)$  und  $(\varepsilon_n)$  mit folgenden Eigenschaften:

$$\overline{K_{\varepsilon_n}(x_n)} \subseteq K_{\varepsilon_{n-1}}(x_{n-1}), \quad K_{\varepsilon_n}(x_n) \cap A_n = \emptyset, \quad \varepsilon_n \in \left(0, \frac{1}{2^{n-1}}\right).$$

Dann ist  $(x_n)$  eine Cauchyfolge, denn für ein beliebiges  $N \in \mathbb{N}$  ist für  $n, m > N$

$$d(x_n, x_m) \leq d(x_n, x_N) + d(x_m, x_N) \leq \varepsilon_N + \varepsilon_N \leq \frac{1}{2^{N-1}} + \frac{1}{2^{N-1}}.$$

Da  $E$  vollständig ist, existiert  $x = \lim_{n \rightarrow \infty} x_n$  und es ist  $x \in E$ . Für ein beliebig gewähltes  $M \in \mathbb{N}$  und  $n \geq M$  ist  $x_n \in K_{\varepsilon_n}(x_n) \subseteq K_{\varepsilon_M}(x_M)$ . Für  $n \rightarrow \infty$  gilt  $x_n \rightarrow x$ , also liegt der Grenzwert  $x$  in  $\overline{K_{\varepsilon_M}(x_M)}$  und deshalb liegt  $x$  auch in allen Umgebungen  $K_{\varepsilon_m}(x_m)$  für alle  $m \geq M$ . Da  $E = \bigcup_{n=1}^{\infty} A_n$ , existiert ein  $k \in \mathbb{N}$  mit  $x \in A_k$ . Dann folgt  $x \in K_{\varepsilon_k}(x_k) \cap A_k$ , im Widerspruch zu  $K_{\varepsilon_k}(x_k) \cap A_k = \emptyset$ .  $\square$

### Beispiel 2.1.5.

- (a) Wir erhalten aus dem Baireschen Kategoriethorem einen neuen Beweis für die Überabzählbarkeit von  $\mathbb{R}$ . Wäre  $\mathbb{R}$  nämlich abzählbar, sagen wir  $\mathbb{R} = \bigcup_{i=1}^{\infty} \{x_i\}$ , dann wäre  $\mathbb{R}$  von 1. Kategorie, was wegen der Vollständigkeit nicht sein kann.
- (b) Es gibt stetige Funktionen auf  $[0, 1]$ , die in keinem Punkt  $x \in [0, 1]$  differenzierbar sind. Weierstraß entdeckte eine solche Funktion als erster, deswegen heißt ein spezielles Beispiel auch Weierstraß-Funktion. Das einfachste explizite Beispiel ist

$$f(x) = \sum_{k=1}^{\infty} \frac{\{10^k x\}}{10^k},$$

wobei  $\{y\}$  der Abstand von  $y$  zur nächsten ganzen Zahl ist. Wir zeigen nun, dass es unendlich viele Funktionen mit der Eigenschaft geben muss, stetig, aber nirgends differenzierbar zu sein. Sei nämlich  $E = \mathcal{C}[0, 2]$  mit der Metrik  $d(f, g) = \sup_{t \in [0, 2]} |f(t) - g(t)|$ .  $E$  ist dann vollständig. Wir definieren

$$M_n = \left\{ f \in E \mid \exists t \in [0, 1] \text{ mit } \sup_{0 < h < 1} \left| \frac{f(t+h) - f(t)}{h} \right| \leq n \right\}.$$

Dann ist  $M_n$  abgeschlossen für jedes  $n \in \mathbb{N}$ . Um dies zu zeigen, sei  $f_k \rightarrow f \in E$  mit  $f_k \in M_n$  für  $k \in \mathbb{N}$ . Sei  $t_k$  der in der Definition von  $M_n$  auftauchende Punkt für die Funktion  $f_k$ . Nach dem Satz von Bolzano-Weierstrass besitzt  $t_k$  eine konvergente Teilfolge. O.B.d.A. gelte  $t_k \rightarrow t_0 \in [0, 1]$ . Sei  $h$  fest und  $\varepsilon > 0$ . Dann gilt

$$\begin{aligned} \left| \frac{f(t_k+h) - f(t_k)}{h} \right| &\leq \left| \frac{f(t_k+h) - f_k(t_k+h)}{h} \right| + \left| \frac{f_k(t_k+h) - f_k(t_k)}{h} \right| + \left| \frac{f_k(t_k) - f(t_k)}{h} \right| \\ &\leq \varepsilon + n + \varepsilon = n + 2\varepsilon, \end{aligned}$$

wenn  $k \geq n_0(\varepsilon)$  ist, denn  $f_k$  konvergiert gleichmäßig gegen  $f$ . Für  $k \rightarrow \infty$  und  $\varepsilon \rightarrow 0$  gilt dann  $|\frac{1}{h}f(t_0+h) - f(t_0)| \leq n$  und damit ist  $f \in M_n$ , also  $M_n$  abgeschlossen. Weiterhin ist  $M_n$  nirgends dicht. Sei nämlich  $g \in E$  und  $\varepsilon > 0$ . Betrachte die Kugel  $K_\varepsilon(g)$  im Raum  $E$ , das heißt  $f \in K_\varepsilon(g)$  genau dann wenn  $\sup_{t \in [0,2]} |f(t) - g(t)| < \varepsilon$ . Wähle  $f \in K_\varepsilon(g)$  wie im folgenden Bild.

Alle Anstiege seien betragsmäßig größer als  $n$ . Dann folgt  $f \notin M_n = \overline{M_n}$ , also ist  $M_n$  nirgends dicht. Daraus folgt  $E \neq \bigcup_{n=1}^{\infty} M_n$ . Damit gibt es aber ein  $H \in E$  mit  $H \notin M_n$  für alle  $n \in \mathbb{N}$ . Also ist  $H$  in keinem Punkt differenzierbar, das heißt für jedes  $t \in [0, 1]$  ist  $\sup_{0 < h < 1} |\frac{1}{h}(H(t+h) - H(t))| = \infty$ .

- (c) Das Bairesche Theorem verleiht gewissermaßen Kriterien, die auf kleinen Umgebungen gelten, auf dem ganzen Raum Gültigkeit. Ein Beispiel ist der Satz von Carominas-Balaguer-Agnon. Sei  $f \in C^\infty(a, b)$  mit  $a, b \in \mathbb{R}$  und  $a < b$ . Zu jedem  $x \in (a, b)$  existiere ein  $n_x \in \mathbb{N}_0$  mit  $f^{(n_x)}(x) = 0$ . Dann ist  $f$  ein Polynom, also ist  $f^{(n)} = 0$  für ein  $n \in \mathbb{N}$ . Dass für jedes  $x$  eine Ableitung von  $f$  existiert, so dass diese bei  $x$  Null ist, ist eine lokale Voraussetzung. Das Bairesche Kategoriethem verleiht dieser lokalen Aussage dann globale Gültigkeit auf dem ganzen Raum, nämlich  $f^{(n)}(x) = 0$  für alle  $x \in (a, b)$ . Der kritische Punkt im Beweis ist zu zeigen, dass  $\sup_{x \in (x_0 - \varepsilon, x_0 + \varepsilon)} |n_x| < \infty$ .

Eine Variante des Satzes 2.1.4 ist die folgende, fast analog zu beweisende Umformulierung. Man beachte dabei, dass aufgrund der Kompaktheit die ausgewählte Folge  $(x_n)$  konvergiert.

**Satz 2.1.6.** *Jeder nichtleere lokalkompakte metrische Raum ist von 2. Kategorie.*

## 2.2 Stetige lineare Abbildungen auf normierten Vektorräumen

Seien  $(E_1, \|\cdot\|_1)$  und  $(E_2, \|\cdot\|_2)$  normierte lineare Räume. Mit  $L(E_1, E_2)$  bezeichnen wir den Raum aller linearen Abbildungen von  $E_1$  nach  $E_2$ . Aus der Algebra ist bekannt, dass es sich dabei um einen Vektorraum über demselben Körper wie  $E_2$  handelt. Zuerst wollen wir den Begriff der Stetigkeit, der aus der reellen Analysis bekannt ist, auf allgemeinere normierte lineare Räume übertragen.

**Definition 2.2.1.** Eine lineare Abbildung  $T \in L(E_1, E_2)$  heißt *stetig*, wenn aus  $x_n \rightarrow x$  in  $(E_1, \|\cdot\|_1)$  stets  $T(x_n) \rightarrow T(x)$  in  $(E_2, \|\cdot\|_2)$  folgt.

Auch für den Begriff der Beschränktheit von Mengen gibt es eine allgemeine Formulierung für Teilmengen von normierten linearen Räumen. Die Beschränktheit einer Abbildung zwischen normierten linearen Räumen wird hingegen ein wenig anders als in der reellen Analysis definiert. Dort bedeutet Beschränktheit, dass die Funktion auch im Unendlichen nicht beliebig groß werden darf, bei der folgenden Definition ist dies nur im Endlichen nötig.

**Definition 2.2.2.** Eine Menge  $M$  aus  $(E_1, \|\cdot\|_1)$  heißt *beschränkt*, wenn  $\sup_{x \in M} \|x\|_1 < \infty$  gilt. Das bedeutet, dass  $M$  in  $\overline{K_\rho(0)}$  enthalten ist, wenn  $\rho := \sup_{x \in M} \|x\|_1$  endlich ist. Weiterhin definieren wir für  $j = 1, 2$  die *Einheitssphäre*  $S_j$  in  $(E_j, \|\cdot\|_j)$ ,

$$S_j = \{x \in E_j \mid \|x\|_j = 1\},$$

und die *abgeschlossene Einheitskugel*  $B_j$  in  $(E_j, \|\cdot\|_j)$ ,

$$B_j = \{x \in E_j \mid \|x\|_j \leq 1\}.$$

Also ist  $M$  beschränkt in  $(E_1, \|\cdot\|_1)$  genau dann, wenn es ein  $\rho > 0$  gibt mit  $M \subseteq \rho B_1$ .

**Definition 2.2.3.** Eine lineare Abbildung  $T \in L(E_1, E_2)$  heißt *beschränkt*, wenn  $T$  beschränkte Teilmengen von  $E_1$  auf beschränkte Teilmengen von  $E_2$  abbildet, die Beschränktheit von Mengen also unter  $T$  invariant ist.

Als eine Übungsaufgabe zeigt man, dass die Beschränktheit einer linearen Abbildung  $T : E_1 \rightarrow E_2$  äquivalent ist zur Beschränktheit von  $T$  auf  $B_1$  und sogar ebenfalls äquivalent zur Beschränktheit auf lediglich  $S_1$ .

Im folgenden wollen wir die Eigenschaften linearer Abbildungen ein wenig herausarbeiten. Da beschränkte lineare Abbildungen, die wir auch beschränkte lineare Operatoren nennen, den Hauptgegenstand weiterer Kapitel bieten werden, ist es wichtig zu wissen, unter welchen Umständen wir Beschränktheit oder Stetigkeit überhaupt von einem Operator erwarten können. Sei  $T \in L(E_1, E_2)$  eine lineare Abbildung. Wir unterscheiden zunächst zwei mögliche Fälle. Gibt es ein  $\lambda > 0$  mit  $\|T(x)\|_2 \leq \lambda \|x\|_1$  für alle  $x \in E$ , dann definieren wir

$$\|T\| := \inf \{ \lambda > 0 \mid \|T(x)\|_2 \leq \lambda \|x\|_1 \text{ für alle } x \in E \}.$$

Existiert kein derartiges  $\lambda > 0$ , so setzen wir  $\|T\| = \infty$ . Dies ist natürlich konsistent mit der vorherigen Definition von  $\|T\|$ , wenn das Infimum der leeren Menge ebenfalls  $\infty$  ist.

**Definition 2.2.4** (Operatornorm). Für einen Operator  $T \in L(E_1, E_2)$  heißt  $\|T\|$  die *Operatornorm* von  $T$ .  $\mathcal{L}(E_1, E_2)$  bezeichnet die Menge aller beschränkten linearen Operatoren und es gilt natürlich  $\mathcal{L}(E_1, E_2) \subseteq L(E_1, E_2)$ .

Da  $\|T\|$  ist die kleinste Zahl  $\lambda \geq 0$  mit  $\|T(x)\|_2 \leq \lambda \|x\|_1$  für alle  $x \in E_1$  ist, gilt folglich für alle  $x \in E_1$

$$\|T(x)\|_2 \leq \|T\| \|x\|_1.$$

Diese Gleichung ist sehr wichtig, wir werden sie oft verwenden. Weiterhin definieren wir

$$\mu_B = \sup_{\|x\|_1 \leq 1} \|T(x)\|_2, \quad \mu_S = \sup_{\|x\|_1 = 1} \|T(x)\|_2.$$

Natürlich ist  $\mu_B \geq \mu_S$ . Weiterhin ist klar, dass  $\|T\| < \infty$  genau dann gilt, wenn  $T$  beschränkt ist. Wir werden nun einen Satz beweisen, der die Endlichkeit von  $\|T\|$  und damit die Beschränktheit des Operators mit  $\mu_B$  und  $\mu_S$  in Verbindung bringt. Damit haben wir ein handliches äquivalentes Kriterium für die Beschränktheit einer linearen Abbildung, denn  $\mu_S$  ist meistens leicht zu bestimmen. Dies wird uns bei späteren Beweisen mehr als nützlich sein.

**Satz 2.2.5.** *Für eine lineare Abbildung  $T \in L(E_1, E_2)$  sind folgende Aussagen äquivalent:*

- (i)  $T$  ist beschränkt,
- (ii)  $\|T\| < \infty$ ,
- (iii)  $\mu_B < \infty$ ,
- (iv)  $\mu_S < \infty$ .

Wenn  $T$  beschränkt ist, gilt weiterhin  $\|T\| = \mu_B = \mu_S$ .

*Beweis.* Sei  $x \in E_1$ ,  $x \neq 0$ . Setze  $y = \frac{x}{\|x\|_1}$ , also ist  $\|y\| = 1$ . Es gilt

$$\|T(x)\|_2 = \|T(\|x\|_1 y)\|_2 = \|T(y)\|_2 \|x\|_1 \leq \mu_S \|x\|_1.$$

Da diese Ungleichung auch für  $x = 0$  gilt, folgt  $\|T\| \leq \mu_S \leq \mu_B$ . Nun zeigen wir noch  $\mu_B \leq \|T\|$ . Dies ist klar, wenn  $\|T\| = \infty$ . Sei also  $\|T\| < \infty$ . Dann existiert ein  $\lambda > 0$  mit  $\|T(x)\|_2 \leq \lambda \|x\|_1$  für alle  $x \in E_1$ . Für  $x \in B_1$ , also  $\|x\|_1 \leq 1$ , gilt dann offensichtlich  $\|T(x)\|_2 \leq \lambda$ . Hieraus folgt

$$\sup_{\|x\|_1 \leq 1} \|T(x)\|_2 = \mu_B \leq \lambda,$$

also  $\mu_B \leq \inf \lambda = \|T\|$  und damit  $\mu_B = \|T\|$ . □

Mit diesen Ergebnissen ausgestattet sind wir nun in der Lage, eine Verknüpfung zwischen den Begriffen der Stetigkeit und der Beschränktheit zu ziehen. Es stellt sich heraus, dass für lineare Abbildungen diese Definitionen äquivalent sind, was es uns ermöglicht, im jeweiligen Kontext die nützlichere oder bequemere Eigenschaft zu verwenden. Wir erinnern zum Beweis daran, dass die Stetigkeit einer Funktion  $f$  zwischen topologischen Räumen dadurch definiert ist, dass die Urbilder offener Mengen unter  $f^{-1}$  wieder offen sind.



**Satz 2.2.6.** Sei  $T \in L(E_1, E_2)$  und  $x_0 \in E_1$  fest. Folgende Aussagen sind äquivalent:

- (i)  $T$  ist stetig in  $x_0$ ,
- (ii)  $T$  ist stetig in jedem Punkt  $x \in E_1$ ,
- (iii)  $T$  ist beschränkt.

*Beweis.* (i)  $\Rightarrow$  (ii): Sei  $(x_n)$  eine Folge in  $E_1$  mit  $x_n \rightarrow x$ . Dann gilt  $x_n - x + x_0 \rightarrow x_0$ . Da  $T$  in  $x_0$  stetig ist, folgt

$$T(x_n) - T(x) + T(x_0) = T(x_n - x + x_0) \rightarrow T(x_0),$$

also  $T(x_n) \rightarrow T(x)$ .

(ii)  $\Rightarrow$  (i): Diese Implikation ist trivial.

(ii)  $\Rightarrow$  (iii):  $T$  ist stetig in  $x = 0$ , das heißt das Urbild jeder offenen Umgebung von  $0 \in E_2$  ist eine offene Umgebung von  $0 \in E_1$ . Deswegen es gibt ein  $\delta > 0$ , so dass die Menge  $K_\delta(0)$  in  $E_1$  von  $T$  in die Menge  $K_1(0)$  in  $E_2$  abgebildet wird, das heißt  $T(K_\delta(0)) \subseteq K_1(0)$ . Sei  $x \in E_1$  mit  $x \neq 0$  und  $0 < r < \delta$ . Setze nun  $y = r \frac{x}{\|x\|_1}$ . Dann ist  $\|y\|_1 = r < \delta$ . Damit folgt, dass  $y \in K_\delta(0)$  ist, sich also  $T(y)$  in  $K_1(0)$  befinden muss, was wiederum bedeutet, dass  $\|T(y)\|_2 < 1$  sein muss, also

$$\|T(y)\|_2 = \frac{r}{\|x\|_1} \|T(x)\|_2 < 1.$$

Hieraus folgt  $\|T(x)\|_2 \leq \frac{1}{r} \|x\|_1$ . Da dies trivialerweise auch für  $x = 0$  gilt, ist  $T$  beschränkt.

(iii)  $\Rightarrow$  (ii): Es gilt für  $y, x \in E_1$ :

$$\|T(y) - T(x)\|_2 = \|T(y - x)\|_2 \leq \|T\| \|y - x\|_1.$$

Wenn nun also  $y_n \rightarrow x$  gilt, dann folgt  $T(y_n) \rightarrow T(x)$ , also ist  $T$  stetig in  $x$ . □

Wir haben weiter oben die Operatornorm definiert und dabei den Begriff der Norm verwendet. Dass es sich tatsächlich um eine solche, nämlich auf dem Raum  $\mathcal{L}(E_1, E_2)$  aller beschränkten linearen Operatoren handelt, besagt der folgende Satz. Er sichert uns auch gleich die Banachraumeigenschaft von  $\mathcal{L}(E_1, E_2)$ , vorausgesetzt, dass es sich bei dem Grundraum  $E_2$  um einen Banachraum handelt.

**Satz 2.2.7.**

- (i) Die Operatornorm ist eine Norm auf dem Vektorraum  $\mathcal{L}(E_1, E_2)$ .
- (ii) Wenn  $E_2$  ein Banachraum ist, dann ist  $\mathcal{L}(E_1, E_2)$  mit der Operatornorm ebenfalls ein Banachraum.

*Beweis.* Da die Normeigenschaften einfach nachzurechnen sind, beweisen wir lediglich (ii). Sei  $A_n \in \mathcal{L}(E_1, E_2)$  für  $n \in \mathbb{N}$ , wobei  $(A_n)_{n \in \mathbb{N}}$  eine Cauchyfolge in  $\mathcal{L}(E_1, E_2)$  sei. Sei weiter  $\varepsilon > 0$ . Dann existiert ein  $n_\varepsilon$  mit  $\|A_n - A_m\| < \varepsilon$  für alle  $n, m \geq n_\varepsilon$ . Sei  $x \in E_1$  fest. Es gilt

$$\|A_n(x) - A_m(x)\|_2 = \|(A_n - A_m)(x)\|_2 \leq \|A_n - A_m\| \|x\|_1.$$

Hieraus folgt, dass  $(A_n(x))_{n \in \mathbb{N}}$  eine Cauchyfolge in  $E_2$ . Da  $E_2$  ein Banachraum ist, konvergiert diese Folge, das heißt es gibt ein  $y_x \in E_2$  mit  $y_x = \lim_{n \rightarrow \infty} A_n(x)$ . Wir sehen, dass die Zuordnung  $x \mapsto y_x$  linear in  $x$  ist, das heißt  $A(x) - y_x$  ist ein linearer Operator von  $E_1$  nach  $E_2$ . Wir definieren einen linearen Operator  $A$  durch  $A(x) = y_x$ . Es gilt nach obigen Voraussetzungen stets

$$\|A_n(x) - A_m(x)\|_2 \leq \|A_n - A_m\| \|x\|_1 \leq \varepsilon \|x\|_1$$

für  $n, m > n_\varepsilon$ , und mit  $m \rightarrow \infty$  folgt dann

$$\lim_{m \rightarrow \infty} \|A_n(x) - A_m(x)\|_2 = \|A_n(x) - A(x)\|_2 \leq \varepsilon \|x\|_1$$

für  $n > n_\varepsilon$ . Weiter haben wir

$$\begin{aligned} \|A(x)\|_2 &= \|A(x) - A_n(x) + A_n(x)\|_2 \leq \|A(x) - A_n(x)\|_2 + \|A_n(x)\|_2 \\ &\leq \varepsilon \|x\|_1 + \|A_n\| \|x\|_1 = (\varepsilon + \|A_n\|) \|x\|_1. \end{aligned}$$

Da  $\|A_n\|_2$  beschränkt ist, folgt daraus, dass  $A$  beschränkt ist. Aus  $\|A(x) - A_n(x)\|_2 = \|(A - A_n)(x)\|_2 \leq \varepsilon \|x\|_1$  für  $n \geq n_\varepsilon$  folgt dann

$$\|A - A_n\| = \sup_{\|x\|_1 \leq 1} \|(A - A_n)(x)\|_2 \leq \varepsilon$$

für  $n \geq n_\varepsilon$ . Damit gilt  $\lim_{n \rightarrow \infty} \|A - A_n\| = 0$ , das heißt  $A_n \rightarrow A$  in der Operatornorm.  $\square$

Nach diesen theoretischen Überlegungen ist es an der Zeit, Beispiele für lineare und beschränkte Operatoren zu bringen. Viele der Operationen, die mit Funktionen durchgeführt werden können, beispielsweise Differentiation und Integration, lassen sich auch durch die Wirkung von gewissen Operatoren beschreiben. Diese Schreibweisen sind meist sehr übersichtlich und der Operatorenkalkül sichert für viele Operationen eine Fülle von Eigenschaften.

### Beispiel 2.2.8.

- (a) Sei  $E = \mathbb{C}^n$  und  $\|x\| = \sqrt{\sum_{i=1}^n \|x_i\|^2}$  für  $x \in E$ . Weiter sei  $A$  eine  $n \times n$ -Matrix. Wir definieren  $T_A(x) = A \cdot x$  für alle  $x \in E$ . Dann ist  $T_A$  ein linearer Operator von  $E$  nach  $E$ , wie man leicht anhand der Rechenregeln für Matrizen sieht.  $T_A$  ist beschränkt und es gilt  $\|T_A\| \leq \sqrt{\sum_{i,j=1}^n |a_{ij}|^2}$ . Als Übungsaufgabe zeigt man, dass jeder lineare Operator  $T: \mathbb{C}^n \rightarrow \mathbb{C}^n$  diese Gestalt hat.
- (b) Wir betrachten Integraloperatoren, die einen quadratisch integrierbaren Kern besitzen. Diese werden Hilbert-Schmidt-Operatoren genannt und bilden eine Klasse von Operatoren, denen man in der Funktionalanalysis häufig begegnet. Seien  $a, b \in \mathbb{R}$  mit  $a < b$ ,  $\Omega = [a, b] \times [a, b]$  und sei  $k(t, s): \Omega \rightarrow \mathbb{R}$  messbar. Es gelte weiterhin

$$\int_{\Omega} |k(t, s)|^2 \lambda(dt, ds) = \int_a^b \int_a^b |k(t, s)|^2 \lambda(dt) \lambda(ds) < \infty,$$

was auf jeden Fall erfüllt ist, wenn  $k$  stetig auf  $\Omega$  ist. Weiter definieren wir nun einen Operator  $T$  durch

$$(Tf)(t) = \int_a^b k(t, s) f(s) \lambda(ds)$$

für  $f \in L^2[a, b]$ . Es gilt unter Benutzung der Hölderschen Ungleichung (1.2.11):

$$\begin{aligned} \int_a^b |(Tf)(t)|^2 \lambda(dt) &= \int_a^b \left| \int_a^b k(t, s) f(s) \lambda(ds) \right|^2 \lambda(dt) \\ &< \int_a^b \left( \int_a^b |k(t, s)|^2 \lambda(ds) \right) \left( \int_a^b |f(s)|^2 \lambda(ds) \right) \lambda(dt) \\ &= \|f\|^2 \left( \int_a^b \int_a^b |k(t, s)|^2 \lambda(ds) \lambda(dt) \right) =: \|f\|^2 c_k < \infty. \end{aligned}$$

Aus  $\int |(Tf)(s)|^2 \lambda(ds) < \infty$  folgt, dass  $|(Tf)(s)|^2$   $\lambda$ -fast überall endlich ist, also erreicht man durch abändern auf einer Nullmenge, dass  $Tf$  in  $L^2[a, b]$  liegt.  $T$  bildet also den  $L^2[a, b]$  in sich selbst ab. Wegen  $\|Tf\|^2 \leq c_k \|f\|^2$  für alle  $f \in L^2[a, b]$  folgt weiterhin, dass  $T$  beschränkt ist, das heißt  $T \in \mathcal{L}(L^2[a, b], L^2[a, b])$ . Für die Norm von  $T$  gilt die Abschätzung

$$\|T\| \leq \left( \int_a^b \int_a^b |k(t, s)|^2 \lambda(ds) \lambda(dt) \right)^{1/2}.$$

- (c) Als nächstes betrachten wir sogenannte Differentialoperatoren, die den Vorgang der Differentiation durch die Wirkung von Operatoren ausdrücken. Es sei  $E_1 = \mathcal{C}^1[0, 1]$  die Menge der einmal stetig differenzierbaren und  $E_2 = \mathcal{C}[0, 1]$  die Menge der stetigen Funktionen auf  $[0, 1]$ . Eine Funktion aus  $E_1$  wird durch einmaliges Differenzieren auf eine Funktion aus  $E_2$  abgebildet. Auf den beiden Räumen seien die Normen

$$\|f\|_1 = \sup_{t \in [0, 1]} |f(t)| + \sup_{t \in [0, 1]} |f'(t)|$$

und

$$\|f\|_2 = \sup_{t \in [0, 1]} |f(t)|$$

vorgegeben. Für den Differentiationsoperator  $T$ , der in der Form  $(Tf)(t) = f'(t)$  wirkt, gilt  $T \in L(E_1, E_2)$  und wir haben die Abschätzung

$$\|Tf\|_2 = \|f'\|_2 = \sup_{t \in [0, 1]} |f'(t)| \leq \|f\|_1.$$

Das bedeutet, dass  $T$  beschränkt ist mit  $\|T\| < 1$ , also  $T \in \mathcal{L}(E_1, E_2)$ .

Ersetzen wir hingegen  $\|\cdot\|_1$  durch

$$\|f\|'_1 = \sup_{t \in [0, 1]} |f(t)| = \|f\|_2,$$

dann ist  $T$  nicht mehr beschränkt. Sei nämlich  $f_n(t) = t^n \in E_1$ . Für jedes  $f_n$  gilt  $\|f_n\|'_1 = 1$ , es handelt sich also um eine beschränkte Folge. Hingegen gilt  $Tf_n = nt^{n-1}$  mit  $\|Tf_n\|_2 = n$ . Also existiert kein  $\lambda > 0$  mit  $\|T(f)\|_2 \leq \lambda \|f\|'_1$ , denn dann wäre  $\|T(f_n)\|_2 = n \leq \lambda \|f_n\|'_1 = \lambda$  für alle  $n \in \mathbb{N}$ , was nicht sein kann. Wir folgern, dass  $T$  zwar ein linearer Operator von  $E_1$  nach  $E_2$  ist, hingegen die Beschränktheitseigenschaft nicht aufweist, wenn man von den normierten linearen Räumen  $(E_1, \|\cdot\|'_1)$  und  $(E_2, \|\cdot\|_2)$  ausgeht.

- (d) Ein weiterer wichtiger Operator ist der Operator der Fouriertransformation. Sei  $f \in L^1(\mathbb{R}^d, \lambda)$ . Wir definieren einen Operator  $F$  durch

$$(Ff)(t) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{-i\langle t, s \rangle} f(s) \lambda(ds).$$

Die Funktion  $e^{-i\langle t, s \rangle} f(s)$  ist dabei  $\lambda$ -integrierbar, da

$$|e^{-i\langle t, s \rangle} f(s)| = |f(s)| \in L^1(\mathbb{R}^d, \lambda).$$

Für den Operator  $F$  haben wir die Abschätzung für den punktweisen Betrag

$$|F(f)(t)| \leq \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} |f(s)| \lambda(ds) = \frac{1}{(2\pi)^{d/2}} \|f\|_{L^1(\mathbb{R}^d)}$$

und schließen daraus die folgende Ungleichung für die Norm:

$$\|F(f)\|_{L^\infty(\mathbb{R}^d)} \leq \frac{1}{(2\pi)^{d/2}} \|f\|_{L^1(\mathbb{R}^d)}.$$

Also ist  $F$  eine stetige lineare Abbildung von dem  $L^1(\mathbb{R}^d)$  in den  $L^\infty(\mathbb{R}^d)$  mit  $\|F\| \leq \frac{1}{(2\pi)^{d/2}}$ .

- (e) Sei  $H = \ell^2(\mathbb{N}_0)$  der Hilbertsche Folgenraum. Wir definieren den *Verschiebungsoperator*  $S$  mit  $S(x_0, x_1 \dots) = (0, x_0, x_1 \dots)$  für  $x \in H$ . Dieser Operator wird auch einseitiger Shift genannt. Allgemeiner kann  $S$  auch zweiseitig durch

$$S(\dots, x_{-1}, \underbrace{x_0}_{\text{Stelle 0}}, x_1, \dots) = (\dots, x_{-2}, \underbrace{x_{-1}}_{\text{Stelle 0}}, x_0, x_1, \dots)$$

auf dem Folgenraum  $\ell^2(\mathbb{Z})$  definiert werden.

- (f) Sei  $(X, \mu)$  ein Maßraum und  $H = L^2(X, \mu)$ . Für  $a \in L^\infty(X, \mu)$  definieren wir ein Funktional  $T_a : H \rightarrow H$  durch  $(T_a \varphi)(x) = a(x)\varphi(x)$  für alle  $\varphi \in L^2(X, \mu)$ . Dann gilt:

$$\int_X |T_a \varphi|^2 = \int_X |a(x)\varphi(x)|^2 d\mu \leq \|a\|_{L^\infty(X, \mu)}^2 \int_X |\varphi(x)|^2 d\mu.$$

Daraus folgt, dass

$$\|T_a \varphi\| \leq \|a\|_{L^\infty(X, \mu)}$$

ist und daraus ergibt sich dann, dass es sich bei  $T_a$  um einen beschränkten Operator handelt, für dessen Norm  $\|T_a\| \leq \|a\|_{L^\infty(X, \mu)}$  gilt.

## 2.2.1 Duale Räume

Der Dualraum ist ein wichtiges Konzept über Funktionale, von dem wir in den nächsten Abschnitten ausführlich Gebrauch machen werden. Die Definition hier unterscheidet sich insofern von der Definition des dualen Raumes in der Algebra, da uns hier durch die Normierung der Vektorräume zusätzliche Eigenschaften der Funktionale wie beispielsweise Stetigkeit zur Verfügung stehen. Wir bringen zunächst eine Definition des Dualraums und wiederholen dabei den Begriff des Funktionals, definieren dann Isomorphismen zwischen normierten linearen Räumen und bringen anschließend einige ausführliche Beispiele, die auch ein wenig strukturelles Licht auf uns bereits bekannte Räume werfen. Insbesondere werden wir sehen, dass die  $L^p$ -Räume stark untereinander zusammenhängen.

**Definition 2.2.9** (Dualraum). Sei  $(E, \|\cdot\|)$  ein normierter linearer Raum und  $E_2 = \mathbb{K}$  mit  $\|x\| = |x|$  der Grundkörper von  $E$ , also  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$ . Die stetigen linearen Abbildungen  $F : E \rightarrow \mathbb{K}$  heißen stetige lineare Funktionale von  $E$ . Mit  $E'$  bezeichnen wir den Vektorraum aller stetigen Funktionale von  $E$ . Der Raum  $(E', \|\cdot\|)$  mit der Operatornorm  $\|F\| = \inf \{\alpha > 0 \mid \|F(x)\| \leq \alpha \|x\|_1 \text{ für alle } x \in E\}$  ist ein normierter linearer Raum und da  $E_2$  vollständig ist, ist  $(E', \|\cdot\|)$  ebenfalls ein Banachraum.  $(E', \|\cdot\|)$  heißt der zu  $(E, \|\cdot\|)$  gehörige *duale Raum*.

**Definition 2.2.10.** Seien  $(E_1, \|\cdot\|)$  und  $(E_2, \|\cdot\|)$  zwei normierte lineare Räume. Diese Räume heißen *isomorph*, wenn es eine bijektive lineare Abbildung  $f$  von  $E_1$  auf  $E_2$  gibt, die  $\|f(x)\|_2 = \|x\|_1$  für alle  $x \in E_1$  erfüllt. Man nennt  $f$  dann einen *isometrischen Isomorphismus*.

### Beispiel 2.2.11.

- (a) Sei  $(H, \langle \cdot, \cdot \rangle)$  ein Hilbertraum. Nach Satz 1.4.8 sind alle stetigen Funktionale auf  $H$  von der Form  $F_y(x) = \langle x, y \rangle$  für  $x, y \in H$ . Wir fassen diese stetigen Funktionale im Dualraum  $H' = \{F_y \mid y \in H\}$  zusammen. Nun definieren wir einen weiteren Vektorraum  $H^-$  wie folgt: Als Menge sei  $H^- = H$ . Die Addition sei identisch mit der in  $H$  und die Multiplikation sei durch  $\lambda y_{H^-} = \bar{\lambda} y_H$  definiert. Die Abbildung  $H' \rightarrow H^-$ , die durch  $F_y \mapsto y$  gegeben ist, ist ein Isomorphismus von Vektorräumen. Es gilt nach der Cauchy-Schwarz-Ungleichung stets

$$|F_y(x)| = |\langle x, y \rangle| \leq \|y\| \|x\|,$$

das heißt, dass  $\|F_y\| \leq \|y\|$  ist. Weiterhin haben wir für  $x = y$

$$|F_y(y)| = |\langle y, y \rangle| = \|y\|^2 \leq \|F_y\| \|y\|$$

und damit  $\|y\| \leq \|F_y\|$ . Kombiniert man die beiden Aussagen, so ergibt sich  $\|F_y\| = \|y\|$ . Die Aussage gilt in allgemeineren Hilberträumen ebenfalls, es ist nämlich stets  $H' \cong H^-$ .

- (b) Sei  $E = L^p(X, \mu)$  mit  $1 < p < \infty$  und sei  $g \in L^q(X, \mu)$  mit einem  $q$  so, dass  $\frac{1}{p} + \frac{1}{q} = 1$  gilt. Wir definieren  $F_g(f) = \int_X fg d\mu$  für  $f \in L^p(X, \mu)$ . Diese Integrale existieren für alle  $g$  und  $f$ , da gilt:

$$\int_X |fg| d\mu \leq \left( \int_X |f|^p d\mu \right)^{1/p} \left( \int_X |g|^q d\mu \right)^{1/q} = \|f\|_{L^p} \|g\|_{L^q} < \infty.$$

Desweiteren haben wir

$$|F_g(f)| = \left| \int_X fg \, d\mu \right| \leq \int_X |fg| \, d\mu \leq \|g\|_{L^q} \|f\|_{L^p}.$$

Daraus folgt, dass  $F_g \in L^p(X, \mu)'$  ist und für die Norm des Funktionals  $\|F_g\| \leq \|g\|_{L^q}$  gilt. Man kann zeigen, dass  $\|F_g\| = \|g\|_{L^q}$  und dass jedes Funktional  $F$  aus  $L^p(X, \mu)'$  von der Form  $F_g$  mit einem  $g \in L^q(X, \mu)$  ist. Das bedeutet, dass die Abbildung von  $L^q(X, \mu) \in L^p(X, \mu)'$ , die durch  $g \mapsto F_g$  gegeben ist, ein isometrischer Isomorphismus ist. Damit gilt der allgemeine Zusammenhang

$$L^p(X, \mu)' \cong L^q(X, \mu), \quad 1 < p < \infty, \quad \frac{1}{p} + \frac{1}{q} = 1.$$

Für  $p = 1$  oder  $p = \infty$  ist diese Aussage nicht gültig. Es ist nämlich  $L^\infty(X, \mu)' \not\cong L^1(X, \mu)$ , aber hingegen  $L^1(X, \mu)' \cong L^\infty(X, \mu)$ .

- (c) Sei  $E = \mathcal{C}(X)$  der Raum der stetigen Funktionen auf  $X$ , wobei  $X$  ein kompakter metrischer Raum sei. Weiterhin sei  $\nu$  ein komplexes Borelmaß auf  $X$ . Wir definieren das Funktional  $F_\nu(f) = \int_X f \, d\nu$ .  $F_\nu$  ist ein stetiges lineares Funktional auf  $E$ . Da  $X$  kompakt ist, ist  $|\nu|$  stets endlich. Ist  $\nu$  weiterhin auch positiv, so gilt dann

$$|F_\nu(f)| \leq \int_X |f| \, d\nu \leq \|f\| \int_X d\nu = \nu(X) \|f\|$$

und daraus folgt, dass  $F_\nu \in E'$  ist und die Norm von  $F_\nu$  durch die Abschätzung  $\|F_\nu\| \leq \nu(X)$  charakterisiert wird. Nach dem 2. Satz von Riesz sind alle stetigen linearen Funktionale auf  $E = \mathcal{C}(X)$  von der Form  $F_\nu$ , wenn  $\nu$  ein komplexes Borelmaß auf  $X$  ist.

## 2.3 Das Theorem von Banach-Steinhaus

Der Satz von Banach-Steinhaus ist das zentrale Thema dieses Abschnittes. Es handelt sich um eine für weitere Untersuchungen sehr wichtige Aussage, da man eine oft benötigte Regularitätsbedingung, die gleichmäßige Beschränktheit einer Menge von linearen Abbildungen, aufgrund von schwächeren Voraussetzungen, nämlich der punktweisen Beschränktheit erhält. Da punktweise Beschränktheit in vielen Fällen nicht schwer zu zeigen ist, können eine Menge Folgerungen elegant bewiesen werden. Auch später in diesem Skript werden wir noch oft auf diesen Satz zurückgreifen.

Da der Satz von Banach-Steinhaus eine lokale Aussage (punktweise Beschränktheit) auf den gesamten Raum (gleichmäßige Beschränktheit) überträgt, liegt es nahe, Verknüpfungen mit dem Baireschen Kategoriethorem zu vermuten. Tatsächlich werden wir dieses im Beweis benötigen.

**Satz 2.3.1** (Satz von Banach-Steinhaus). *Sei  $E$  ein Banachraum und  $F$  ein normierter linearer Raum. Sei weiter  $T_i \in \mathcal{L}(E, F)$  für  $i \in I$ , wobei  $I$  eine beliebige Indexmenge sei.  $\{T_i \mid i \in I\}$  sei punktweise beschränkt, das heißt für jedes  $x \in E$  sei*

$$c_x := \sup_{i \in I} \|T_i(x)\| < \infty.$$

*Dann ist  $\{T_i \mid i \in I\}$  gleichmäßig beschränkt, das heißt*

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

*Beweis.* Sei  $A_n = \{x \in E \mid \|T_i(x)\| \leq n \text{ für alle } i \in I\}$ , wobei  $n \in \mathbb{N}$  sein soll. Da jedes  $T_i$  stetig ist, sind die Mengen  $A_n$  sämtlich abgeschlossen. Weiterhin ist die Menge der Abbildungen  $\{T_i \mid i \in I\}$  punktweise

beschränkt, das heißt es gibt für jedes  $x \in E$  ein  $n \in \mathbb{N}$  mit  $x \in A_n$ , so dass also  $E = \bigcup_{n=1}^{\infty} A_n$  gilt. Nach dem Baireschen Kategoriethemem (Satz 2.1.4) existiert, da  $E$  ein Banachraum ist, ein  $x_0 \in E$ , ein  $\varepsilon > 0$  und ein  $n \in \mathbb{N}$  mit  $K_\varepsilon(x_0) \subseteq A_n = A_n$ .

Sei  $x \in E$  mit  $\|x\| \leq 1$ . Wir wählen uns ein  $r$  mit  $0 < r < \varepsilon$  und setzen  $y = x_0 + rx$ . Dann gilt  $\|y - x_0\| = \|rx\| \leq r\|x\| < \varepsilon$ , also ist  $y \in K_\varepsilon(x_0)$ . Für  $T_i(x)$  ergibt sich somit

$$\|T_i(x)\| = \left\| T_i \left( \frac{y - x_0}{r} \right) \right\| \leq \frac{1}{r} (\|T_i(y)\| + \|T_i(x_0)\|) \leq \frac{1}{r}(n + c_{x_0}),$$

da  $y \in K_\varepsilon(x_0) \subseteq A_n$  und damit  $\|T_i(y)\| \leq n$  ist. Aus dieser Aussage, die nicht von  $i$  und vom konkret gewählten  $x$  abhängt, solange nur  $\|x\| \leq 1$  ist, folgt die Behauptung, da

$$\sup_{i \in I} \|T_i\| = \sup_{i \in I} \sup_{\|x\| \leq 1} \|T_i(x)\| \leq \frac{1}{r}(n + c_{x_0}) < \infty$$

□

Aus dem Beweis kann man als einfache Übungsaufgabe zusätzlich noch eine weitere Aussage herleiten. Seien  $E, F$  wie im Satz 2.3.1 und  $T_i \in \mathcal{L}(E, F)$  für alle  $i$  aus einer Indexmenge  $I$ . Sei weiter  $E_1 = \{x \in E \mid c_x = \sup_{i \in I} \|T_i(x)\| < \infty\}$ . Dann ist  $E_1 = E$  oder  $E_1$  ist eine Menge der ersten Kategorie.

Kommen wir nun zu einigen Anwendungen und Beispielen des Banach-Steinhaus-Theorems. Der Satz schafft viel strukturelle Klarheit innerhalb wichtiger Banachräume.

**Beispiel 2.3.2.** Zunächst bringen wir ein Beispiel, bei dem die Voraussetzungen des Banach-Steinhaus-Theorems nicht erfüllt sind, denn dann kann man auch die Implikationen des Satzes nicht erwarten. Sei  $E$  der Vektorraum der finiten Folgen, also aller Folgen, die nur endlich viele Einträge haben, mit der Norm  $\|(x_n)_{n \in \mathbb{N}}\| := \sup_{n \in \mathbb{N}} |x_n|$  für  $(x_n)_{n \in \mathbb{N}} \in E$ . Sei weiter  $(E, \|\cdot\|)$  ein normierter linearer Raum, der jedoch kein Banachraum sein möge. Dann können wir eine punktweise beschränkte Menge von linearen Abbildungen konstruieren, die nicht gleichmäßig beschränkt ist. Wir wählen ein  $k \in \mathbb{N}$  und definieren

$$T_k(x_1, \dots, x_n, 0, \dots) = (1 \cdot x_1, 2 \cdot x_2, \dots, k \cdot x_k, 0, \dots).$$

Dann liegt  $T_k$  in  $\mathcal{L}(E, E)$  und die Gleichung  $\|T_k\| = k$  ergibt sich aus

$$\begin{aligned} \|T_k(x)\| &= \sup_{n \in \mathbb{N}} \{|1 \cdot x_1|, \dots, |k \cdot x_k|\} \leq k \cdot \sup_{n \in \mathbb{N}} \{|x_1|, \dots, |x_k|\} \\ &\leq k \cdot \sup_{n \in \mathbb{N}} \{|x_n|\} = k \cdot \|(x_n)_{n \in \mathbb{N}}\|. \end{aligned}$$

Da  $k \in \mathbb{N}$  beliebig groß wird, kann die Menge  $\{T_k \mid k \in \mathbb{N}\}$  nicht gleichmäßig beschränkt sein.

**Korollar 2.3.3.** Sei  $E$  ein Banachraum und  $F$  ein normierter linearer Raum. Weiter sei  $A_k \in \mathcal{L}(E, F)$  für  $k \in \mathbb{N}$  und es existiere ein nach Voraussetzung nicht notwendig beschränktes  $A \in \mathcal{L}(E, F)$  mit  $\lim_{k \rightarrow \infty} A_k(x) = A(x)$  für alle  $x \in E$ . Dann ist  $A \in \mathcal{L}(E, F)$ , das heißt, dass  $A$  tatsächlich doch beschränkt ist, und für die Norm von  $A$  gilt  $\|A\| \leq \lim_{k \rightarrow \infty} \|A_k\|$ .

*Beweis.* Aus der Voraussetzung folgt, dass  $\{A_k \mid k \in \mathbb{N}\}$  punktweise beschränkt ist, das heißt wir können Satz 2.3.1 anwenden. Damit folgt  $\sup_{k \in \mathbb{N}} \|A_k\| < \infty$ . Aus dieser gleichmäßigen Beschränktheit ergibt sich nun

$$\|A_k(x)\| \leq \|A_k\| \|x\| \leq C \|x\|$$

für alle  $x \in E$  mit einem  $C \in \mathbb{R}$ . Im Limes  $k \rightarrow \infty$  ergibt sich damit  $\|A(x)\| \leq C \|x\|$  und daraus folgt, dass  $A$  beschränkt ist. Wir wählen nun eine Teilfolge  $(A_{k_n})_{n \in \mathbb{N}}$  mit  $\lim_{n \rightarrow \infty} \|A_{k_n}\| = \liminf_{n \rightarrow \infty} \|A_n\|$  und wenden das gleiche Argument wie oben bei  $A_k$  an. Daraus ergibt sich  $\|A(x)\| \leq \liminf_{n \rightarrow \infty} \|A_n\| \|x\|$  und damit  $\|A\| \leq \lim_{k \rightarrow \infty} \|A_k\|$ . □

Insbesondere haben wir folgendes gezeigt. Wenn  $F = \mathbb{C}$  und  $(E, \|\cdot\|)$  ein Banachraum ist und wir lineare stetige Funktionale  $F_k \in E'$  für alle  $k \in \mathbb{N}$  definiert haben, für die  $\lim_{k \rightarrow \infty} F_k(x) =: F(x)$  existiert, dann ist  $F \in E'$  ebenfalls ein stetiges lineares Funktional auf  $E$ .

**Korollar 2.3.4** (Satz von Mazur-Orlicz). Seien  $E$  und  $F$  Banachräume.  $B(\cdot, \cdot)$  sei eine Bilinearform auf  $E \times F$ . Weiterhin sei  $B$  partiell stetig in beiden Variablen, das heißt wenn  $x_n \rightarrow x$  in  $E$  und  $y_n \rightarrow y$  in  $F$  gilt, so ist  $\lim_{n \rightarrow \infty} B(x_n, y_0) = B(x, y_0)$  und  $\lim_{n \rightarrow \infty} B(x_0, y_n) = B(x_0, y)$  bei beliebigen festen  $y_0 \in F$  und  $x_0 \in E$ .

Dann ist  $B$  gleichmäßig stetig, das heißt wenn  $x_n \rightarrow x$  in  $E$  und  $y_n \rightarrow y$  in  $F$  konvergiert, dann gilt  $\lim_{n \rightarrow \infty} B(x_n, y_n) = B(x, y)$ .

*Beweis.* Es gilt

$$B(x_n, y_n) - B(x, y) = B(x_n - x, y_n) + B(x, y_n - y),$$

wobei nach Voraussetzung  $x_n - x \rightarrow 0$  und  $y_n - y \rightarrow 0$  konvergieren. O.B.d.A. kann man also annehmen, dass  $x = 0$  oder  $y = 0$  gilt. Sei  $y = 0$ , also  $y_n \rightarrow 0$ . Wir definieren  $T_n(y) = B(x_n, y)$  für  $y \in F$ .  $T_n$  ist eine lineare Abbildung von  $F$  nach  $\mathbb{C}$ . Weil  $B$  partiell stetig ist, ist  $T_n$  stetig für jedes  $n \in \mathbb{N}$ . Es gilt  $T_n(y) = B(x_n, y) \rightarrow B(x, y) = T(y)$  wegen der partiellen Stetigkeit für alle  $y \in F$ . Daraus folgt, dass  $\sup_{n \in \mathbb{N}} \|T_n(y)\| < \infty$  für jedes  $y \in E$  gilt. Nach dem Banach-Steinhaus-Theorem (Satz 2.3.1) ist damit  $C = \sup_{n \in \mathbb{N}} \|T_n\| < \infty$ . Dann haben wir

$$|B(x_n, y_n)| = |T_n(y_n)| \leq \|T_n\| \|y_n\| \leq C \|y_n\| \rightarrow 0,$$

da  $y_n \rightarrow 0$ . Also gilt  $B(x_n, y_n) \rightarrow 0 = B(x, 0)$ . □

## 2.4 Der Satz von der offenen Abbildung und der Satz vom abgeschlossenen Graphen

Zuerst wollen wir den Satz von der offenen Abbildung beweisen, der auch Satz von Banach-Schauder heißt. Der Beweis dieses Satzes hat strukturelle Ähnlichkeit zum Beweis des Satzes von der Umkehrabbildung im  $\mathbb{R}^n$ .

**Satz 2.4.1** (Satz von Banach-Schauder). *Jede stetige surjektive lineare Abbildung  $T$  eines Banachraumes  $E$  auf einen Banachraum  $F$  ist offen, das heißt  $T(\mathcal{O})$  ist eine offene Menge in  $F$  für jede offene Menge  $\mathcal{O}$  aus  $E$ .*

*Beweis.* Wir wollen zeigen, dass es für jedes  $\varepsilon > 0$  und jedes  $x_0 \in E$  ein  $\delta > 0$  gibt, so dass  $T(U_\varepsilon(x_0)) \supseteq U_\delta(T(x_0))$  gilt. Da  $T$  linear ist, sei o.B.d.A.  $x_0 = 0$ . Zu zeigen ist dann, dass es für jedes  $\varepsilon > 0$  ein  $\delta_\varepsilon > 0$  gibt mit  $T(U_\varepsilon(0)) \supseteq U_{\delta_\varepsilon}(0)$ . Sei  $\varepsilon > 0$ . Der erste Schritt ist zu beweisen, dass  $\overline{T(U_\varepsilon(0))}$  eine Kugel  $U_\beta(0)$  in  $F$  für ein  $\beta > 0$  umfasst. Sei dazu  $U = U_\varepsilon(0)$ . Aus  $E = \bigcup_{n=1}^{\infty} nU$  folgt  $T(E) = \bigcup_{n=1}^{\infty} nT(U) = F$ , weil  $T$  surjektiv ist. Nach dem Baireschen Kategoriethemem (Satz 2.1.4) existiert ein  $m \in \mathbb{N}$  mit  $\overline{mT(U)} \supseteq U_\alpha(y)$ ,  $y \in F$ ,  $\alpha > 0$ . Das ist äquivalent zu  $\overline{T(U)} \supseteq \frac{1}{m}U_\alpha(y)$ . Aus  $U_\varepsilon(0) \supseteq U_{\varepsilon/2}(0) - U_{\varepsilon/2}(0)$  folgt

$$\overline{T(U)} \supseteq \overline{T(U_{\varepsilon/2}(0)) - T(U_{\varepsilon/2}(0))} \supseteq \overline{T(U_{\varepsilon/2}(0))} - \overline{T(U_{\varepsilon/2}(0))}.$$

Die letzte Inklusion sieht man so ein: Wenn  $x_n \rightarrow x$  und  $y_n \rightarrow y$  mit  $x_n, y_n \in T(U_{\varepsilon/2}(0))$  und  $x, y \in \overline{T(U_{\varepsilon/2}(0))}$ , so gilt  $x - y \in \overline{T(U_{\varepsilon/2}(0))} - \overline{T(U_{\varepsilon/2}(0))}$ . Weiter konvergiert aber  $x_n - y_n \rightarrow x - y$ , wobei für die Differenzen stets  $x_n - y_n \in T(U_{\varepsilon/2}) - T(U_{\varepsilon/2})$  und  $x - y \in \overline{T(U_{\varepsilon/2}(0))} - \overline{T(U_{\varepsilon/2}(0))}$  ist.

Damit haben wir

$$\overline{T(U)} \supseteq \overline{T(U_{\varepsilon/2}(0))} - \overline{T(U_{\varepsilon/2}(0))} \supseteq \frac{1}{m}(U_{\alpha/2}(y) - U_{\alpha/2}(y)) = \bigcup_{z \in \frac{1}{m}U_{\alpha/2}(y)} \left\{ z - \frac{1}{m}U_{\alpha/2}(y) \right\} = \tilde{U},$$

wobei  $\tilde{U}$  als Vereinigung offener Mengen offen ist.  $\tilde{U}$  enthält die Null, also enthält  $\tilde{U}$  eine kleine Kugel  $U_\beta(0)$  um Null.

Im zweiten Schritt zeigen wir, dass  $T(U_\varepsilon(0))$  eine Kugel um 0 in  $F$  enthält. Wähle dazu eine Folge  $(r_n)_{n \in \mathbb{N}}$  positiver Zahlen mit  $\sum_{n=1}^{\infty} r_n < \frac{\varepsilon}{2}$ . Sei  $E_\alpha = U_\alpha(0)$  in  $E$ ,  $F_\delta = U_\delta(0)$  in  $F$  und sei  $r_0 = \frac{\varepsilon}{2}$ . Nach dem im ersten Schritt Gezeigten existiert für jedes  $n \in \mathbb{N}$  ein  $\delta_n > 0$  mit  $F_{\delta_n} \subseteq \overline{T(E_{r_n})}$ . O.B.d.A. sei  $\lim_{n \rightarrow \infty} \delta_n = 0$ . Wir zeigen nun  $F_{\delta_0} \subseteq T(E_\varepsilon)$ . Sei dafür  $y \in F_{r_0}$ . Dann ist also  $y \in \overline{T(E_{r_0})}$ . Damit existiert ein  $x_0 \in E_{r_0}$  mit  $\|y - T(x_0)\| < \delta_1$  und damit gilt  $y - T(x_0) \in F_{\delta_1} \subseteq \overline{T(E_{r_1})}$ . Also existiert ein  $x_1 \in E_{r_1}$  mit  $\|y - T(x_0) - T(x_1)\| < \delta_2$ . Induktiv konstruieren wir eine Folge  $(x_n)$  in  $E_{r_n}$  mit  $\|y - T(x_0) - \dots - T(x_n)\| < \delta_{n+1}$  und  $\|x_n\| \leq r_n$ . Weil  $\sum_{n=1}^{\infty} r_n < \infty$ , konvergiert  $\sum_{n=1}^{\infty} r_n$  im Banachraum  $E$ . Sei  $x = \sum_{n=0}^{\infty} x_n \in E$ . Dann gilt

$$\|x\| \leq \sum_{n=0}^{\infty} \|x_n\| \leq \sum_{n=1}^{\infty} r_n + \|x_0\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

also gilt sogar  $x \in E_\varepsilon$  und

$$\left\| y - T \left( \sum_{k=0}^n x_k \right) \right\| \leq \delta_{n+1}.$$

Wegen der Stetigkeit der Norm und von  $T$  folgt daraus

$$\begin{aligned} \lim_{n \rightarrow \infty} \left\| y - T \left( \sum_{k=0}^n x_k \right) \right\| &= \left\| y - \lim_{n \rightarrow \infty} T \left( \sum_{k=0}^n x_n \right) \right\| \leq \lim_{n \rightarrow \infty} \delta_{n+1} = 0 \\ &= \left\| y - T \left( \lim_{n \rightarrow \infty} \sum_{k=0}^n x_n \right) \right\| = \|y - T(x)\| \leq 0, \end{aligned}$$

und damit ist  $y = T(x)$ . □

**Korollar 2.4.2.** Sei  $T$  eine stetige bijektive lineare Abbildung des Banachraums  $E$  auf den Banachraum  $F$ . Dann ist die Umkehrabbildung  $T^{-1}$  stetig, das heißt  $T^{-1} \in \mathcal{L}(F, E)$ .

*Beweis.*  $T^{-1}$  existiert wegen der Bijektivität und bildet  $F$  eineindeutig auf  $E$  ab. Sei  $U$  eine offene Menge in  $E$ . Dann ist  $V := T(U)$  nach Satz 2.4.1 eine offene Menge in  $F$ . Das Urbild  $T^{-1}(V) = U$  ist offen in  $E$ , also ist  $T^{-1}$  stetig. □

Nun kommen wir zum Satz vom abgeschlossenen Graphen. Dazu benötigen wir erst einmal ein paar Definitionen. Stellt man sich beispielsweise eine gezeichnete Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  vor, so handelt es sich im Grunde genommen um ein zweidimensionales Bild des Graphen der Funktion, nämlich um die Menge  $(x, f(x))$  für alle betrachteten  $x \in \mathbb{R}$ . Ebenso verhält es sich mit Abbildungen  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ , wo wir uns Flächen im dreidimensionalen Raum vorstellen, also einen dreidimensionalen Graphen. Dieses Konzept eines Graphen lässt sich stets durch Paarung von Argument und Funktionswert mathematisch fassen, auch in allgemeinen Räumen.

**Definition 2.4.3.** Sei  $T$  eine Abbildung der Menge  $M$  in die Menge  $N$ . Die Menge  $\{(x, T(x)) \mid x \in M\}$  heißt der *Graph* von  $T$ . Man schreibt  $G(T) = \{(x, T(x)) \mid x \in M\}$ .

**Definition 2.4.4.** Sei  $E$  ein normierter linearer Raum und sei  $D$  ein Unterraum von  $E$ . Eine lineare Abbildung des normierten linearen Raumes  $D$  in den normierten linearen Raum  $F$  heißt *abgeschlossen*, wenn der Graph  $G(T)$  von  $T$  eine abgeschlossene Teilmenge des normierten linearen Raumes  $E \oplus F$  ist.

Der normierte lineare Raum  $E \oplus F$  ist dabei der Vektorraum  $E + F$  der direkten Summe von  $E$  und  $F$ , versehen mit der Norm  $\|(x, y)\| = \|x\| + \|y\|$  für  $x \in E$  und  $y \in F$ . Sind  $E$  und  $F$  Banachräume, dann ist auch  $E \oplus F$  ein Banachraum.

**Bemerkung 2.4.5.** Es gelten folgende Implikationen:



- (i) Jede stetige lineare Abbildung von  $E$  nach  $F$  ist abgeschlossen.
- (ii) Sei  $T \in L(E, F)$ . Dann ist  $T$  abgeschlossen genau dann, wenn aus  $x_n \rightarrow 0$  in  $E$  und  $T(x_n) \rightarrow y$  in  $F$  stets  $y = 0$  folgt. Man beachte, dass an nicht konvergierende Folgen  $T(x_n)$  keine Anforderungen gestellt werden.

Die Bemerkung 2.4.5 (i) lässt sich umkehren unter der Voraussetzung, dass es sich bei den betrachteten Räumen  $E$  und  $F$  um Banachräume handelt. Dies ist der Inhalt des folgenden Satzes. Handelt es sich auch nur bei einem der beiden Räume nicht um einen Banachraum, so kann man nicht mehr erwarten, dass aus der Abgeschlossenheit auch die Stetigkeit folgt. Gegenbeispiele kann man sich bei Bedarf ohne weiteres konstruieren.

**Satz 2.4.6** (Satz vom abgeschlossenen Graphen). *Jede abgeschlossene lineare Abbildung  $T$  eines Banachraums  $E$  in einen Banachraum  $F$  ist stetig.*

*Beweis.* Da  $E$  und  $F$  Banachräume sind, ist auch  $E \oplus F$  ein Banachraum. Nach Voraussetzung ist  $G(T)$  abgeschlossen in  $E \oplus F$ , also ist  $G(T)$  in der Norm von  $E \oplus F$  ebenfalls ein Banachraum. Definiere die folgende lineare Abbildung  $P$  von  $G(T)$  auf  $E$ :

$$P(x, T(x)) = x, \quad (x, T(x)) \in G(T), x \in E.$$

Dann ist  $P$  stetig, denn wenn  $x_n \rightarrow x$  und  $T(x_n) \rightarrow y$  für ein beliebiges  $y$ , dann ist

$$P(x_n, T(x_n)) = x_n \rightarrow x = P(x, T(x)),$$

wobei das zweite Argument von  $P$  keine Rolle spielt. Weiterhin ist  $P$  bijektiv, denn wenn  $P(x, T(x)) = P(x', T(x'))$  ist, dann ist also  $x = x'$  und damit  $T(x) = T(x')$ . Nach Folgerung 2.4.2 ist die Abbildung  $P^{-1}$  stetig, das heißt wenn  $x_n \rightarrow x$  gilt, dann gilt auch  $(x_n, T(x_n)) = P^{-1}(x_n) \rightarrow P^{-1}(x) = (x, T(x))$ , insbesondere also  $T(x_n) \rightarrow T(x)$  und damit ist  $T$  ebenfalls stetig.  $\square$

## 2.5 Das Hahn-Banach-Theorem

Ebenso wichtig und strukturfördernd wie das Banach-Steinhaus-Theorem ist das Hahn-Banach-Theorem. Es garantiert uns unter gewissen Voraussetzungen die Existenz von stetigen linearen Funktionalen mit gewünschten Eigenschaften auf Banachräumen. Eine grundlegende Frage, die sich mit dem Hahn-Banach-Theorem beantworten lässt, ist die nach stetigen linearen Funktionalen, die in einem bestimmten Punkt nicht Null werden. Genauer gibt es tatsächlich zu jedem  $x \in E$  ein  $f \in E'$  mit  $f(x) \neq 0$ .

Eine andere Frage, die wir beantworten wollen, handelt von der Trennung konvexer Mengen. Für zwei konvexe und disjunkte Mengen im  $\mathbb{R}^2$  gibt es eine Gerade, die diese Mengen trennt, so dass die eine Menge auf einer Seite der Geraden liegt und die andere auf der anderen Seite. Es wäre nun interessant zu wissen, ob dies auch in allgemeineren Banachräumen der Fall ist. Hier werden wir jedoch relativ starke Voraussetzungen an die betrachteten Mengen stellen müssen, damit wir diese Frage positiv beantworten können.

Um diese Probleme in Angriff nehmen zu können, definieren wir uns zunächst einige wichtige Begriffe. Es folgen als erstes einige mengentheoretische Resultate, die wir für unsere Untersuchungen gut gebrauchen können.

**Definition 2.5.1** (Halbordnung). Es sei  $M$  eine beliebige Menge. Eine *Halbordnung* auf  $M$  ist eine Relation  $\leq$  auf  $M$ , so dass für alle  $x, y, z \in M$  folgende Eigenschaften erfüllt sind:

- (i) Aus  $x \leq y$  und  $y \leq x$  folgt  $x = y$  , (Antisymmetrie)
- (ii)  $x \leq x$  , (Reflexivität)
- (iii) ist  $x \leq y$  und  $y \leq z$ , dann ist auch  $x \leq z$ . (Transitivität)

Eine Halbordnung ist beispielsweise die Mengeninklusion  $\subseteq$  auf einem Mengensystem. Man beachte, dass bei einer Halbordnung auf einer Menge zwei Elemente dieser Menge nicht vergleichbar sein müssen. Mengen, bei denen dies jedoch der Fall ist, nennt man total geordnet.

**Definition 2.5.2.** Eine Teilmenge  $K \subseteq M$  einer Menge  $M$  mit einer Halbordnung  $\leq$  heißt *total geordnet*, wenn für alle  $x, y \in K$  entweder  $x \leq y$  oder  $y \leq x$  gilt, wenn also je zwei Elemente aus  $K$  miteinander vergleichbar sind.

**Definition 2.5.3.** Ein Element  $x_0$  einer halbgeordneten Menge  $M$  heißt

- (i) *obere Schranke* für  $K \subseteq M$ , wenn für alle  $x \in K$  gilt, dass  $x \leq x_0$  ist.
- (ii) *maximales Element*, wenn aus  $x_0 \leq x \in K$  stets  $x_0 = x$  folgt.

**Beispiel 2.5.4.**

- (i) Sei  $M = \mathcal{C}[0, 1]$  die Menge der stetigen Funktionen von  $\mathbb{R}$  nach  $\mathbb{R}$ . Definiere eine Halbordnung auf  $M$  durch  $f \leq g \Leftrightarrow f(x) \leq g(x)$  für alle  $x \in [0, 1]$ . Man sieht leicht, dass  $M$  mit dieser Halbordnung nicht total geordnet ist.
- (ii) Die Menge  $M = \mathbb{R}^2$  mit der Halbordnung  $(x_1, x_2) \leq (y_1, y_2) \Leftrightarrow x_1 \leq x_2$  und  $x_2 \leq y_2$  ist ebenfalls nicht total geordnet.
- (iii) Sei  $M = \{(x, y) \in \mathbb{R}^2 \mid 0 \leq x \leq 1, 0 \leq y \leq 1 - x\}$ . Diese Menge stellt ein ausgefülltes „Dreieck“ im ersten Quadranten dar. Jeder Punkt auf der Diagonalen  $x = 1 - y$  ist für diese Menge maximal.

Von außerordentlicher Wichtigkeit ist das folgende Zornsche Lemma. Es ist unerlässlich beispielsweise für den Beweis, dass jedes Ideal eines Ringes in einem maximalen Ideal enthalten ist. Der Beweis des Zornschen Lemmas, den wir hier nicht geben werden, benutzt jedoch das Auswahlaxiom, ist also aus den „intuitiven“ Axiomen der Mengenlehre nicht ableitbar. In der Mengenlehre wird gezeigt, dass Auswahlaxiom und Zornsches Lemma sogar äquivalent zueinander sind: Setzt man eines der beiden voraus, kann man damit das jeweils andere beweisen. Eine dritte zu diesen beiden äquivalente Aussage ist die des Zermeloschen Wohlordnungssatzes.

**Lemma 2.5.5** (Lemma von Zorn). Wenn jede totalgeordnete Teilmenge einer halbgeordneten Menge  $M$  eine obere Schranke in  $M$  besitzt, dann gibt es in  $M$  ein maximales Element.

Im folgenden sei  $E$  ein reeller Vektorraum. Wir kennen bereits lineare Funktionale auf  $E$ . Diese besitzen jedoch zu starke Eigenschaften für einen effizienten Einsatz des Hahn-Banach-Theorems. Wir werden stattdessen die schwächeren sublinearen Funktionale benutzen. Ebenfalls werden wir keine Norm mehr benötigen, sondern lediglich eine Halbnorm. Diese haben wir bereits in Definition 1.2.1 eingeführt, wir wiederholen den Begriff aber hier aus Gründen der Vollständigkeit und Übersichtlichkeit. Man beachte die subtilen Unterschiede zwischen einer Halbnorm und einem sublinearen Funktional.

**Definition 2.5.6.** Ein *sublineares Funktional* auf  $E$  ist eine Abbildung  $q : E \rightarrow \mathbb{R}$  mit folgenden Eigenschaften:

- (i)  $q(x_1 + x_2) \leq q(x_1) + q(x_2)$  für alle  $x_1, x_2 \in E$ .
- (ii)  $q(\lambda x) = \lambda q(x)$  für alle  $x \in E$  und  $\lambda \geq 0$ .

**Definition 2.5.7** (Halbnorm). Eine *Halbnorm* auf  $E$  ist eine Abbildung  $p : E \rightarrow [0, \infty)$  mit den Eigenschaften

- (i)  $p(x_1 + x_2) \leq p(x_1) + p(x_2)$  für alle  $x_1, x_2 \in E$ . (Dreiecksungleichung)
- (ii)  $p(\lambda x) = |\lambda| p(x)$  für  $x \in E$  und  $\lambda \in \mathbb{R}$ .

**Beispiel 2.5.8.**

- (i) Sei  $E$  ein  $\mathbb{R}$ -Vektorraum und seien  $f_1, f_2, \dots, f_r$  lineare Funktionale auf  $E$ . Wir definieren  $p(x) = |f_1(x)| + \dots + |f_r(x)|$ , dann ist  $p$  sowohl eine Halbnorm als auch ein sublineares Funktional. Definiert man hingegen  $q(x) = f_1(x)$ , dann ist  $q$  keine Halbnorm, sofern  $f_1 \neq 0$  ist, aber dennoch ein sublineares Funktional. Jede Halbnorm ist natürlich ein sublineares Funktional.
- (ii) Sei  $E = \mathcal{C}[0, 1]$  der Vektorraum aller stetigen Abbildungen von  $\mathbb{R}$  nach  $\mathbb{R}$  und es sei  $p(f) = \sup_{t \in [0, 1]} |f(t)|$ . Dann ist  $p$  eine Halbnorm und ein sublineares Funktional auf  $E$ .

Wir formulieren und beweisen nun den Satz von Hahn-Banach für sublineare Funktionale im Fall eines reellen Vektorraums. Man kann die sublinearen Funktionale durch Halbnormen ersetzen und die reellen durch komplexe Vektorräume. Diese Aussagen werden wir wenig später ebenfalls beweisen.

**Satz 2.5.9** (Hahn-Banach-Theorem für reelle Vektorräume). *Es sei  $F$  ein linearer Teilraum eines reellen Vektorraumes  $E$  und  $q$  ein sublineares Funktional auf  $E$ . Weiter sei  $f$  ein lineares Funktional auf  $F$  mit  $f(x) \leq q(x)$  für alle  $x \in F$ . Dann gibt es ein lineares Funktional  $g$  auf ganz  $E$  mit folgenden Eigenschaften:*

- (i)  $f(x) = g(x)$  für alle  $x \in F$ .
- (ii)  $g(x) \leq q(x)$  für alle  $x \in E$ .

*Beweis.* Wir führen den Beweis in zwei Schritten durch. Im ersten zeigen wir, dass es für jedes  $x_0 \in E$  ein Funktional  $g$  mit den gewünschten Eigenschaften auf der Menge  $H = \{\lambda x_0 + y \mid \lambda \in \mathbb{R}, y \in F\}$ , also auf allen „Linien“ durch Punkte in  $F$  mit der Richtung  $x_0$  gibt. Im zweiten Schritt zeigen wir, dass diese Funktionale dann auf dem ganzen Raum übereinstimmen.

Sei also zunächst  $x_0 \in E$ , aber  $x_0 \notin F$  und  $H = \{\lambda x_0 + y \mid \lambda \in \mathbb{R}, y \in F\}$ . Wir wollen zeigen, dass  $f$  eine Fortsetzung zu einem linearen Funktional  $g$  auf  $H$  besitzt, das  $g(x) \leq q(x)$  für alle  $x \in H$  erfüllt. Ist  $g$  ein solches lineares Funktional auf  $H$ , so muss sicher gelten

$$g(\lambda x_0 + y) = \lambda g(x_0) + g(y) = \lambda g(x_0) + f(y),$$

denn auf  $F$  stimmt  $g$  ja mit  $f$  überein. Wenn wir  $c := g(x_0)$  so wählen können, dass  $g(x) \leq q(x)$  für alle  $x \in H$  ist, dann können wir obige Gleichung einfach zur Definition von  $g$  nutzen. Seien hierzu  $u, v \in F$ . Es gilt dann

$$f(u) - f(v) = f(u - v) \leq q(u - v) = q(u + x_0 - x_0 - v) \leq q(u + x_0) + q(-x_0 - v).$$

Hieraus folgt  $-q(-x_0 - v) - f(v) \leq q(x_0 + u) - f(u)$ , also auch

$$m_1 := \sup_{v \in F} (-q(-x_0 - v) - f(v)) \leq \inf_{u \in F} (q(x_0 + u) - f(u)) := m_2.$$

Wähle ein  $c \in [m_1, m_2]$  und setze  $g(x_0) = c$ . Insbesondere gilt dann  $c \leq q(x_0 + \frac{y}{\lambda}) - f(\frac{y}{\lambda})$ , indem wir in der obigen Gleichung  $u = \frac{y}{\lambda}$  gesetzt haben. Mit  $z = \lambda x_0 + y$  folgt dann für  $\lambda > 0$

$$\begin{aligned} g(z) &= f(y) + \lambda c \leq f(y) + \lambda \left( q\left(x_0 + \frac{y}{\lambda}\right) - f\left(\frac{y}{\lambda}\right) \right) \\ &= f(y) + q\left(\lambda x_0 + \frac{\lambda y}{\lambda}\right) - f\left(\frac{\lambda y}{\lambda}\right) = f(y) + q(z) - f(y) = q(z), \end{aligned}$$

Analog folgt mit die Aussage  $g(z) \leq q(z)$  für  $\lambda < 0$ , wenn man beachtet, dass  $f(y) + \lambda c = f(y) + (-\lambda)(-c)$  und  $-c \in [-m_2, -m_1]$  ist. Dann ist nämlich  $-c \leq q(-x_0 - v) - f(v)$  für alle  $v \in F$ , insbesondere für  $v = \frac{y}{\lambda}$ . Aus diesen beiden Ungleichungen entnehmen wir, dass  $g(z) \leq q(z)$  für alle  $z \in H$  gilt.

Den zweiten Schritt beweisen wir mit dem Lemma von Zorn. Wir definieren eine halbgeordnete Menge  $\mathcal{M}$  durch  $\mathcal{M} = \{(h, H) \mid \}$ , wobei  $H$  ein linearer Teilraum von  $E$  ist, der  $F$  enthält, also  $F \subseteq H$ , und  $h$  ein lineares Funktional auf  $H$  mit  $f(x) = h(x)$  für alle  $x \in F$  sei. Die Halbordnungsrelation definieren wir wie folgt. Es gelte  $(h_1, H_1) \leq (h_2, H_2)$  genau dann, wenn  $H_1$  ein linearer Teilraum von  $H_2$  ist und die Einschränkung von  $h_2$  auf  $H_1$  gerade  $h_1$  ergibt, also  $h_2|_{H_1} = h_1$ . Sei  $\mathcal{K}$  eine total geordnete Teilmenge von  $\mathcal{M}$ . Wir müssen zeigen, dass für  $\mathcal{K}$  eine obere Schranke in  $\mathcal{M}$  existiert.

Definiere dafür  $H_s = \bigcup_{(h,H) \in \mathcal{K}} H$  und  $h_s(x) = h(x)$ , falls  $(h, H) \in \mathcal{K}$  und  $x \in H$  ist. Weil  $\mathcal{K}$  total geordnet ist, ist  $H_s$  ein Vektorraum und  $h_s$  ist wohldefiniert, denn alle in Frage kommenden Funktionale haben in  $x$  denselben Wert. Im folgenden wird gezeigt, dass  $(h_s, H_s) \in \mathcal{M}$  ist und  $(h_s, H_s)$  eine obere Schranke für  $\mathcal{K}$  darstellt. Nach dem Lemma von Zorn besitzt damit  $\mathcal{M}$  ein maximales Element  $(h_m, H_m)$ . Dann muss  $H_m = E$  sein, denn ansonsten würde ein Vektorraum  $H_0$  mit  $H_m \subset H_0 \subseteq E$  und eine Erweiterung  $h_0$  von  $h_m$  mit  $h_0(x) \leq q(x)$  für  $x \in H_0$  existieren, was einen Widerspruch zur Maximalität darstellt. Mit  $g = h_m$  erhalten wir also unser gesuchtes Funktional.  $\square$

Wir geben also auf einem Teilraum  $F \subseteq E$  ein lineares Funktional  $f$  vor, das durch ein auf ganz  $E$  definiertes sublineares Funktional beschränkt ist. Dann kann man nach dem eben bewiesenen Satz das Funktional  $f$  auf den ganzen Raum so erweitern, dass es immer noch überall durch das sublineare Funktional beschränkt ist. Wie angekündigt beweisen wir nun dasselbe Theorem für den Fall, dass eine Halbnorm statt eines sublinearen Funktionals auf  $E$  gegeben ist. Es handelt sich lediglich eine Folgerung, da der Hauptteil des Beweises vom vorigen Satz übernommen wird.

**Korollar 2.5.10.** Sei  $F$  ein linearer Teilraum des  $\mathbb{R}$ -Vektorraumes  $E$  und  $p$  eine Halbnorm auf  $E$ .  $f$  sei ein lineares Funktional auf  $F$  mit  $|f(x)| \leq p(x)$  für alle  $x \in F$ . Dann existiert ein lineares Funktional  $g$  auf  $E$  mit folgenden Eigenschaften:

- (i)  $f(x) = g(x)$  für alle  $x \in F$ .
- (ii)  $|g(x)| \leq p(x)$  für alle  $x \in E$ .

*Beweis.* Da  $p$  als Halbnorm ebenfalls ein sublineares Funktional ist, können wir nach Satz 2.5.9 eine Erweiterung  $g$  von  $f$  finden. Es gilt  $g(x) \leq p(x)$  für alle  $x \in E$ , also auch  $-g(x) = g(-x) \leq p(-x) = p(x)$  aufgrund der Halbnormeigenschaft von  $p$ .  $\square$

Das Hahn-Banach-Theorem kann man auch für komplexe Vektorräume formulieren. Dies wollen wir im folgenden tun. Sei dazu  $E$  ein komplexer Vektorraum und  $f : E \rightarrow \mathbb{C}$  ein lineares Funktional auf  $E$ , so dass also für alle  $\lambda_1, \lambda_2 \in \mathbb{C}$  und  $x_1, x_2 \in E$  gilt:  $f(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 f(x_1) + \lambda_2 f(x_2)$ . Da sich  $f$  in Real- und Imaginärteil zerlegen lässt, gilt  $f(x) = f_1(x) + i f_2(x)$  mit  $f_1(x), f_2(x) \in \mathbb{R}$ . Wegen der Linearität von  $f$  folgt

$$f(ix) = f_1(ix) + i f_2(ix) = i f(x) = i f_1(x) - f_2(x),$$

woraus man mit Koeffizientenvergleich schließt, dass  $f_2(ix) = f_1(x)$  und  $f_1(ix) = -f_2(x)$  gilt. Damit lässt sich  $f$  schreiben als

$$f(x) = f_1(x) - i f_1(ix).$$

Es lässt sich also ein komplexes Funktional durch ein reelles ausdrücken, wobei insbesondere  $f_1 : E \rightarrow \mathbb{R}$  ein lineares Funktional auf dem als reell aufgefassten Vektorraum  $E$  ist. Sei umgekehrt  $f_1$  ein lineares Funktional auf dem reellen Vektorraum  $E$ . Dann definieren wir  $f(x) = f_1(x) - i f_1(ix)$  für  $x \in E$ . Dann zeigt man als Übungsaufgabe, dass  $f$  ein lineares Funktional auf dem komplexen Vektorraum  $E$  ist. Damit haben wir einen eindeutigen Zusammenhang zwischen  $\mathbb{C}$ -linearen Funktionalen  $f$  und den entsprechenden  $\mathbb{R}$ -linearen Funktionalen  $f_1$  auf dem als komplex oder reell aufgefassten Vektorraum  $E$ . Mit diesen Vorbereitungen sind wir nun in der Lage, den Satz von Hahn-Banach im komplexen Fall zu beweisen.

**Satz 2.5.11** (Satz von Hahn-Banach, komplexer Fall). *Sei  $E$  ein komplexer Vektorraum und  $p$  eine Halbnorm auf  $E$ . Weiter sei  $f$  ein lineares Funktional auf einem linearen Teilraum  $F$  von  $E$ , für das  $|f(x)| \leq p(x)$  für alle  $x \in F$  gilt. Dann existiert ein lineares Funktional  $g$  auf  $E$  mit den Eigenschaften*

$$(i) \quad f(x) = g(x) \text{ für } x \in F,$$

$$(ii) \quad |g(x)| \leq p(x) \text{ für } x \in E.$$

*Beweis.* Sei  $f(x) = f_1(x) - if_1(ix)$  für ein  $\mathbb{R}$ -lineares Funktional auf dem reellen Vektorraum  $F$ . Da  $|f(x)| = |f_1(x) - if_1(ix)| \leq p(x)$ , gilt auch  $|f_1(x)| \leq p(x)$  für  $x \in F$ . Nach dem Hahn-Banach-Theorem für reelle Vektorräume (Satz 2.5.9) existiert ein lineares Funktional  $g_1$  auf dem reellen Vektorraum  $E$  mit  $f_1(x) = g_1(x)$  für  $x \in F$  und  $|g_1(x)| \leq p(x)$  für alle  $x \in E$ . Sei  $g(x) = g_1(x) - ig_1(ix)$  für  $x \in E$ . Dann ist  $g$  ein lineares Funktional auf dem komplexen Vektorraum  $E$ . Wie wir in der Vorbereitung zu diesem Satz sahen, gilt  $f = g$  auf  $F$ . Zu zeigen bleibt noch  $|g| \leq p$  auf ganz  $E$ . Sei dazu  $x \in E$  fest. Wähle ein  $\varphi \in \mathbb{R}$  mit  $|g(x)| = e^{i\varphi}g(x) = g(e^{i\varphi}x)$ . Dann gilt

$$|g(x)| = g(e^{i\varphi}x) = |g_1(e^{i\varphi}x)| \leq p(e^{i\varphi}x) = p(x),$$

wobei wir Satz 2.5.9 verwendet und die Halbnormeigenschaften von  $p$  ausgenutzt haben.  $\square$

Diese Ergebnisse wollen wir nun auf stetige lineare Funktionale anwenden. Sei dazu  $(E, \|\cdot\|)$  ein normierter linearer Raum über einem Körper  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$ . Wir können dann einige wichtige Folgerungen beweisen, insbesondere die Existenz von Funktionalen mit angenehmen Eigenschaften. Wir gewinnen auch einige Aussagen über Dualräume.

**Satz 2.5.12.** *Sei  $F$  ein linearer Teilraum von  $E$  und  $f$  ein stetiges lineares Funktional auf  $F$ . Dann existiert ein stetiges lineares Funktional  $g$  auf  $E$  mit  $f(x) = g(x)$  für  $x \in F$  und  $\|f\| = \|g\|$ .*

*Beweis.* Da  $f \in F'$  ist, gilt  $|f(x)| \leq \|f\| \|x\|$  für alle  $x \in F$ . Sei  $p(y) = \|f\| \|y\|$ , wobei  $y \in E$ . Dann ist  $p$  eine Halbnorm und es gilt  $|f(x)| \leq p(x)$ , sofern  $x \in F$  ist. Sei  $g$  die Erweiterung von  $f$  nach Satz 2.5.11 oder 2.5.9. Dann gilt

$$|g(x)| \leq p(x) = \|f\| \|x\|$$

für  $x \in E$  und damit  $\|g\| \leq \|f\|$ . Trivialerweise ist auch  $\|f\| \leq \|g\|$ , da die Operatornorm das Supremum über die Einheitskugel ist, die bei Erweiterung des Raumes „größer“ wird oder gleich bleibt. Also ist  $\|f\| = \|g\|$ .  $\square$

**Korollar 2.5.13.** Sei  $x \in E$ . Dann existiert ein stetiges lineares Funktional  $f_x$  auf  $E$  mit  $f_x(x) = \|x\|$  und  $\|f_x\| = 1$ . Insbesondere ist  $f_x(x) \neq 0$ , wenn  $x \neq 0$  ist.

*Beweis.* Sei  $F = \mathbb{K} \cdot x$  und sei ein stetiges lineares Funktional  $f$  auf  $F$  definiert durch  $f(\lambda x) = \lambda \|x\|$ , wobei  $\lambda \in \mathbb{K}$ . Dann gilt  $\|f\| = 1$ . Sei weiter  $f_x = g$  die Erweiterung von  $f$  auf  $E$  nach Satz 2.5.12. Dann gilt auch  $\|f_x\| = 1$  und  $\|x\| = g(x) = f_x(x)$ .  $\square$

Das eben bewiesene Korollar beantwortet die eingangs gestellte Frage, ob es auf einem normierten linearen Raum  $E$  „genügend“ lineare Funktionale gibt. Man beachte, dass dafür keine weiteren Voraussetzungen an den Raum  $E$  gestellt werden müssen. Das folgende Korollar gibt uns hingegen einen Zusammenhang zwischen der Norm von  $x$  und dem Wert der normierten Funktionale in  $x$ .

**Korollar 2.5.14.** Es gilt  $\|x\| = \max \{|f(x)| \mid f \in E', \|f\| = 1\}$ .

*Beweis.* Wenn  $f$  ein stetiges lineares Funktional ist mit  $\|f\| \leq 1$ , so gilt  $f(x) \leq \|f\| \|x\| \leq \|x\|$ . Deshalb haben wir  $\sup_{\|f\| \leq 1} |f(x)| \leq \|x\|$ . Sei  $f = f_x$  aus Folgerung 2.5.13. Dann gilt natürlich  $\|x\| = f_x(x)$  und damit wird das Supremum in  $f_x$  angenommen und ist deshalb sogar ein Maximum.  $\square$

**Korollar 2.5.15.** Für  $x \in E$  sei  $h_x(f) = f(x)$  für alle  $f \in E'$ . Dann ist  $h_x \in (E')' = E''$ . Die Abbildung  $x \mapsto h_x$  ist eine bijektive isometrische lineare Abbildung von  $E$  auf einen linearen Teilraum  $E_0$  von  $E''$ .

*Beweis.*  $\|h_x\| = \sup_{\|f\| \leq 1} |f(x)| = \|x\|$  nach Folgerung 2.5.14. □

Man beachte, dass keinesfalls  $E_0 = E''$  gelten muss. Es gibt jedoch eine Klasse von Räumen, die diese Bedingung erfüllen. Solche Räume nennen wir reflexiv. Wir identifizieren fortan  $x \in E$  mit  $h_x \in E''$ . Damit ist  $E$  ein linearer Teilraum von  $E''$ .

**Definition 2.5.16.** Ein normierter linearer Raum  $(E, \|\cdot\|)$  heißt *reflexiv*, wenn  $E_0 = E''$  beziehungsweise nach Identifizierung  $E = E''$  gilt, das heißt wenn jedes stetige lineare Funktional  $h$  auf  $E'$  von der Form  $h(f) = h_x(f)$  ist, wobei  $f \in E'$  und  $x \in E$  ist.

Nach Hilberträumen, bei denen bis auf komplexe Konjugation die Räume  $E$  und  $E'$  identifiziert werden können, sind die reflexiven Räume mit  $E = E''$  also die „nächstbesten“. Es gilt immer

$$E \subseteq E'' \subseteq E^{(4)} \subseteq E^{(6)} \subseteq \dots$$

sowie

$$E' \subseteq E''' \subseteq E^{(5)} \subseteq \dots$$

Man kann zeigen, dass im Falle der Reflexivität von  $E$  überall Gleichheit herrscht, während niemals Gleichheit gilt, wenn  $E$  nicht reflexiv ist.

**Beispiel 2.5.17.**

- (a) Wir wissen bereits, dass  $L^p(X, \mu)'' \cong L^q(X, \mu)' \cong L^p(X, \mu)$ , falls  $p \in (1, \infty)$  und  $\frac{1}{p} + \frac{1}{q} = 1$  ist. Damit ist  $L^p(X, \mu)$  reflexiv für  $p \in (1, \infty)$ .
- (b)  $L^1(X, \mu)$  und  $L^\infty(X, \mu)$  sind für allgemeines  $X$  und  $\mu$  nicht reflexiv. Als Übungsaufgabe zeigt man, dass speziell  $\ell^1(\mathbb{N})$  nicht reflexiv ist.

Nun soll es uns um die Trennung konvexer Mengen gehen. Auch hier können wir Gewinn aus dem Hahn-Banach-Theorem ziehen. Zuerst werden wir dazu das Minkowski-Funktional auf geeigneten Mengen definieren. Sei  $E$  ein reeller Vektorraum und  $M \subseteq E$ . Die Definition einer konvexen Menge gaben wir zwar bereits in 1.1.3, wiederholen sie hier aber aus Gründen der Übersichtlichkeit.

**Definition 2.5.18.**  $M$  heißt

- (i) *konvex*, wenn aus  $x, y \in M$  und  $\lambda \in [0, 1]$  stets auch  $\lambda x + (1 - \lambda)y \in M$  folgt,
- (ii) *kreisförmig*, wenn aus  $x \in M$  und  $\lambda \in \mathbb{R}$  mit  $|\lambda| \leq 1$  stets  $\lambda x \in M$  folgt,
- (iii) *absorbierend*, wenn  $E = \bigcup_{\lambda > 0} \lambda M$ .

Bei einer konvexen Menge  $M$  liegt zu zwei Punkten  $x$  und  $y$  also auch immer die Verbindungsstrecke zwischen  $x$  und  $y$  in  $M$ . Ist  $M$  kreisförmig, so liegt zu  $x$  auch die Verbindungsstrecke von  $x$  und  $-x$  durch  $0$  in  $M$ . Dass  $M$  absorbierend ist, ist, wie man zeigen kann, gleichwertig dazu, dass  $M$  eine Umgebung um die Null enthält. Eine absorbierende Menge kann also stets so weit „aufgeblasen“ werden, dass sie ein beliebiges  $x \in E$  enthält und dieses durch weiteres Aufblasen auch nicht wieder verliert, es also „absorbiert“.

Sei  $K$  eine absorbierende Teilmenge von  $E$ . Für  $x \in E$  definieren wir

$$p_K(x) = \inf \{ \lambda > 0 \mid x \in \lambda K \}.$$

Diese Definition ist sinnvoll, denn weil  $K$  absorbierend ist, gibt es zu jedem  $x \in E$  ein  $\lambda > 0$  mit  $x \in \lambda K$ .  $p_K$  heißt *Minkowski-Funktional* der Menge  $K$ .

**Satz 2.5.19.** Sei  $K$  eine konvexe absorbierende Teilmenge von  $E$ , insbesondere  $0 \in K$ . Dann gilt

- (i)  $p_K$  ist ein nichtnegatives sublineares Funktional auf  $E$ .
- (ii) Wenn  $K$  kreisförmig ist, dann ist  $p_K$  eine Halbnorm.
- (iii) Es gelten die Inklusionen  $\{x \in E \mid p_K(x) < 1\} \subseteq K \subseteq \{x \in E \mid p_K(x) \leq 1\}$ .
- (iv) Wenn  $(E, \|\cdot\|)$  ein reeller normierter linearer Raum und  $K$  offen ist, dann gilt

$$K = \{x \in E \mid p_K(x) < 1\}.$$

*Beweis.* (i): Wir beweisen  $p_K(x+y) \leq p_K(x) + p_K(y)$ . Sei dazu  $x \in \lambda K$  und  $y \in \mu K$  mit  $\lambda > 0$  und  $\mu > 0$ . Dann gilt

$$\frac{1}{\lambda + \mu}(x+y) = \frac{\lambda}{\lambda + \mu} \frac{x}{\lambda} + \frac{\mu}{\lambda + \mu} \frac{y}{\mu} \in K,$$

da  $\frac{x}{\lambda} \in K$  und  $\frac{y}{\mu} \in K$  und  $K$  konvex ist. Also gilt  $x+y \in (\lambda+\mu)K$  und damit ist  $p_K(x+y) \leq \lambda + \mu$ . Bildet man nun das Infimum über  $\lambda$  und  $\mu$  auf der rechten Seite unter der Nebenbedingung  $x \in \lambda K$  und  $y \in \mu K$ , so erhält man daraus  $p_K(x+y) \leq p_K(x) + p_K(y)$ . Man zeigt weiterhin leicht, dass  $p_K(\alpha x) = \alpha p_K(x)$  für  $\alpha \geq 0$  gilt.

(ii): Wir müssen zeigen, dass  $p_K(\alpha x) = |\alpha| p_K(x)$  für  $x \in E$  und  $\alpha \in \mathbb{R}$  gilt. Für  $\alpha = 0$  ist das klar. Sei deshalb  $\alpha \neq 0$ . Für  $\lambda > 0$  mit  $x \in \frac{\lambda}{\alpha} K$  gilt, da  $K$  kreisförmig ist, auch  $x \in \frac{\lambda}{|\alpha|} K$ . Setze  $\lambda' = \frac{\lambda}{|\alpha|}$ . Dann gilt

$$p_K(\alpha x) = \inf \{\lambda > 0 \mid \alpha x \in \lambda K\} = \inf \{|\alpha| \lambda' > 0 \mid x \in \lambda' K\} = |\alpha| \inf \{\lambda' > 0 \mid x \in \lambda' K\} = |\alpha| p_K(x).$$

(iii): Sei  $p_K(x) < 1$ . Dann gibt es ein  $\lambda < 1$  mit  $x \in \lambda K$ . Es gilt weiter

$$x = \lambda \frac{x}{\lambda} + (1-\lambda) \cdot 0,$$

das heißt  $x$  ist als Konvexkombination von Elementen  $\frac{x}{\lambda}$  und  $0$  aus  $K$  ebenfalls in  $K$  enthalten und damit ist  $\{x \in E \mid p_K(x) < 1\} \subseteq K$ . Wenn  $x \in K$  ist, dann ist  $p_K(x) \leq 1$  und damit folgt  $K \subseteq \{x \in E \mid p_K(x) \leq 1\}$ .

(iv): Sei  $p_K(x) \geq 1$ . Dann ist  $x \notin \lambda K$  für alle  $\lambda \in (0, 1)$  und damit ist für diese  $\lambda$  auch  $\frac{x}{\lambda} \in E \setminus K$ . Weil  $K$  offen ist, ist  $E \setminus K$  abgeschlossen. Deshalb gilt  $\lim_{\lambda \rightarrow 1} \frac{x}{\lambda} = x \in E \setminus K$ , also  $x \notin K$ . Damit gilt  $K \subseteq \{x \in E \mid p_K(x) < 1\}$ . Weil  $\{x \in E \mid p_K(x) < 1\} \subseteq K$  nach (iii) gilt, folgt die behauptete Gleichheit.  $\square$

Im folgenden sei  $(E, \|\cdot\|)$  ein normierter linearer Raum. Wir bereiten nun den Trennungssatz für konvexe Mengen vor. Wir werden konvexe Mengen nicht wie im einführenden Motivationsbeispiel des  $\mathbb{R}^2$  durch Geraden trennen, sondern durch stetige lineare Funktionale. Im vorbereitenden Lemma werden wir eine konvexe Menge, die die Null nicht enthält, durch ein solches Funktional von der Null trennen.

**Lemma 2.5.20.** Sei  $M$  eine offene, konvexe Menge im normierten linearen Raum  $(E, \|\cdot\|)$  und  $0 \notin M$ . Dann existiert ein stetiges lineares Funktional  $f \in E'$  mit  $f(x) < 0$  für alle  $x \in M$ .

*Beweis.* Fixiere zunächst ein beliebiges  $x \in M$ . Sei  $U = M - x = \{m - x \mid m \in M\}$  die um  $x$  verschobene Menge  $M$ . Die Menge  $U$  ist ebenfalls offen und konvex und es gilt  $0 = x - x \in U$ . Sei  $y := -x$ . Weil  $0 \notin M$  ist, ist  $y \notin U$ .  $p_U$  sei das Minkowski-Funktional auf  $U$ . Da  $U$  offen und absorbierend ist, ist  $p_U$  nach Satz 2.5.19 ein nichtnegatives sublineares Funktional.

Sei  $F = \mathbb{R} \cdot y$ . Dann ist  $F$  ein eindimensionaler Teilraum von  $M$ . Wir definieren auf  $F$  ein lineares Funktional  $h$  durch  $h(ty) = t p_U(y)$  für alle  $t \in \mathbb{R}$ . Für  $t \leq 0$  gilt insbesondere  $h(ty) = t p_U(y) \leq 0 \leq p_U(y)$ . Für  $t > 0$  hingegen gilt  $h(ty) = t p_U(y) = p_U(ty)$ , das heißt  $h(z) \leq p_U(z)$  für alle  $z \in F$ . Nach dem Hahn-Banach-Theorem für sublineare Funktionale (Satz 2.5.9) existiert dann ein Funktional  $f$  auf  $E$  mit

- (i)  $f(z) = h(z)$  für  $z \in F$  und  
(ii)  $f(v) \leq p_U(v)$  für alle  $v \in E$ .

Wir zeigen, dass  $f$  stetig auf  $E$  ist. Da  $U$  offen und  $0 \in U$  ist, existiert ein  $\delta > 0$  mit  $K_\delta(0) \subseteq U$ . Daraus folgt, dass  $p_{K_\delta(0)} \geq p_U$  auf ganz  $E$  ist. Nach Satz 2.5.19 (iv) ist  $p_{K_\delta(0)}(x) < 1$  genau dann, wenn  $x \in K_\delta(0)$  ist, das heißt wenn  $\|x\| < \delta$  ist. Damit folgt  $p_{K_\delta(0)}(x) = \frac{1}{\delta} \|x\|$  für alle  $x \in E$ , also

$$f(v) \leq p_U(v) \leq p_{K_\delta(0)}(v) = \frac{1}{\delta} \|v\|,$$

für alle  $v \in E$ . wegen  $-f(v) = f(-v) \leq \frac{1}{\delta} \|-v\|$  gilt deshalb  $|f(v)| \leq \frac{1}{\delta} \|v\|$  und damit ist  $f$  beschränkt, also stetig.

Da  $y \notin U$  und  $U = \{v \in E \mid p_U(v) < 1\}$  ist, folgt nach Satz 2.5.19 (iv), dass  $p_U(y) \geq 1$  sein muss. Sei  $z \in M$ . Dann ist  $u := z - x \in U$ , also

$$f(z) = p_U(z) \leq p_U(u) - p_U(y) < 1 - 1 = 0.$$

□

**Satz 2.5.21** (Erster Trennungssatz für konvexe Mengen). *Seien  $M_1$  und  $M_2$  zwei disjunkte konvexe Mengen, wobei  $M_1$  offen in  $(E, \|\cdot\|)$  sei. Dann gibt es ein stetiges lineares Funktional  $f \in E'$  mit  $f(x_1) < f(x_2)$  für alle  $x_1 \in M_1$  und alle  $x_2 \in M_2$ .*

*Beweis.* Sei  $M = M_1 - M_2 = \{m_1 - m_2 \mid m_1 \in M_1, m_2 \in M_2\}$ . Als Vereinigung offener, konvexer Mengen ist  $M = \bigcup_{x_2 \in M_2} \{M_1 - x_2\}$  offen und konvex. Wegen  $M_1 \cap M_2 = \emptyset$  gilt  $0 \notin M$ . Damit sind die Voraussetzungen für Lemma 2.5.20 erfüllt und es gibt ein  $f \in E'$  mit  $f(x) < 0$  für alle  $x \in M$ . Seien nun  $x_1 \in M_1$  und  $x_2 \in M_2$ . Aus  $f(x_1 - x_2) < 0$  folgt dann  $f(x_1) < f(x_2)$ . □

**Satz 2.5.22** (Zweiter Trennungssatz für konvexe Mengen).  *$M$  sei eine abgeschlossene, konvexe, nicht-leere Teilmenge von  $(E, \|\cdot\|)$ . Sei weiter  $x \in E$ , aber  $x \notin M$ . Dann existiert ein lineares Funktional  $f \in E'$  und ein  $c > 0$  mit  $f(x) < f(x) + c \leq f(z)$  für alle  $z \in M$ .*

*Beweis.* Da  $M$  abgeschlossen und  $x \notin M$  ist, gibt es ein  $\varepsilon > 0$ , so dass  $K_\varepsilon(x) \cap M = \emptyset$  ist. Weil  $K_\varepsilon$  eine offene und konvexe Menge ist, gibt es nach dem ersten Trennungssatz, Satz 2.5.21 mit  $M_1 = K_\varepsilon(x)$  und  $M_2 = M$  ein stetiges lineares Funktional  $f \in E'$  mit  $f(x + u) < f(z)$  für alle  $u \in K_\varepsilon(0)$  und alle  $z \in M$ . Ersetzt man  $u$  durch  $-u$ , so folgt aus beiden Ungleichungen  $f(x) + |f(u)| < f(z)$ . Nun bilden wir das Supremum über alle  $u \in K_\varepsilon$  und erhalten

$$f(x) + \|f\| \varepsilon \leq f(z).$$

Die Konstante  $c$  ist dann  $c := \|f\|$ . □

Der Unterschied zum ersten Trennungssatz besteht hauptsächlich darin, dass  $f$  hier den Punkt  $x$  und die Menge  $M$  strikt trennt. Dafür müssen wir allerdings auch stärkere Voraussetzungen an die Menge stellen, genauer gesagt die Abgeschlossenheit. Im ersten Trennungssatz könnte dafür hingegen  $\sup_{x \in M_1} f(x) = \inf_{x \in M_2} f(x)$  sein.

**Bemerkung 2.5.23.**



- (i) Satz 2.5.22 besitzt folgende Verallgemeinerung:  $M_1$  sei eine abgeschlossene, konvexe Teilmenge und  $M_2$  sei eine kompakte, konvexe Teilmenge von  $(E, \|\cdot\|)$ . Es gelte  $M_1 \cap M_2 = \emptyset$ . Dann existiert ein lineares Funktional  $f \in E'$  und ein  $c > 0$  mit  $f(x_1) < f(x_1) + c \leq f(x_2)$  für  $x_1 \in M_1$  und  $x_2 \in M_2$ . Damit man wie im ersten Trennungssatz zwei Mengen, diesmal aber strikt, trennen kann, werden also sehr starke Voraussetzungen an die Mengen benötigt, namentlich Abgeschlossenheit und Kompaktheit.
- (ii) Es gibt ebenso Trennungssätze in komplexen Räumen. Deren Ungleichungen betreffen die Realteile der Funktionale.

Eine Verallgemeinerung des Satzes 2.5.21 ist das Eidelheit-Theorem. Wir werden es ohne Beweis angeben, benötigen vorher aber noch eine kurze Definition.

**Definition 2.5.24.**  $M$  sei eine Teilmenge eines reellen Vektorraumes  $E$ .  $x_0$  heißt *algebraisch innerer Punkt* von  $M$ , wenn ein  $\delta_x > 0$  existiert, so dass  $x_0 + \lambda x \in M$  für alle  $\lambda \in \mathbb{R}$  mit  $|\lambda| < \delta_x$  ist. Die Menge aller algebraischen inneren Punkte von  $M$  bezeichnen wir mit  $M^0$ .

**Lemma 2.5.25** (Eidelheit-Theorem).  $A, B$  seien konvexe Mengen im reellen Vektorraum  $E$ . Es sei  $A^0 \neq \emptyset$ , das heißt  $A$  besitze mindestens einen algebraisch inneren Punkt, und  $A \cap B = \emptyset$ . Dann gibt es ein lineares Funktional  $f$  auf  $E$  mit  $f(a) \leq f(b)$  für alle  $a \in A$  und alle  $b \in B$ .



# Kapitel 3

## Beschränkte lineare Operatoren im Hilbertraum

### 3.1 Adjungierte, selbstadjungierte und unitäre Operatoren

#### 3.1.1 Adjungierte Operatoren

Im folgenden sei  $(H, \langle \cdot, \cdot \rangle)$  immer ein komplexer Hilbertraum.

**Satz 3.1.1** (Definition und Eigenschaften adjungierter Operatoren).

(i) Sei  $T \in \mathcal{L}(H)$ . Dann existiert ein eindeutig bestimmter linearer Operator  $S$  von  $H$  nach  $H$  mit  $\langle Tx, y \rangle = \langle x, Sy \rangle$  für alle  $x, y \in H$ .  $S$  heißt der zu  $T$  adjungierte Operator. Wir schreiben ab sofort  $S = T^*$ .

(ii)  $T^*$  ist beschränkt und es gilt  $\|T^*\| = \|T\| = \|T^*T\|^{1/2}$

(iii) Für  $T_1, T_2 \in \mathcal{L}(H)$  und  $\lambda_1, \lambda_2 \in \mathbb{C}$  gilt:

$$(\lambda_1 T_1 + \lambda_2 T_2)^* = \overline{\lambda_1} T_1^* + \overline{\lambda_2} T_2^*.$$

Weiterhin ist  $(T_1^*)^* = T_1$  sowie  $(T_1 T_2)^* = T_2^* T_1^*$ .

*Beweis.* (i): Seien  $x \in H$  fest. Wir definieren ein lineares Funktional  $F_x(y) := \langle Ty, x \rangle$  für  $y \in H$ . Mit Hilfe der Cauchy-Schwarz-Ungleichung (1.3.3) sehen wir, da  $T$  beschränkt ist:

$$|F_x(y)| = |\langle Ty, x \rangle| \leq \|Ty\| \|x\| \leq \|T\| \|y\| \|x\| = (\|T\| \|x\|) \|y\|.$$

Also ist  $F_x$  stetig und nach dem 2. Satz von Riesz (Satz 1.4.8) gibt es ein  $x^* \in H$  mit  $F_x(y) = \langle y, x^* \rangle$  für alle  $y \in H$ . Wegen  $\langle Ty, x \rangle = \langle y, x^* \rangle$  und  $\langle x, Ty \rangle = \langle x^*, y \rangle$  ist die Zuordnung  $S : x \mapsto x^*$  eine wohldefinierte lineare Abbildung von  $H$  nach  $H$ . Damit haben wir die Existenz des adjungierten Operators gezeigt, wir beweisen nun seine Eindeutigkeit. Sei  $S_1$  ein anderer linearer Operator, für den  $\langle Ty, x \rangle = \langle y, S_1 x \rangle$  für alle  $x, y \in H$  gelte. Dann folgt  $\langle Ty, x \rangle = \langle y, Sx \rangle = \langle y, S_1 x \rangle$ , also  $\langle y, Sx - S_1 x \rangle = 0$  für alle  $x, y \in H$ . Mit  $y = Sx - S_1 x$  folgt dann aber  $\|y\| = 0$  und somit  $Sx = S_1 x$ .

Die Aussagen (ii) und (iii) beweisen wir gemeinsam. Dafür zeigen wir zunächst, dass  $T^*$  beschränkt ist und  $(T^*)^* = T$  gilt. Die Beschränktheit ergibt sich durch die Betrachtung von

$$|\langle Ty, x \rangle| = |\langle y, T^* x \rangle| \leq \|T\| \|x\| \|y\|$$

und wir erhalten daraus die Ungleichung  $\|T^*\| \leq \|T\|$ . Wir wissen, dass  $(T^*)^*$  existiert. Nun gilt weiterhin:

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

und

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

für alle  $x, y \in H$ . Hiermit folgt aus der Eindeutigkeitsaussage in (i), dass  $T = (T^*)^*$  sein muss. Die Aussage  $\|T\| = \|T^*\|$  bekommen wir leicht, indem wir in der Ungleichung  $\|T^*\| \leq \|T\|$  den Operator  $T$  mit  $T^*$  ersetzen. Dann ist nämlich  $\|(T^*)^*\| = \|T\| \leq \|T^*\|$  und daraus folgt die Gleichheit. Zu guter letzt gilt nach der Cauchy-Schwarz-Ungleichung

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

und somit ergibt sich  $\|T\|^2 \leq \|T^*T\| \leq \|T^*\| \|T\| \leq \|T\|^2$ , wobei wir wieder die etwas weiter oben bewiesene Ungleichung  $\|T^*\| \leq \|T\|$  verwendeten. Damit haben wir auch  $\|T\| = \|T^*T\|^{1/2}$  gezeigt.  $\square$

**Satz 3.1.2** (Satz von Hellinger-Toeplitz). *Seien  $A, B$  lineare Operatoren auf einem Hilbertraum  $H$  mit  $\langle Ax, y \rangle = \langle x, By \rangle$  für alle  $x, y \in H$ . Dann sind  $A$  und  $B$  beschränkte Operatoren und es gilt  $A^* = B$ .*

*Beweis.* Nach dem Satz vom abgeschlossenen Graphen (Satz 2.4.6) reicht es zu zeigen, dass  $A$  ein abgeschlossener Operator ist. Dann folgt, dass  $A$  beschränkt ist und es gilt  $A^* = B$  nach Satz 3.1.1. Sei also  $(x_n)_{n \in \mathbb{N}}$  eine Folge aus  $H$  mit  $x_n \rightarrow x$  und  $Ax_n \rightarrow y$  mit  $x, y \in H$ . Wir müssen also zeigen, dass  $y = Ax$  ist. Es gilt  $\langle A(x_n - x), z \rangle = \langle x_n - x, Bz \rangle$  für jedes  $z \in H$ . Mit dem Grenzübergang  $n \rightarrow \infty$  folgt dann:  $\langle y - Ax, z \rangle = \langle x - x, Bz \rangle = 0$  für alle  $z \in H$ . Daraus folgt, dass  $y - Ax = 0$  sein muss und somit ist  $y = Ax$ .  $\square$

Nun erweitern wir unseren Strukturhorizont mit einigen Definitionen nützlicher Algebren. Eine Banachalgebra ist eine Struktur, die Banach selbst nie zu Gesicht bekommen hat, da sie von Gelfand nach Banachs Tod entwickelt wurde.

**Definition 3.1.3** (\*-Algebra). Eine \*-Algebra  $A$  ist eine Algebra, die mit einer Algebreninvolution  $a \mapsto a^*$  versehen ist, das heißt  $A$  ist ein komplexer Vektorraum und ein Ring, für den gilt:

- (i)  $(\lambda a)b = \lambda(ab) = a(\lambda b)$  für alle  $\lambda \in \mathbb{C}$  und alle  $a, b \in A$ .
- (ii)  $a \mapsto a^*$  ist eine lineare Abbildung von  $A$  nach  $A$  mit den Eigenschaften
  - 1)  $(\alpha_1 a_1 + \alpha_2 a_2)^* = \overline{\alpha_1} a_1^* + \overline{\alpha_2} a_2^*$  für alle  $\alpha_1, \alpha_2 \in \mathbb{C}$  und alle  $a_1, a_2 \in A$ ,
  - 2)  $(a^*)^* = a$  für alle  $a \in A$ ,
  - 3)  $(a_1 a_2)^* = a_2^* a_1^*$  für alle  $a_1, a_2 \in A$ .

**Definition 3.1.4** (Banachalgebra). Eine Banachalgebra ist eine Algebra mit einer Norm  $\|\cdot\|$ , so dass gilt:

- (i)  $(A, \|\cdot\|)$  ist ein Banachraum,
- (ii)  $\|ab\| \leq \|a\| \|b\|$  für alle  $a, b \in A$ .

**Definition 3.1.5** ( $C^*$ -Algebra). Eine  $C^*$ -Algebra ist eine Banachalgebra  $(A, \|\cdot\|)$ , die außerdem eine \*-Algebra ist und für die  $\|a^*a\| = \|a\|^2$  gilt. Diese letzte Eigenschaft wird auch die  $C^*$ -Bedingung genannt.

**Beispiel 3.1.6.**

- (a) Sei  $H$  ein Hilbertraum. Dann ist  $A = \mathcal{L}(H)$  mit der Operatornorm  $\|\cdot\|$  und der Abbildung  $T \mapsto T^*$  eine  $C^*$ -Algebra.
- (b) Sei  $X$  ein kompakter metrischer Raum. Wir betrachten den Raum der stetigen Funktionen auf  $X$  mit dem Operator  $(f^*)(x) = \overline{f(x)}$  und der gängigen Norm  $\|f\| = \sup_{x \in X} |f(x)|$ . Dieser Raum ist eine kommutative  $C^*$ -Algebra mit Einselement. Die  $C^*$ -Bedingung können wir leicht nachrechnen:

$$\|f^*f\| = \sup_{x \in X} |f^*f(x)| = \sup_{x \in X} |f(x)^2| = \sup_{x \in X} |f(x)|^2 = \|f\|^2.$$

Nach dem *Theorem von Gelfand* gilt weiterhin: Zu jeder kommutativen  $C^*$ -Algebra  $A$  mit Einselement existiert ein kompakter topologischer Raum  $X$  derart, dass  $A$  als  $C^*$ -Algebra isomorph zur  $C^*$ -Algebra der stetigen Funktionen  $\mathcal{C}(X \rightarrow \mathbb{C})$  ist. Auf den Beweis dieses Satzes werden wir aus Zeitgründen verzichten.

### 3.1.2 Selbstadjungierte Operatoren

**Definition 3.1.7** (Selbstadjungierter Operator). Ein Operator  $T \in \mathcal{L}(H)$  eines Hilbertraums  $H$  heißt *selbstadjungiert*, wenn  $T = T^*$  ist.

Selbstadjungierte Operatoren spielen ungeachtet ihres sehr theoretischen mathematischen Hintergrundes in der Physik eine überaus wichtige Rolle. Wie wir sehen werden, besitzen diese Operatoren lediglich reelle Eigenwerte, weshalb sie sich in der Quantenmechanik zur Darstellung von Messgrößen, sogenannten *Observablen* verwenden lassen. Die dort betrachteten Operatoren sind jedoch unbeschränkt.

**Definition 3.1.8** (Normaler Operator). Ein Operator  $T \in \mathcal{L}(H)$  eines Hilbertraums  $H$  heißt *normal*, wenn  $T^*T = TT^*$  ist.

**Satz 3.1.9** (Eigenschaften von selbstadjungierten Operatoren). *Die folgenden Aussagen geben uns eine breite Basis um mit selbstadjungierten Operatoren effektiv arbeiten zu können.*

- (i) Ein beschränkter linearer Operator  $T \in \mathcal{L}(H)$  auf einem Hilbertraum  $H$  ist genau dann normal, wenn  $\|Tx\| = \|T^*x\|$  für alle  $x \in H$  gilt.
- (ii) Wenn  $T_1, T_2$  selbstadjungierte Operatoren sind, dann ist auch jede Linearkombination  $\alpha_1 T_1 + \alpha_2 T_2$  mit  $\alpha_1, \alpha_2 \in \mathbb{R}$  selbstadjungiert.
- (iii) Wenn  $T_1, T_2$  selbstadjungierte Operatoren sind, dann ist  $T_1 T_2$  ist genau dann selbstadjungiert, wenn  $T_1 T_2 = T_2 T_1$  ist.
- (iv) Sei  $T \in \mathcal{L}(H)$ . Dann ist  $T$  genau dann selbstadjungiert, wenn  $\langle Tx, x \rangle$  für alle  $x \in H$  reell ist.

*Beweis.* Wir widmen uns lediglich dem Beweis von Aussage (iv). Die anderen Beweise sind Übungsaufgaben. Sei  $T$  ein selbstadjungierter Operator. Dann gilt

$$\langle Tx, x \rangle = \langle x, T^*x \rangle = \langle x, Tx \rangle = \overline{\langle Tx, x \rangle},$$

und daraus folgt, dass  $\langle Tx, x \rangle$  reellwertig ist. Für die Rückrichtung machen wir uns die Polarisierungsidentität zu nutze. Für  $B \in \mathcal{L}(H)$  gilt:

$$\langle Bx, y \rangle = \frac{1}{4} \langle B(x+y), x+y \rangle - \langle B(x-y), x-y \rangle + i \langle B(x+iy), x+iy \rangle - i \langle B(x-iy), x-iy \rangle.$$

Mit dieser Identität folgt aus  $\langle Tx, x \rangle = \langle x, Tx \rangle = \langle T^*x, x \rangle$  für alle  $x \in H$  auch  $\langle Tx, y \rangle = \langle T^*x, y \rangle$  für alle  $x \in H$ , denn wir haben die Gleichungskette

$$\begin{aligned} \langle Tx, y \rangle &= \frac{1}{4} \langle T(x+y), x+y \rangle - \langle T(x-y), x-y \rangle + i \langle T(x+iy), x+iy \rangle - i \langle T(x-iy), x-iy \rangle \\ &= \frac{1}{4} \langle T^*(x+y), x+y \rangle - \langle T^*(x-y), x-y \rangle + i \langle T^*(x+iy), x+iy \rangle - i \langle T^*(x-iy), x-iy \rangle \\ &= \langle T^*x, y \rangle. \end{aligned}$$

Da dies nun für beliebige  $x, y \in H$  galt, muss  $T = T^*$  sein. □

Für selbstadjungierte Operatoren kann man eine Halbordnung definieren, die diese Operatoren also vergleichbar macht. Dies muss über Skalarprodukte geschehen, damit man die Ordnungsstruktur des dem Hilbertraum zugrundeliegenden Körpers ausnutzen kann. Wir werden später, im Abschnitt 3.2 über Projektionsoperatoren, noch einmal auf diese Halbordnung zurückkommen und damit einige Projektionsoperatoren sogar charakterisieren können. Für vergleichbare selbstadjungierte Operatoren hat man aufgrund der Vektorraumstruktur von  $\mathcal{L}(H)$  sogar einige Rechenregeln, die denen auf geordneten Körpern verdächtig ähneln.

**Definition 3.1.10.** Seien  $A, B \in \mathcal{L}(H)$  selbstadjungierte Operatoren. Wir schreiben  $A \geq B$ , wenn  $\langle Ax, x \rangle \geq \langle Bx, x \rangle$  für alle  $x \in H$  gilt. Ein selbstadjungierter Operator  $A \in \mathcal{L}(H)$  heißt *positiv*, wenn  $A \geq 0$  ist, wenn also  $\langle Ax, x \rangle \geq 0$  für alle  $x \in H$  ist. Die Ungleichung ist wohldefiniert, da nach Satz 3.1.9 (iv)  $\langle Ax, x \rangle$  reellwertig ist.

Die obige Definition ist etwas unglücklich gewählt, da so auch der Nulloperator positiv genannt wird, allerdings hat es sich so als Standard etabliert.

**Satz 3.1.11** (Rechenregeln für selbstadjungierte Operatoren).

- (i) Seien  $A_1, A_2$  zwei selbstadjungierte Operatoren mit  $A_1 \geq A_2$ . Dann gilt auch  $A_1 + B \geq A_2 + B$  für jeden selbstadjungierten Operator  $B \in \mathcal{L}(H)$ .
- (ii) Wenn für selbstadjungierte Operatoren  $A_1, A_2, B_1, B_2 \in \mathcal{L}(H)$  die Ungleichungen  $A_1 \geq A_2$  und  $B_1 \geq B_2$  gelten, dann gilt auch  $A_1 + B_1 \geq A_2 + B_2$ .
- (iii) Wenn  $A_1 \geq A_2$  gilt, so folgt für alle  $\lambda \geq 0 \in \mathbb{C}$ , dass  $\lambda A_1 \geq \lambda A_2$  ist.
- (iv) Aus der Polarisierungsidentität folgt: Wenn  $A \geq B$  und  $B \geq A$  gilt, so folgt  $A = B$ .

### 3.1.3 Isometrische Operatoren und unitäre Operatoren

**Definition 3.1.12.** Ein Operator  $T \in \mathcal{L}(H)$  eines Hilbertraums  $H$  heißt *isometrisch*, wenn  $\|Tx\| = \|x\|$  für alle  $x \in H$  gilt.

Ein weiterer wichtiger Begriff ist der des unitären Operators. Wir werden in Kürze feststellen, dass hier enge Parallelen beispielsweise zum Begriff der unitären Matrix vorliegen. Zunächst werden wir aber einige Zusammenhänge zwischen den verschiedenen definierten Begriffen herstellen.

**Definition 3.1.13.** Ein Operator  $T \in \mathcal{L}(H)$  eines Hilbertraums  $H$  heißt *unitär*, wenn  $T^*T = TT^* = I$  ist, wobei  $I$  die identische Abbildung ist.

Ein isometrischer Operator erhält nicht nur die Norm, sondern sogar das Skalarprodukt mit verschiedenen Werten in den Eingängen.

**Satz 3.1.14.** Ein Operator  $T \in \mathcal{L}(H)$  ist genau dann isometrisch, wenn  $\langle Tx, Ty \rangle = \langle x, y \rangle$  für alle  $x, y \in H$  gilt.

**Satz 3.1.15.** Sei  $T \in \mathcal{L}(H)$ . Dann ist  $T$  genau dann unitär, wenn  $T$  isometrisch und surjektiv ist.

*Beweis.* Sei  $T$  unitär, das heißt  $T^*T = TT^* = I$ . Dann gilt  $T(T^*(x)) = Ix = x$  für alle  $x \in H$  und somit bildet  $T$  den Raum  $H$  surjektiv auf  $H$  ab. Die Isometrie erhält man aus folgender Betrachtung:

$$\|T(x)\|^2 = \langle Tx, Tx \rangle = \langle T^*Tx, x \rangle = \langle Ix, x \rangle = \|x\|^2.$$

Sei für die Rückrichtung  $T$  isometrisch und surjektiv. Die Injektivität von  $T$  folgt leicht, Wenn nämlich  $Tx = 0$  ist folgt, dann folgt auch  $\|Tx\| = \|x\| = 0$  und das bedeutet, dass  $x = 0$  sein muss. Also ist  $T$  bijektiv und deshalb existiert die Umkehrabbildung  $T^{-1}$ . Nach Satz 3.1.14 gilt  $\langle Tx, Ty \rangle = \langle x, y \rangle$  für alle  $x, y \in H$ . Wir wählen  $y = T^{-1}z$  für ein  $z \in H$ . Daraus ergibt sich nun die Gleichung  $\langle Tx, z \rangle = \langle x, T^{-1}z \rangle$  für alle  $x, z \in H$  und deshalb muss  $T^* = T^{-1}$  sein.  $\square$

Wenn  $U$  und  $V$  unitäre Operatoren sind, so sind auch  $UV$  und  $U^{-1}$  unitär, was man durch einfaches Nachrechnen sofort sieht. Damit bilden die unitären Operatoren aus  $\mathcal{L}(H)$  mit der Hintereinanderschaltung als Verknüpfung eine Gruppe. Wir betrachten nun einige wichtige Operatoren und stellen ihre Eigenschaften fest.

**Beispiel 3.1.16.**

- (a) Wir können Matrizen als Operatoren auffassen. Sei  $H = \mathbb{C}^n$  mit dem üblichen Skalarprodukt  $\langle x, y \rangle = \sum_{j=1}^n x_j \overline{y_j}$ , wobei jeweils  $x = (x_1, \dots, x_n)$  und  $y = (y_1, \dots, y_n)$  sei. Sei weiter  $A = (a_{ij}) \in \mathbb{C}^{n \times n}$  eine Matrix. Wir definieren dann einen Operator  $T_A$  durch  $T_A(x) = Ax$ . Sei weiter  $A^* = (b_{ij})$  mit  $b_{ij} = \overline{a_{ji}}$  für  $i, j = 1, \dots, n$  die adjungierte Matrix. Es gilt für  $x, y \in H$ :

$$\langle T_A x, z \rangle = \sum_{i=1}^n \left( \sum_{j=1}^n a_{ij} x_j \right) \overline{z_i} = \sum_{i,j=1}^n a_{ij} x_j \overline{z_i}.$$

Also gilt haben wir für den  $i$ -ten Eintrag des Vektors  $T_A x$  die Gleichung  $(T_A x)_i = \sum_{j=1}^n a_{ij} x_j$ . Für die adjungierte Matrix  $A^*$  und den zugehörigen Operator  $T_{A^*}$  haben wir die Gleichung

$$\langle x, T_{A^*} z \rangle = \sum_{i=1}^n x_i \left( \sum_{j=1}^n \overline{b_{ij} z_j} \right) = \sum_{i,j=1}^n \overline{b_{ij} z_j} x_i = \sum_{i,j=1}^n a_{ji} x_i \overline{z_j} = \sum_{i,j=1}^n a_{ij} x_j \overline{z_i}.$$

Damit gilt  $\langle T_A(x), z \rangle = \langle x, T_{A^*} z \rangle$ , sprich  $(T_A)^* = T_{A^*}$ . Hiermit haben wir den Zusammenhang zwischen adjungierten Matrizen und dem adjungierten Matrixoperator gezeigt.

Es gilt aufgrund der Rechenregeln für Matrizen, dass  $T_A T_B = T_{AB}$  ist, also ist die Abbildung von  $\mathbb{C}^{n \times n}$  nach  $\mathcal{L}(H)$ , die durch  $A \mapsto T_A$  definiert ist, ein Homomorphismus von  $*$ -Algebren. Desweiteren ist  $T_A$  selbstadjungiert genau dann, wenn  $A = A^*$  ist, wenn also die Matrix  $A$  hermitesch ist.  $T_A$  ist unitär genau dann, wenn  $AA^* = A^*A = E$  gilt, wenn also  $A$  unitär ist.

- (b) Betrachten wir nun ein Beispiel eines Integraloperators. Sei  $I$  ein reelles Intervall und  $H = L^2(I, \lambda)$ . Beispielsweise könnten wir  $I = [0, 1]$  und das Lebesgue-Maß  $\lambda$  auf  $[0, 1]$  wählen. Sei weiter  $K(s, t)$  eine Lebesgue-messbare Funktion auf  $I \times I$ , die quadratisch integrierbar ist, also  $K(\cdot, \cdot) \in L^2(I \times I, \lambda)$ . Wir definieren einen Operator  $T_K$  durch

$$(T_K f)(s) = \int_I K(s, t) f(t) \lambda(dt)$$

für alle  $f \in H$ . Dann ist  $T_K \in \mathcal{L}(H)$  und es gilt für die Operatornorm

$$\|T_K\| \leq \left( \int_I \int_I |K(s, t)|^2 \lambda(ds) \lambda(dt) \right)^{1/2}.$$

Sei  $K^*(s, t) := \overline{K(t, s)}$  für  $s, t \in I$ . Man zeigt analog zum ersten Beispiel, dass  $(T_K)^* = T_{K^*}$  gilt. Damit ist  $T_K$  selbstadjungiert genau dann, wenn  $K(s, t) = \overline{K(t, s)}$   $\lambda$ -fast überall ist, wenn also  $K(s, t)$   $\lambda$ -fast überall reell ist.

- (c) Wir betrachten nun Multiplikationsoperatoren. Sei wieder  $I$  ein reelles Intervall. Es sei weiter  $a \in L^\infty(I)$  und  $H = L^2(I, \lambda)$ . Wir definieren  $(T_a(f))(t) = a(t)f(t)$  für alle  $f \in H$ . Dann ist  $T_a \in \mathcal{L}(H)$ . Man zeigt weiter, dass  $\|T_a\| = \|a\|_{L^\infty(I)}$  gilt. Wenn  $a$  stetig und beschränkt auf  $I$  ist, dann ist  $\|T_a\| = \sup_{t \in I} |a(t)|$ .
- (d) Als weiteres Beispiel wollen wir die Fouriertransformation betrachten. Sei  $H = L^2(\mathbb{R}, \lambda)$ . Für  $f \in L^1(\mathbb{R}^d, \lambda)$  sei

$$(\mathcal{F}f)(t) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{-i(t,s)} f(s) \lambda(ds),$$

wobei  $(t, s) = \sum_{k=1}^d t_k s_k$  das Skalarprodukt auf dem reellen Hilbertraum  $\mathbb{R}^d$  sei. Wir definieren den Schwartzraum  $\mathcal{S}(\mathbb{R}^d)$  durch

$$\mathcal{S}(\mathbb{R}^d) = \left\{ f \in C^\infty(\mathbb{R}^d) \mid \sup_{x \in \mathbb{R}^d} |D^\alpha(x^\beta f(x))| < \infty \text{ für alle } \alpha, \beta \in \mathbb{N}_0^d \right\}.$$

Für die Multiindexnotation siehe Abschnitt 1.2.2. Der Operator  $\mathcal{F}$  hat die folgenden Eigenschaften.

- 1)  $\mathcal{F}$  bildet  $\mathcal{S}(\mathbb{R}^d)$  bijektiv  $\mathcal{S}(\mathbb{R}^d)$  ab und es gilt  $(\mathcal{F}^{-1}g)(s) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{i(s,t)} f(t) \lambda(dt)$ . Damit ist  $\mathcal{F}$  ein Automorphismus auf  $\mathcal{S}(\mathbb{R}^d)$ .
- 2) Es gilt das *Plancherel-Theorem*:  $\mathcal{F}$  lässt Normen invariant, das heißt  $\|\mathcal{F}f\|_{L^2(\mathbb{R}^d)} = \|f\|_{L^2(\mathbb{R}^d)}$ .

$\mathcal{S}(\mathbb{R}^d)$  ist ein dichter linearer Teilraum des  $L^2(\mathbb{R}^d)$ .  $\mathcal{F}$  besitzt aufgrund der eben festgestellten Eigenschaften eine stetige Fortsetzung zu einem isometrischen Operator auf  $L^2(\mathbb{R}^d)$ , den wir ebenfalls mit  $\mathcal{F}$  bezeichnen. Dieser bildet  $L^2(\mathbb{R}^d)$  bijektiv auf  $L^2(\mathbb{R}^d)$  ab und ist damit ein unitärer Operator im Hilbertraum  $L^2(\mathbb{R}^d)$ . Wir meinen in Zukunft mit  $\mathcal{F}$  stets diesen fortgesetzten Operator. Man kann zeigen, dass gilt:

$$(\mathcal{F}f)(t) = \frac{1}{(2\pi)^{d/2}} \lim_{R \rightarrow \infty} \int_{\{x \mid \|x\| \leq R\}} e^{-i(t,s)} f(s) \lambda(ds).$$

- (e) Als letztes Beispiel wollen wir die Verschiebungsoperatoren (Shiftoperatoren) betrachten. Sei  $H = \ell^2$  und  $S(x_1, x_2, x_3, \dots) = (0, x_1, x_2, x_3, \dots)$ . Es gilt

$$\|S(x_n)\|^2 = 0^2 + \sum_{n=1}^{\infty} |x_n|^2 = \|(x_n)\|^2,$$

also ist  $S$  isometrisch. Der Vektor  $(1, 0, 0, \dots)$  liegt jedoch nicht im Bild von  $S$ , also bildet  $S$  den Raum  $H$  nicht surjektiv auf sich selbst ab und demnach ist  $S$  nicht unitär. Man nennt  $S$  auch *unilateral shift*. Setzen wir hingegen  $H = \ell^2(\mathbb{Z})$  und  $S(\dots, x_{-1}, x_0, x_1, x_2, \dots) = (\dots, x_{-2}, x_{-1}, x_0, x_1, \dots)$ , so ist  $S$  wiederum isometrisch und diesmal sogar unitär. Dieser Operator  $S$  heißt *bilateral shift*.

Eine wichtige Konstruktion ist der Hardyraum  $H^2(\mathbb{D})$  auf der offenen Kreisscheibe  $\mathbb{D} = \{z \in \mathbb{C} \mid |z| < 1\}$ . Der Raum  $H^2(\mathbb{D})$  ist die Menge aller holomorphen Funktionen auf  $\mathbb{D}$ , für die

$$\|f\|_2 := \sup_{0 < r < 1} \left( \frac{1}{2\pi} \int_0^{2\pi} |f(re^{i\varphi})|^2 d\varphi \right)^{1/2} < \infty$$

ist. Sei  $f(z) = \sum_{n=0}^{\infty} a_n z^n$  eine Potenzreihe. Es gilt dann

$$|f(re^{i\varphi})|^2 = \left| \sum_{n=0}^{\infty} a_n r^n e^{in\varphi} \right|^2 = \sum_{n,k=0}^{\infty} a_n \bar{a}_k r^{n+k} e^{i(n-k)\varphi}$$

Hiermit folgt:

$$\begin{aligned} \int_0^{2\pi} |f(re^{i\varphi})|^2 d\varphi &= \int_0^{2\pi} \sum_{n,k=0}^{\infty} a_n \bar{a}_k r^{n+k} e^{i(n-k)\varphi} d\varphi = \sum_{n,k=0}^{\infty} a_n \bar{a}_k r^{k+n} \underbrace{\int_0^{2\pi} e^{i(n-k)\varphi} d\varphi}_{=2\pi \delta_{kn}} \\ &= 2\pi \sum_{n=0}^{\infty} |a_n|^2 r^{2n} \end{aligned}$$

und damit ist

$$\sup_{0 < r < 1} \frac{1}{2\pi} \int_0^{2\pi} |f(re^{i\varphi})|^2 d\varphi = \sum_{n=0}^{\infty} |a_n|^2.$$



Für eine Funktion  $f \in H^2(D)$ , die durch die Potenzreihe  $f(z) = \sum_{n=0}^{\infty} a_n z^n$  dargestellt ist, muss also die Folge  $(a_n)_{n \in \mathbb{N}} \in \ell^2(\mathbb{N}_0)$  sein. Ist umgekehrt  $(b_n) \in \ell^2(\mathbb{N}_0)$  eine solche Folge, so definieren wir eine Funktion  $g$  durch die Potenzreihe  $g(z) = \sum_{n=0}^{\infty} b_n z^n$  für alle  $z \in \mathbb{D}$ . Dann gilt mit  $r = |z| < 1$  nach der Cauchy-Schwarz-Ungleichung

$$\sum_{n=0}^{\infty} |b_n z^n| = \sum_{n=0}^{\infty} |b_n| r^n \leq \left( \sum_{n=0}^{\infty} |b_n|^2 \right)^{1/2} \left( \sum_{n=0}^{\infty} r^{2n} \right)^{1/2} < \infty.$$

Da die Potenzreihe somit überall auf  $\mathbb{D}$  konvergiert, definiert  $g(z) = \sum_{n=0}^{\infty} b_n z^n$  eine holomorphe Funktion auf  $\mathbb{D}$ , für die

$$\sup_{0 < r < 1} \frac{1}{2\pi} \int_0^{2\pi} |g(re^{i\varphi})|^2 d\varphi = \sum_{n=0}^{\infty} |b_n|^2$$

ist. Also gilt  $g \in H(\mathbb{D})$  und der Hardyraum hat damit folgende Gestalt:

$$H^2(D) = \left\{ f \mid f \text{ ist holomorph auf } \mathbb{D}, f(z) = \sum_{n=0}^{\infty} a_n z^n, (a_n)_{n \in \mathbb{N}} \in \ell^2(\mathbb{N}_0) \right\}.$$

Für zwei Potenzreihen  $f(z) = \sum_{n=0}^{\infty} a_n z^n$  und  $g(z) = \sum_{n=0}^{\infty} b_n z^n$  ist

$$\langle f, g \rangle = \lim_{R \uparrow 1} \frac{1}{2\pi} \int_0^{2\pi} f(re^{i\varphi}) \overline{g(re^{i\varphi})} d\varphi = \sum_{n=0}^{\infty} a_n \overline{b_n}.$$

Damit ist  $\langle \cdot, \cdot \rangle$  ein Skalarprodukt auf  $H^2(\mathbb{D})$  und damit der Hardyraum ist ein Hilbertraum, der nach Korollar 1.5.13 natürlich isomorph zum  $\ell^2(\mathbb{N}_0)$  ist. Dennoch ist der Hardyraum als eigenständiger Raum wichtig, denn er trägt zusätzliche Struktur. Die Funktionen im  $H^2(\mathbb{D})$  sind holomorph und  $\{z^n \mid n \in \mathbb{N}_0\}$  bildet ein vollständiges Orthonormalsystem. Der Shiftoperator wirkt auf eine Funktion  $z^n$  durch  $S_0 z^n := z^{n+1}$ , also durch Multiplikation mit der Variablen  $z$ . Dies ist ein Beispiel für einen häufig auftretenden und sehr angenehmen Fall, nämlich dass man mathematische Strukturen unter verschiedenen Gesichtspunkten betrachten kann und Resultate unterschiedlicher mathematischer Teildisziplinen anwenden kann, deren Implikationen sich dann beispielsweise durch Isomorphismen auf andere „Betrachtungsweisen“ übertragen.

## 3.2 Projektionsoperatoren

Wir bringen zunächst eine kurze Wiederholung. Wenn  $H_1$  ein abgeschlossener, linearer Teilraum des Hilbertraumes  $H$  ist, dann ist nach dem 1. Satz von Riesz (Satz 1.4.6)  $H = H_1 \oplus H_1^\perp$ , das heißt für alle  $x \in H$  existieren eindeutig bestimmte  $x_1 \in H_1$  und  $x_2 \in H_1^\perp$  mit  $x = x_1 + x_2$ . Wir definieren nun den *Projektionsoperator* auf den Vektorraum  $H_1$  durch  $P_{H_1} x := x_1$ . Aufgrund der Gleichung

$$\|P_{H_1}\|^2 = \|x_1\|^2 \leq \|x_1\|^2 + \|x_2\|^2 = \|x_1 + x_2\|^2 = \|x\|^2,$$

in der wir den Satz des Pythagoras benutzt haben, ist  $P_{H_1} \in \mathcal{L}(H)$  und  $\|P_{H_1}\| \leq 1$ .

**Satz 3.2.1.** *Ein Operator  $T \in \mathcal{L}(\mathcal{H})$  ist ein Projektionsoperator genau dann, wenn  $T = T^*$  gilt, also  $T$  selbstadjungiert ist, und wenn  $T^2 = T$  ist. Die Menge  $\{T(x) \mid x \in H\}$  ist dann der Vektorraum, auf den  $T$  projiziert.*

*Beweis.* Wir beweisen zunächst die Hinrichtung. Sei  $T = P_{H_1}$  ein Projektionsoperator auf einen Unterraum  $H_1$  des Hilbertraums  $H$  und sei  $x = x_1 + x_2 \in H$  die Zerlegung eines Elementes  $x \in H$  in  $x_1 \in H_1$  und  $x_2 \in H_1^\perp$ . Dann gilt

$$T^2(x) = T(T(x)) = T(x_1) = x_1 = T(x),$$

also  $T^2 = T$ . Weiter haben wir

$$\begin{aligned}\langle T(x), y \rangle &= \langle P_{H_1}(x), y \rangle = \langle x_1, y_1 + y_2 \rangle = \langle x_1, y_1 \rangle + \langle x_1, y_2 \rangle = \langle x_1, y_1 \rangle \\ &= \langle x_1, y_1 \rangle + \langle x_2, y_1 \rangle = \langle x_1 + x_2, y_1 \rangle = \langle x, P_{H_1}(y) \rangle = \langle x, T(y) \rangle,\end{aligned}$$

wobei wir benutzt haben, dass  $\langle x_1, y_2 \rangle = 0$  gilt, da  $x_1 \in H$  und  $y_2 \in H_1^\perp$  ist. Damit ist also  $T = T^*$ .

Nun beweisen wir die Rückrichtung. Sei  $T\mathcal{L}(H)$  mit  $T^* = T$  und  $T^2 = T$  und sei  $H_1 = \{x \in H \mid Tx = x\}$  das Bild von  $H$  unter  $T$ . Es ist  $H_1$  ein abgeschlossener linearer Teilraum von  $H$ , da  $T$  linear und stetig ist. Für  $x \in H$  gilt  $T(T(x)) = T^2x = Tx$ , das heißt, dass  $Tx \in H_1$  ist. Weiter gilt  $H = H_1 \oplus H_1^\perp$  nach dem ersten Satz von Riesz (Satz 1.4.6). Sei  $x \in H$  und  $z \in H_1$ . Wir zeigen, dass  $x - Tx \in H_1^\perp$  ist. Es gilt nämlich

$$\langle x - Tx, z \rangle = \langle x, z \rangle - \langle Tx, z \rangle = \langle x, z \rangle - \langle x, Tz \rangle = \langle x, z \rangle - \langle x, z \rangle = 0$$

aufgrund der Selbstadjungiertheit von  $T$ , also ist tatsächlich  $x - Tx \in H_1^\perp$ . Da  $x = Tx + x - Tx$  eine Zerlegung jedes Elementes  $x \in H$  ist, bildet  $T$  den Raum  $H$  auf  $H_1$  ab und ist der Projektionsoperator  $P_{H_1}$ .  $\square$

Es gibt daher genau so viele Projektionsoperatoren in einem Raum, wie es Unterräume in ihm gibt. Die trivialen Projektionsoperatoren sind dabei die Identität auf dem gesamten Raum und der Nulloperator. Für Projektionsoperatoren gibt es viele Charakterisierungsmöglichkeiten. Die folgenden Sätze stellen einige Resultate über Summen und Produkte von Projektionsoperatoren zusammen.

**Satz 3.2.2.** *Seien  $P_1$  und  $P_2$  Projektionsoperatoren auf Unterräume  $H_1$  und  $H_2$  von  $H$ . Folgende Aussagen sind äquivalent:*

- (i)  $P_1 + P_2$  ist ein Projektionsoperator,
- (ii)  $P_1P_2 = 0$ ,
- (iii)  $H_1 \perp H_2$ .

Ist eine dieser Bedingungen erfüllt, dann ist  $P_1 + P_2$  der Projektionsoperator von  $H$  auf den Raum  $H_1 \oplus H_2$ .

*Beweis.* Wir zeigen zunächst die Implikation (ii)  $\Rightarrow$  (i). Sei dazu  $P = P_1 + P_2$ . Dann gilt  $P^* = (P_1 + P_2)^* = P_1^* + P_2^* = P_1 + P_2 = P$ . Wegen  $P_1P_2 = 0$  gilt auch  $0 = (P_1P_2)^* = P_2^*P_1^* = P_2P_1$ . Damit folgt

$$P^2 = (P_1 + P_2)(P_1 + P_2) = P_1^2 + P_1P_2 + P_2P_1 + P_2^2 = P_1 + P_2 = P.$$

Damit ist  $P$  ein Projektionsoperator nach Satz 3.2.1. Nun zeigen wir die umgekehrte Richtung, (i)  $\Rightarrow$  (ii). Sei  $P = P_1 + P_2$  ein Projektionsoperator. Dann gilt

$$P_1 + P_2 = P = P^2 = P_1^2 + P_1P_2 + P_2P_1 + P_2^2 = P_1 + P_2 + P_1P_2 + P_2P_1$$

und damit  $P_1P_2 + P_2P_1 = 0$ . Multiplikation dieser Gleichung von links und von rechts mit  $P_1$  liefert

$$P_1P_2P_1 + P_2P_1^2 = P_1P_2P_1 + P_2P_1 = 0,$$

sowie

$$P_1^2P_2 + P_1P_2P_1 = P_1P_2 + P_1P_2P_1 = 0.$$

Zusammen liefern diese Gleichungen, dass  $P_1P_2 = P_2P_1$  sein muss. Damit und aus obigen Gleichungen haben wir

$$P_1P_2 + P_2P_1 = 2P_1P_2 = 0,$$

also  $P_1P_2 = 0$ . Es bleiben noch zwei Äquivalenzen zu zeigen. Für (ii)  $\Rightarrow$  (iii) sei  $x_1 \in H_1$  und  $x_2 \in H_2$ . Es gilt

$$\langle x_1, x_2 \rangle = \langle P_1x_1, P_2x_2 \rangle = \langle x_1, P_1P_2x_2 \rangle = \langle x_1, 0 \rangle = 0.$$

Die Richtung (iii)  $\Rightarrow$  (ii) folgt analog.  $\square$

**Satz 3.2.3.** Seien  $P_1, P_2$  wie in Satz 3.2.2 definiert.  $P_1P_2$  ist ein Projektionsoperator genau dann, wenn  $P_1P_2 = P_2P_1$  gilt. Ist dies erfüllt, dann ist der projizierte Raum gleich  $H_1 \cap H_2$ .

**Satz 3.2.4.** Seien  $P_1, P_2$  wie in Satz 3.2.2 definiert. Folgende Aussagen sind äquivalent.

- (i)  $P_1 - P_2$  ist Projektionsoperator,
- (ii)  $P_1 \geq P_2$ ,
- (iii)  $H_1 \supseteq H_2$ ,
- (iv)  $P_1P_2 = P_2$ ,
- (v)  $P_2P_1 = P_2$ .

Ist eine dieser Bedingungen erfüllt, dann ist der zu  $P_1 - P_2$  gehörende projizierte Unterraum von  $H$  gleich  $H_1 \ominus H_2$ , dem orthogonalen Komplement von  $H_2$  im Hilbertraum  $H_1$ .

*Beweis.* Siehe Übungen, Seminar. □

Insbesondere besagt dieser Satz, dass sowohl das links- als auch das rechtsseitige Produkt eines kleineren Projektionsoperators mit einem größeren immer den kleineren liefert (Dies besagen die Implikationen (ii)  $\Rightarrow$  (iv) beziehungsweise (ii)  $\Rightarrow$  (v)).

### 3.3 Konvergenzarten im Hilbertraum

Zuerst wollen wir uns mit der Konvergenz von Elementen im Hilbertraum beschäftigen. Sei also  $(H, \langle \cdot, \cdot \rangle)$  ein Hilbertraum und sei  $(x_n)_{n \in \mathbb{N}}$  eine Folge von Elementen in  $H$ . Den gewöhnlichen Konvergenzbegriff haben wir bereits kennengelernt. Es gilt  $x_n \rightarrow x$  genau dann, wenn  $\lim_{n \rightarrow \infty} \|x_n - x\| = 0$  ist. Nun werden wir eine zweite Konvergenzart definieren.

**Definition 3.3.1.** Wir sagen, dass eine Folge  $(x_n)_{n \in \mathbb{N}}$  *schwach* gegen  $x$  konvergiert und schreiben  $x_n \rightharpoonup x$ , wenn  $\lim_{n \rightarrow \infty} \langle x_n, y \rangle = \langle x, y \rangle$  für alle  $y \in H$  gilt. Eine Folge  $(x_n)_{n \in \mathbb{N}}$  heißt eine *schwache Cauchyfolge*, wenn für jedes  $y \in H$  die Folge  $(\langle x_n, y \rangle)_{n \in \mathbb{N}}$  eine Cauchyfolge ist.

#### Bemerkung 3.3.2.

- (i) Aus  $x_n \rightarrow x$  folgt  $x_n \rightharpoonup x$ . Normkonvergenz impliziert im Hilbertraum also die schwache Konvergenz.
- (ii) Aus  $x_n \rightharpoonup x$  und  $x_n \rightharpoonup x'$  folgt  $x = x'$ , denn ist  $\langle x_n, y \rangle \rightarrow \langle x, y \rangle = \langle x', y \rangle$  für alle  $y \in H$ , so folgt  $\langle x - x', y \rangle = 0$ . Setzt man dann  $y = x - x'$ , so folgt  $\|y\|^2 = 0$  und damit  $x = x'$ .
- (iii) Es gilt  $x_n \rightarrow x$  genau dann, wenn  $\|x_n\| \rightarrow \|x\|$  und gleichzeitig  $x_n \rightharpoonup x$  gilt.

**Beispiel 3.3.3.** Sei  $(e_n)_{n \in \mathbb{N}}$  ein normiertes Orthogonalsystem im Hilbertraum  $H$ . Dann gilt  $e_n \rightharpoonup 0$ . Sei nämlich ein beliebiges  $y \in H$  gewählt. Nach der Besselschen Ungleichung (Satz 1.5.5) gilt

$$\sum_{n=1}^{\infty} |\langle y, e_n \rangle|^2 \leq \|y\|^2.$$

Die Reihe konvergiert, also gilt  $\langle y, e_n \rangle \rightarrow 0$  und damit  $\langle e_n, y \rangle \rightarrow 0 = \langle 0, y \rangle$ . Demzufolge gilt  $e_n \rightharpoonup 0$ , da  $y$  beliebig war. Hingegen ist  $\|e_n\| = 1$  für alle  $n \in \mathbb{N}$ , also  $e_n \not\rightarrow 0$ .

**Satz 3.3.4.**

- (i) Jede schwach konvergente Folge  $(x_n)_{n \in \mathbb{N}}$  ist normbeschränkt, das heißt  $\sup_{n \in \mathbb{N}} \|x_n\| < \infty$ .
- (ii) Jede beschränkte Folge  $(x_n)_{n \in \mathbb{N}}$  in  $H$  besitzt eine schwach konvergente Teilfolge.
- (iii) Eine schwache Cauchyfolge konvergiert stets schwach gegen ein Element des Hilbertraumes.

*Beweis.* Wir beschränken uns auf die Aussage (i). Der Beweis wird mit dem Satz von Banach-Steinhaus (Satz 2.3.1) geführt. Sei  $(x_n)_{n \in \mathbb{N}}$  eine Folge in  $H$  mit  $x_n \rightharpoonup x$ . Es gilt dann also  $\langle x_n, y \rangle \rightarrow \langle x, y \rangle$  für jedes  $y \in H$ . Die Folge  $\mathcal{F}_n(y) = \langle y, x_n \rangle$  von stetigen linearen Funktionalen auf  $H$  ist demnach punktweise beschränkt. Damit folgt aus dem Satz von Banach-Steinhaus, dass die Folge  $(\mathcal{F}_n)_{n \in \mathbb{N}}$  normbeschränkt ist.  $\square$

Als nächstes wenden wir uns der Untersuchung von Folgen von Operatoren zu. Sei  $T_n \in \mathcal{L}(H)$  für jedes  $n \in \mathbb{N}$  und sei außerdem  $T \in \mathcal{L}(H)$ . Wir wissen bereits, was es heißt, wenn  $T_n$  in der Operatornorm gegen  $T$  konvergiert, also wenn  $\lim_{n \rightarrow \infty} \|T_n - T\| = 0$  ist. Man sagt dazu auch, dass  $(T_n)$  gleichmäßig gegen  $T$  konvergiert, weil die gleichmäßige Konvergenz auf der Einheitskugel vorliegt. Wir werden dafür  $T_n \Rightarrow T$  schreiben. Ein weiterer Begriff ist der der starken Konvergenz, die jedoch in Wirklichkeit eine vergleichsweise schwache Aussage über Konvergenz liefert. Es gibt jedoch auch noch den Begriff der schwachen Konvergenz für Operatoren, ähnlich wie bei Folgen von Elementen des Hilbertraumes.

**Definition 3.3.5.** Wir sagen, eine Folge  $(T_n)_{n \in \mathbb{N}}$  konvergiere

- (i) in der starken Operatortopologie oder konvergiere *stark* gegen  $T$ , wenn  $\lim_{n \rightarrow \infty} T_n x = T x$  für alle  $x \in H$  ist. Wir schreiben dafür  $T_n \xrightarrow{s} T$ .
- (ii) in der schwachen Operatortopologie oder konvergiere *schwach* gegen  $T$ , wenn  $T_n x \rightharpoonup T x$  für alle  $x \in H$  gilt, das heißt wenn  $\lim_{n \rightarrow \infty} \langle T_n x, y \rangle = \langle T x, y \rangle$  für alle  $x, y \in H$  ist. Wir schreiben dafür  $T_n \xrightarrow{w} T$ , wobei das  $w$  für „weak“ steht.

**Bemerkung 3.3.6.** Für die Konvergenzen gelten folgende einfache Eigenschaften:

- (i) Aus  $T_n \Rightarrow T$  folgt  $T_n \xrightarrow{s} T$  und aus  $T_n \xrightarrow{s} T$  folgt  $T_n \xrightarrow{w} T$ , das heißt Konvergenz impliziert starke Konvergenz und diese wiederum die schwache Konvergenz.
- (ii) Aus  $T_n \xrightarrow{w} T$  und  $T_n \xrightarrow{w} T'$  folgt  $T = T'$ , das heißt der schwache Limes einer Folge von Operatoren ist eindeutig.

Einige technisch wichtige Aussagen über schwache Konvergenz von Operatoren stellt der folgende Satz zusammen.

**Satz 3.3.7.**

- (i) Jede schwach konvergente Folge  $(T_n)_{n \in \mathbb{N}}$  ist in der Operatornorm beschränkt, das heißt

$$\sup_{n \in \mathbb{N}} \|T_n\| < \infty.$$

- (ii)  $H$  sei separabel. Sei  $T_n \in \mathcal{L}(H)$  für  $n \in \mathbb{N}$  und sei  $\sup_{n \in \mathbb{N}} \|T_n\| < \infty$ . Dann existiert eine Teilfolge  $(T_{n_k})_{k \in \mathbb{N}}$  von  $(T_n)_{n \in \mathbb{N}}$ , die in der schwachen Operatortopologie gegen einen Operator  $T \in \mathcal{L}(H)$  konvergiert.
- (iii) Sei  $T_n \in \mathcal{L}(H)$  für jedes  $n \in \mathbb{N}$ . Für alle  $x, y \in H$  sei die Zahlenfolge  $(\langle T_n x, y \rangle)_{n \in \mathbb{N}}$  eine Cauchyfolge. Dann existiert ein  $T \in \mathcal{L}(H)$  mit  $T_n \xrightarrow{w} T$ .

*Beweis.* Den Beweis findet man in jedem guten Buch über Hilberträume.  $\square$

Wir wollen uns nun kurz mit monotonen, beschränkten Folgen selbstadjungierter Operatoren beschäftigen. Der in Kürze folgende Satz ist eine Art Verallgemeinerung des entsprechenden Konvergenzsatzes für monotone, beschränkte Folgen reeller Zahlen. Wir benötigen jedoch zuerst einen Hilfssatz.

**Lemma 3.3.8.** Sei  $T \in \mathcal{L}(H)$  ein selbstadjungierter Operator mit  $T \geq 0$ . Dann gilt

$$|\langle Tx, y \rangle|^2 \leq \langle Tx, x \rangle \langle Ty, y \rangle$$

für alle  $x, y \in H$ .

*Beweis.* Sei auf  $H \times H$  eine Sesquilinearform  $(\cdot, \cdot)$  durch  $(x, y) = \langle Tx, y \rangle$  für alle  $x, y \in H$  definiert. Für diese gilt die Cauchy-Schwarzsche Ungleichung  $|(x, y)|^2 \leq (x, x)(y, y)$ . Durch Einsetzen der Definition von  $(\cdot, \cdot)$  folgt die Behauptung.  $\square$

**Satz 3.3.9.** Es sei  $(T_n)_{n \in \mathbb{N}}$  eine monoton wachsende und beschränkte Folge von selbstadjungierten Operatoren, das heißt  $T_n \leq T_m$  für  $n \leq m$  und  $\sup_{n \in \mathbb{N}} \|T_n\| < \infty$ . Dann existiert ein selbstadjungierter Operator  $T \in \mathcal{L}(H)$  mit  $T_n \xrightarrow{s} T$ .

*Beweis.* Sei  $c = \sup_{n \in \mathbb{N}} \|T_n\|$  und  $n \leq m$ . Dann ist  $T_m - T_n \geq 0$  aufgrund der Monotonie. Wir wenden Lemma 3.3.8 auf  $T = T_m - T_n$  an. Dies ergibt

$$|\langle (T_m - T_n)x, y \rangle|^2 \leq \langle (T_m - T_n)x, x \rangle \langle (T_m - T_n)y, y \rangle \leq \langle (T_m - T_n)x, x \rangle \cdot 2c \|y\|^2$$

durch Abschätzung mit der Operatornorm. Setze nun  $y = (T_m - T_n)x$ . Dann gilt

$$\|(T_m - T_n)x\|^4 \leq \langle (T_m - T_n)x, x \rangle \cdot 2c \|(T_m - T_n)x\|^2$$

und daraus ergibt sich

$$\|(T_m - T_n)x\|^2 \leq 2c \langle (T_m - T_n)x, x \rangle,$$

indem man durch  $\|(T_m - T_n)x\|^2$  dividiert, falls dies  $\neq 0$  ist. Ansonsten gilt die Ungleichung trivialerweise. Weil die Folge  $(\langle T_n x, x \rangle)_{n \in \mathbb{N}}$  als monotone, beschränkte Zahlenfolge konvergiert, folgt daraus, dass  $(T_n x)_{n \in \mathbb{N}}$  eine Cauchyfolge in  $H$  ist. Es existiert also ein  $y_x$  mit  $\lim_{n \rightarrow \infty} T_n x = y_x$ . Wir definieren  $T(x) := y_x$ . Dann ist  $T$  ein linearer Operator. Nach einer Folgerung aus dem Banach-Steinhaus-Theorem (Korollar 2.3.3) ist  $T \in \mathcal{L}(H)$ . Es bleibt noch zu zeigen, dass  $T^* = T$  gilt, dass  $T$  also selbstadjungiert ist. Weil jedoch  $T_n^* = T_n$  ist, gilt  $\langle T_n x, y \rangle = \langle x, T_n y \rangle$  für alle  $x, y \in H$ . Daraus folgt aber durch Grenzübergang, dass auch für die Grenzwerte die Gleichheit  $\langle Tx, y \rangle = \langle x, Ty \rangle$  gelten muss.  $\square$

## 3.4 Spektrum mit Resolvente

Um eine Motivation für die folgenden Definitionen zu bekommen, betrachten wir die Gleichung

$$\int_0^1 k(s, t) f(t) \lambda(dt) - \lambda f(s) = g(s) \quad (3.4.1)$$

Dabei sei  $g \in L^2[0, 1]$  eine gegebene Funktion und  $\lambda \in \mathbb{C}$  ein Parameter. Das Problem besteht darin, die Funktion  $f \in L^2[0, 1]$ , welche die Gleichung löst, zu suchen. Man kann das Problem durch Operatoren umformulieren. Definieren wir  $T$  durch  $(Tf)(s) = \int_0^1 k(s, t) f(t) \lambda(dt)$  für  $f \in L^2[0, 1]$ , so lautet obige Gleichung  $Tf - \lambda f = (T - \lambda I)f = g$ . Für eine Lösung  $f$  dieser Gleichung ergeben sich, um eine sinnvolle mathematische Behandlung zu ermöglichen, einige Anforderungen. Diese übertragen sich auf gewünschte Eigenschaften des Operators  $T$ . Wir wollen fordern:

- (i) Eindeutigkeit der Lösung, falls eine solche existiert. Eine Lösung erfüllt  $(T - \lambda I)f = g$ , und wenn diese eindeutig sein soll, dann muss der Operator  $T - \lambda I$  injektiv sein.

- (ii) Existenz einer Lösung für beliebige vorgegebene  $g \in L^2[0, 1]$ . Daraus folgt, dass  $T - \lambda I$  surjektiv sein muss. Damit fordern wir also zusammen mit (i), dass  $(T - \lambda I)^{-1}$  existiert und auf ganz  $L^2[0, 1]$  definiert ist. Die eindeutig bestimmte Lösung ist dann gegeben durch

$$f = (T - \lambda I)^{-1}g.$$

- (iii) Stetigkeit. Wenn  $g_n \rightarrow g$  für die vorgegebenen Funktionen gilt, dann soll auch für die Lösungen  $f_n \rightarrow f$  folgen. Dies bedeutet, dass  $(T - \lambda I)^{-1}$  stetig und damit beschränkt sein muss.

Diese Forderungen und den dahinterstehenden Kalkül wollen wir nun mathematisch formulieren und präzisieren. Folgende Definitionen mögen etwas technisch wirken, man sollte aber nicht vergessen, dass es sich zunächst lediglich um Begriffsbildungen handelt, mit denen man weiterrechnen und die man anwenden kann. Die Bedeutung und der Sinngehalt dieser Begriffe werden uns später ausreichend klar werden.

Sei  $H$  ein Hilbertraum und  $T \in \mathcal{L}(H)$  ein beschränkter linearer Operator auf  $H$ .

**Definition 3.4.1** (Resolvente). Eine Zahl  $\lambda \in \mathbb{C}$  gehört zur sogenannten *Resolventenmenge*  $\rho(T)$  von  $T$ , wenn es einen auf ganz  $H$  definierten, beschränkten, linearen Operator  $B_\lambda$  gibt, für den  $B_\lambda(T - \lambda I) = (T - \lambda I)B_\lambda = I$  gilt. Ist diese Gleichung erfüllt, dann ist  $B_\lambda = (T - \lambda I)^{-1}$  die Umkehrabbildung des Operators  $T - \lambda I$ .  $B_\lambda$  heißt dann die *Resolvente* von  $T$  im Punkt  $\lambda \in \rho(T)$  und wir schreiben  $R_\lambda := B_\lambda = (T - \lambda I)^{-1}$ .

**Bemerkung 3.4.2.** Sei  $B_\lambda$  ein auf ganz  $H$  definierter, linearer Operator mit  $B_\lambda(T - \lambda I) = (T - \lambda I)B_\lambda = I$ . Dann ist  $B_\lambda = (T - \lambda I)^{-1}$ . Nach einer Folgerung aus dem Satz von der offenen Abbildung (Lemma 2.4.2) ist  $B_\lambda$  beschränkt, denn  $T - \lambda I$  ist stetig und bildet  $H$  surjektiv auf  $H$  ab. Also gilt  $B_\lambda = R_\lambda$  und damit  $\lambda \in \rho(T)$ .

Wir werden  $B_\lambda$  immer für mögliche Resolventenkandidaten schreiben und  $R_\lambda$  nur dann benutzen, wenn es sich tatsächlich um die geforderte Resolvente handelt. Dies mag die Verwendung eines doppelten Symbols rechtfertigen. Eine Zahl  $\lambda \in \mathbb{C}$  gehört also genau dann zur Resolventenmenge  $\rho(T)$ , wenn sich der Operator  $T - \lambda I$  tatsächlich invertieren lässt, die Resolvente  $R_\lambda$  also existiert. Mit dieser Definition ausgestattet, kommen wir nun zur Definition des Spektrums eines Operators. Dieser Begriff ist einer der fruchtbarsten in der gesamten Operatoretheorie, denn er ermöglicht in vielen Fällen die Charakterisierung der wichtigsten Eigenschaften eines Operators. Den Rest dieses Skriptes werden wir uns fast ausschließlich mit Theorien beschäftigen, die direkt oder indirekt das Spektrum eines Operators zum Gegenstand haben.

**Definition 3.4.3** (Spektrum). Die Menge  $\sigma(T) := \mathbb{C} \setminus \rho(T)$  heißt das *Spektrum* des Operators  $T$ .

Das Spektrum einer Matrix  $A$  besteht in der linearen Algebra aus allen Eigenwerten, also aus allen  $\lambda \in \mathbb{C}$ , für die das Gleichungssystem  $A - \lambda I = 0$  nicht nur die triviale Lösung besitzt, für die also  $\det(A - \lambda I) = 0$  ist. Dies bedeutet auch, dass die Abbildung  $A - \lambda I$  nicht injektiv ist, was äquivalent dazu ist, dass sie nicht surjektiv ist. Auf unendlichdimensionalen Räumen gilt diese Äquivalenz nicht mehr. Aus diesem Grund benötigten wir obige erweiterte Definition des Spektrums. Eigenwerte sind in der Funktionalanalysis lediglich Sonderfälle von Einträgen aus dem Spektrum, sie werden jedoch ähnlich wie bei Matrizen definiert.

**Definition 3.4.4** (Eigenwert, Eigenvektor).  $\lambda \in \mathbb{C}$  heißt ein *Eigenwert* von  $T$ , wenn es ein  $x \in H$  mit  $x \neq 0$  gibt, so dass  $Tx = \lambda x$  ist.  $x$  heißt dann ein *Eigenvektor* von  $T$  zum Eigenwert  $\lambda$ .

**Bemerkung 3.4.5.** Dass  $\lambda \in \mathbb{C}$  ein Eigenwert ist, ist äquivalent dazu, dass  $(T - \lambda I)$  nicht injektiv ist, denn es gilt  $(T - \lambda I)x = 0$  für die Eigenvektoren  $x \in H$ . Jeder Eigenwert liegt damit also im Spektrum  $\sigma(T)$ .

Man kann das Spektrum folgendermaßen einteilen:

- (i) Als Punktspektrum  $\sigma_p(T)$  bezeichnet man die Menge der Eigenwerte:

$$\sigma_p(T) = \{\lambda \in \mathbb{C} \mid T - \lambda I \text{ ist nicht injektiv}\}.$$

- (ii) Als stetiges Spektrum  $\sigma_c(T)$  bezeichnet man die Menge

$$\sigma_c(T) = \{\lambda \in \mathbb{C} \mid T - \lambda I \text{ ist injektiv, aber nicht surjektiv, } (T - \lambda I)H \text{ ist dicht in } H\}.$$

Die Menge  $(T - \lambda I)H$  ist dabei das Bild von  $T$ .

### Beispiel 3.4.6.

- (a) Gegeben sei eine Matrix  $A \in \mathbb{C}^{N \times N}$  mit dem zugehörigen Operator  $T_A(x) = Ax$ . Sei  $\lambda \in \mathbb{C}$  ein Eigenwert der Matrix  $A$ . Dann existiert ein  $x \in \mathbb{C}^n$  mit  $x \neq 0$ , so dass  $Ax = \lambda x$  gilt. Natürlich ist dann auch  $\lambda$  ein Eigenwert und  $x$  ist ein Eigenvektor von  $T_A$ . Jeder Wert  $\lambda \in \sigma(T)$  im Spektrum von  $T_A$  ist umgekehrt auch ein Eigenwert der Matrix  $A$ . Damit folgt

$$\sigma(T_A) = \sigma_p(T_A) = \{\lambda \in \mathbb{C} \mid \lambda \text{ ist ein Eigenwert von } A\}.$$

Da alle endlichdimensionalen Hilberträume isomorph zu  $\mathbb{C}^N$  oder  $\mathbb{R}^N$  sind, ist somit für alle endlichdimensionalen Räume das Spektrum genau das Punktspektrum.

- (b) Sei  $H = L^2[0, 1]$  bezüglich des Lebesgue-Maßes. Wir definieren einen multiplikativen Operator  $T$  durch  $(Tf)(t) = tf(t)$ . Dann ist  $T \in \mathcal{L}(H)$  und es gilt  $\|T\| = \sup_{t \in [0, 1]} t = 1$ . Wir behaupten, dass  $T$  keine Eigenwerte besitzt. Zum Beweis nehmen wir zunächst an, dass  $\lambda$  ein Eigenwert und  $f$  ein dazugehöriger Eigenvektor sei. Es gilt also  $Tf = \lambda f$ , das heißt  $tf(t) = \lambda f$ , was nichts anderes heißt, als dass  $(t - \lambda)f(t) = 0$  für alle  $t \in [0, 1]$  sein muss. Daraus folgt, dass  $f(t) = 0$  für alle  $t \neq \lambda$  sein muss, also ist  $\text{supp } f \subseteq \{\lambda\}$ . Da das Lebesgue-Maß letzterer Menge 0 ist, ist  $f = 0$  fast überall und damit ist  $f = 0$  im Raum  $L^2[0, 1]$ , was für einen Eigenvektor nicht sein darf. Demzufolge ist  $\lambda$  auch kein Eigenwert.

Eine weitere Behauptung ist, dass  $\lambda \in \rho(T)$  gilt, wenn  $\lambda \notin [0, 1]$  ist. Dies beweisen wir nun. Sei  $\lambda \notin [0, 1]$ . Dann ist die Funktion  $\frac{1}{t - \lambda}$  auf  $[0, 1]$  ein beschränkter linearer Operator, denn sie besitzt dort keine Polstelle. Mit  $B_\lambda f = \frac{1}{t - \lambda} f$  gilt also

$$(B_\lambda(T - \lambda I)f)(t) = \left(\frac{1}{t - \lambda}\right)(t - \lambda)f(t) = f(t)$$

sowie

$$((T - \lambda I)B_\lambda f)(t) = (t - \lambda) \left(\frac{1}{t - \lambda}\right) f(t) = f(t),$$

also ist  $B_\lambda = R_\lambda$  eine Resolvente für  $\lambda$  und damit muss  $\lambda \in \rho(T)$  sein. Wir haben also  $\sigma(T) \subseteq [0, 1]$ .

Als nächstes wollen wir die umgekehrte Richtung,  $[0, 1] \subseteq \sigma(T)$  zeigen. Dazu beweisen wir also, dass  $\lambda \in \sigma(T)$  für  $\lambda \in [0, 1]$  ist. Angenommen, für diese  $\lambda$  wäre  $\lambda \in \rho(T)$ . Dann existiert  $R_\lambda \in \mathcal{L}(H)$  mit  $R_\lambda(T - \lambda I)f = (T - \lambda I)R_\lambda f = f$  für alle  $f \in H$ . Wir entnehmen daraus, dass  $\|f\| \leq \|R_\lambda\| \|(T - \lambda I)f\|$  gelten muss. Nun definieren wir eine Funktion  $f_\varepsilon$  durch  $f(t) = 1$ , falls  $t \in (\lambda - \varepsilon, \lambda + \varepsilon)$  ist und  $f(t) = 0$  sonst. Für dieses  $f_\varepsilon$  gilt nach obiger Annahme

$$\|f_\varepsilon\| \leq \|R_\lambda\| \|(T - \lambda I)f_\varepsilon\| \leq \|R_\lambda\| \|(t - \lambda)f(t)\| \leq \|R_\lambda\| \varepsilon \|f_\varepsilon\|,$$

da  $(t - \lambda)f(t)$  nur dann nicht Null ist, wenn  $|t - \lambda| < \varepsilon$  ist. Da dann aber  $\|R_\lambda\| \varepsilon \geq 1$  gilt und man  $\varepsilon$  beliebig klein machen kann, folgt ein Widerspruch, da  $\|R_\lambda\| = 0$  sein müsste, um die Ungleichung zu erfüllen. Also ist  $\lambda \in \sigma(T)$  und wir halten insgesamt fest, dass  $\sigma(T) = [0, 1]$  und  $\sigma_p(T) = \emptyset$  gilt.

- (c) Es gibt Operatoren, deren Bilder auf hochgradig nichttriviale Weise von den Ausgangswerten abhängen. Wir betrachten beispielsweise den sogenannten *Volterraschen Integraloperator*  $T$ , der durch  $(T_k f)(t) = \int_0^t k(t, s) f(s) \lambda(ds)$  für alle  $f \in H = L^2[0, 1]$  definiert ist. Dabei sei  $k(t, s)$  eine stetige Funktion auf  $[0, 1] \times [0, 1]$ . Man kann zeigen, dass  $\sigma(T_k) = \{0\}$  gilt, dass also das Spektrum des Integraloperators nur die Null umfasst, egal wie  $k$  gewählt wurde. Ein Spezialfall dieser Operatoren ist der bekannte Fall  $k(t, s) = 1$ . Hier haben wir ein Beispiel vorliegen, in dem das Spektrum eines Operators tatsächlich nichts über seine Natur aussagt.

Wir kommen nun zu einigen Eigenschaften des Spektrums beziehungsweise der Resolventenmenge. Sei hierzu  $T \in \mathcal{L}$  ein stetiger, linearer Operator auf  $H$ . Die Aussagen über die Resolventenidentitäten sind eher technischer Natur, bei vielen Beweisen in weiterführenden Themen jedoch von großem Nutzen.

**Bemerkung 3.4.7.**

- (i) Für das Spektrum des adjungierten Operators  $T^*$  gilt

$$\sigma(T^*) = \{\bar{\lambda} \mid \lambda \in \sigma(T)\}, \quad \rho(T^*) = \{\bar{\lambda} \mid \lambda \in \rho(T)\}.$$

- (ii) Es gelten die folgenden beiden Resolventenidentitäten:

$$R_\lambda(T) - R_{\lambda'}(T) = (\lambda - \lambda')R_\lambda(T)R_{\lambda'}(T) = (\lambda - \lambda')R_{\lambda'}(T)R_\lambda(T)$$

für  $\lambda, \lambda' \in \rho(T)$ , sowie

$$R_\lambda(T) - R_\lambda(S) = R_\lambda(T)(S - T)R_\lambda(S)$$

für  $\lambda \in \rho(T) \cap \rho(S)$ .

*Beweis.* Der Beweis der Aussage (i) ist eine Übungsaufgabe. Für die erste Resolventenidentität rechnen wir nach

$$\begin{aligned} (T - \lambda I)[R_{\lambda'}(T) + (\lambda - \lambda')R_\lambda(T)R_{\lambda'}(T)] &= (T - \lambda I)R_{\lambda'}(T) + (\lambda - \lambda') \underbrace{(T - \lambda I)R_\lambda(T)}_{=I} R_{\lambda'}(T) \\ &= [(T - \lambda I + (\lambda - \lambda')I)R_{\lambda'}(T)] = (T - \lambda' I)R_{\lambda'}(T) = I. \end{aligned}$$

Analog folgt  $[R_{\lambda'}(T) + (\lambda' - \lambda)R_\lambda(T)R_{\lambda'}(T)](T - \lambda I) = I$  und damit

$$R_\lambda(T) = R_{\lambda'}(T) + (\lambda - \lambda')R_\lambda(T)R_{\lambda'}(T).$$

□

**Satz 3.4.8.** Für  $T \in \mathcal{L}(H)$  gelten folgende Eigenschaften:

- (i)  $\rho(T)$  ist offen und  $\sigma(T)$  ist abgeschlossen.
- (ii) Wenn  $\lambda_0 \in \rho(T)$  und  $|\lambda - \lambda_0| < \|R_{\lambda_0}(T)\|^{-1}$  ist, dann ist auch  $\lambda \in \rho(T)$  und die Resolventenabbildung  $R : \lambda \mapsto R_\lambda$  von  $T$  ist eine analytische operatorwertige Funktion  $\rho(T) \rightarrow \mathcal{L}(H)$  mit der Potenzreihenentwicklung

$$R_\lambda(T) = \sum_{n=0}^{\infty} (\lambda - \lambda_0)^n R_{\lambda_0}(T)^{n+1}. \quad (3.4.2)$$

- (iii) Wenn  $|\lambda| > \|T\|$  ist, dann ist  $\lambda \in \rho(T)$  und es gilt die Entwicklung

$$R_\lambda(T) = - \sum_{n=0}^{\infty} \lambda^{-n-1} T^n.$$



*Beweis.* Wie man als Übungsaufgabe nachrechnen kann, folgt (i) aus (ii). Beweisen wir also nun Aussage (ii). Wir erinnern uns zunächst an folgende Tatsache:  $\sum_{n=0}^{\infty} A_n$  mit  $A_n \in \mathcal{L}(H)$  konvergiert in der Operatornorm und definiert einen neuen Operator aus  $\mathcal{L}(H)$  genau dann, wenn  $\sum_{n=0}^{\infty} \|A_n\| < \infty$  ist. Hiermit ergibt sich nun, dass die Reihe in (3.4.2) konvergiert, denn es gilt

$$\|(\lambda - \lambda_0)^n R_{\lambda_0}(T)^{n+1}\| \leq \underbrace{(|\lambda - \lambda_0| \|R_{\lambda_0}(T)\|)^n}_{=:q < 1} \|R_{\lambda_0}(T)\|,$$

wobei wir benutzen, dass  $\|A^n\| \leq \|A\|^n$  für jeden Operator  $A$  gilt. Damit besitzt die Norm der Reihe die Gestalt  $\sum_{n=0}^{\infty} \|R_{\lambda_0}(T)\| q^n$  und die Reihe selbst konvergiert daher. Sie definiert nach obiger Bemerkung einen beschränkten, linearen Operator  $B_\lambda = \sum_{n=0}^{\infty} (\lambda - \lambda_0)^n R_{\lambda_0}(T)^{n+1}$ . Für diesen folgt

$$\begin{aligned} (T - \lambda I)B &= (T - \lambda_0 I)B + (\lambda_0 - \lambda)B \\ &= \sum_{n=0}^{\infty} (\lambda - \lambda_0)^n \underbrace{(T - \lambda_0 I)R_{\lambda_0}(T)^{n+1}}_{R_{\lambda_0}(T)^n} - \sum_{n=0}^{\infty} (\lambda - \lambda_0)^{n+1} R_{\lambda_0}(T)^{n+1} \\ &= (\lambda - \lambda_0)^0 R_{\lambda_0}(T)^0 = I. \end{aligned}$$

Analog folgt die Aussage  $B(T - \lambda I) = I$  und damit gilt  $B = (T - \lambda I)^{-1} = R_\lambda(T)$ .

Wir zeigen nun (iii). Es gilt für jedes Glied der Reihe

$$\|\lambda^{-n-1} T^n\| \leq |\lambda^{-1}| \underbrace{(|\lambda|^{-1} \|T\|)^n}_{=:q < 1}$$

und somit konvergiert diese. Wir definieren nun

$$S := - \sum_{n=0}^{\infty} \lambda^{-n-1} T^n$$

und wenden den Operator  $T - \lambda I$  auf  $S$  an. Wir erhalten

$$(T - \lambda I)S = - \sum_{n=0}^{\infty} \lambda^{-n-1} T^{n+1} + \sum_{n=0}^{\infty} \lambda^{-n} T^n = \lambda^0 T^0 = I,$$

da wir eine Teleskopsumme vorliegen haben. Analog ergibt sich auch die Aussage für  $S(T - \lambda I)$  und somit folgt, dass  $S = R_\lambda(T)$  sein muss. Dies beendet den Beweis.  $\square$

Mit  $r(T)$  wollen wir den *Spektralradius* von  $T$  bezeichnen, also den Radius des kleinsten Kreises um den Nullpunkt, der das Spektrum von  $T$  vollständig enthält. Er lässt sich oft auf sehr angenehme Weise berechnen.

**Satz 3.4.9.** Für  $T \in \mathcal{L}(H)$  gelten folgende Aussagen.

(i) Das Spektrum von  $T$  ist nicht leer.

(ii)  $r(T) = \sup \{|\lambda| \mid \lambda \in \sigma(T)\} = \lim_{n \rightarrow \infty} \|T^n\|^{1/n}$ .

*Beweis.* Die erste Aussage beweisen wir indirekt mit ein paar Erkenntnissen aus der Funktionentheorie. Sei angenommen, dass  $\sigma(T) = \emptyset$  leer ist und somit  $\rho(T) = \mathbb{C}$  gilt. Wir wählen nun ein  $\lambda \in \mathbb{C}$  mit  $|\lambda| > \|T\|$  und fixieren ein  $x \in H$ . Es gilt

$$\|(T - \lambda I)x\| \geq \|\lambda x\| - \|Tx\| \geq |\lambda| \|x\| - \|T\| \|x\| = (|\lambda| - \|T\|) \|x\|.$$

Dieselbe Abschätzung kann man aufgrund der Symmetrie in der Dreiecksungleichung auch für  $\|T\| > |\lambda|$  durchführen und erhält dann zusammen, dass  $\|(T - \lambda I)x\| \geq ||\lambda| - \|T|| \|x\|$  ist. Nun setzen wir  $y =$

$(T - \lambda I)x$ . Da  $\lambda \in \rho(T)$  ist, existiert die Resolvente  $R_\lambda$  und es ist  $x = R_\lambda(T)y$ . Setzen wir dies in die obige Abschätzung ein, so erhalten wir, dass  $\|R_\lambda(T)\| \neq T|\lambda - \|T\|^{-1}$  ist, also ist nach Satz 3.4.8 (iii)  $R_\lambda(T)$  analytisch und somit  $\langle R_\lambda(T)x, y \rangle$  holomorph auf  $\rho(T) = \mathbb{C}$ . Es gilt weiter

$$|\langle R_\lambda(T)x, y \rangle| \leq \frac{1}{|\lambda| - \|T\|} \|x\| \|y\| \quad (3.4.3)$$

für alle  $|\lambda| \neq \|T\|$  nach der Cauchy-Schwarz-Ungleichung. Daraus und aus der lokalen Stetigkeit der Resolventenabbildung folgt, dass  $\langle R_\lambda(T)x, y \rangle$  beschränkt auf ganz  $\mathbb{C}$  ist. Nach dem Satz von Liouville ist die Funktion damit konstant, also  $\langle R_\lambda(T)x, y \rangle = c$  für alle  $\lambda \in \mathbb{C}$ . Wegen Gleichung (3.4.3) ist

$$c = \lim_{|\lambda| \rightarrow \infty} c = \lim_{|\lambda| \rightarrow \infty} \langle R_\lambda(T)x, y \rangle = 0$$

für alle  $x, y \in H$  und somit folgt, dass  $R_\lambda(T) = 0$  ist. Dies ist ein Widerspruch zu  $R_\lambda(T) = (T - \lambda I)^{-1}$ , also kann das Spektrum von  $T$  nicht leer sein.  $\square$

Wir wissen mit Satz 3.4.8 (iii) bereits, dass  $|\lambda| \leq \|T\|$  ist, also muss der Spektralradius ebenfalls kleiner als  $\|T\|$  sein. Es kann nicht immer  $r(T) = \|T\|$  gelten, wie das Beispiel des Volterraschen Integraloperators zeigt, denn dort ist  $r(T) = 0$ . Der folgende Satz liefert uns jedoch unter bestimmten Bedingungen die Gleichheit.

**Satz 3.4.10.** *Wenn  $T \in \mathcal{L}(H)$  normal ist, dann gilt  $r(T) = \|T\|$ .*

### 3.4.1 Das Spektrum selbstadjungierter Operatoren

Sei  $T = T^* \in \mathcal{L}(H)$  ein selbstadjungierter Operator. Da  $T$  damit auch normal ist, gilt damit nach Satz 3.4.10 für den Spektralradius  $r(T) = \|T\|$ .

**Satz 3.4.11.** *Es ist  $\lambda \in \rho(T)$  für  $\lambda \in \mathbb{C}$  genau dann, wenn eine Konstante  $c_\lambda > 0$  existiert, so dass  $\|(T - \lambda I)x\| \geq c_\lambda \|x\|$  für alle  $x \in H$  ist. Ist dies erfüllt, dann gilt  $\|R_\lambda(T)\| \leq c_\lambda^{-1}$ .*

*Beweis.* Wir beginnen mit der einfacheren Richtung. Sei  $\lambda \in \rho(T)$ . Dann gilt:

$$\|x\| = \|R_\lambda(T - \lambda I)x\| \leq \|R_\lambda(T)\| \|(T - \lambda I)x\|$$

und wir setzen  $c_\lambda = \|R_\lambda(T)\|^{-1}$ . Die Rückrichtung ist nicht ganz so einfach und wir werden den Beweis in drei Schritten durchführen.

- (i) Zuerst zeigen wir die Injektivität von  $(T - \lambda I)$ . Sei  $(T - \lambda I)x_1 = (T - \lambda I)x_2$ . Dann gilt auch  $0 = (T - \lambda I)(x_1 - x_2)$  und  $0 = \|(T - \lambda I)(x_1 - x_2)\| \geq c_\lambda \|x_1 - x_2\|$  nach Voraussetzung. Hieraus ergibt sich, dass  $x_1 - x_2 = 0$  sein muss und somit ist  $(T - \lambda I)$  injektiv.
- (ii) Als nächstes zeigen wir die Abgeschlossenheit des Bildes  $(T - \lambda I)H$ . Es konvergiere dafür  $(T - \lambda I)x_n \rightarrow y$ , wobei  $x_n, y \in H$  seien. Wir müssen zeigen, dass  $y \in (T - \lambda I)H$  ist. Es gilt die Abschätzung

$$\|(T - \lambda I)(x_n - x_m)\| = \left\| \underbrace{(T - \lambda I)x_n}_{:=y_n} - \underbrace{(T - \lambda I)x_m}_{:=y_m} \right\| \geq c_\lambda \|x_n - x_m\|,$$

also gilt  $\|x_n - x_m\| \leq c_\lambda^{-1} \|y_n - y_m\|$ . Unsere Folge  $(y_n)_{n \in \mathbb{N}}$  ist konvergent, also eine Cauchyfolge und deshalb muss auch  $(x_n)_{n \in \mathbb{N}}$  eine Cauchyfolge sein, konvergiert also in  $H$ , sagen wir  $x_n \rightarrow x$ . Da  $(T - \lambda I)$  stetig ist, folgt die Konvergenz  $(T - \lambda I)x_n \rightarrow (T - \lambda I)x$ . Da  $(T - \lambda I)x_n$  auch gegen  $y$  konvergiert, gilt aufgrund der Eindeutigkeit des Grenzwerts  $y = (T - \lambda I)x$  und somit ist  $(T - \lambda I)H$  abgeschlossen.

(iii) Zu guter letzt beweisen wir, dass  $(T - \lambda I)$  den Hilbertraum  $H$  wieder auf ganz  $H$  abbildet, also surjektiv ist. Nach Schritt (ii) wissen wir bereits, dass  $H_1 = (T - \lambda I)H$  ein abgeschlossener linearer Teilraum von  $H$  ist. Nach dem ersten Satz von Riesz (Satz 1.4.6) haben wir dadurch eine orthogonale Zerlegung von  $H$  in  $H = H_1 \oplus H_1^\perp$ .

Sei  $u \in H_1^\perp$ . Dann gilt  $\langle (T - \lambda I)x, u \rangle = 0$  für alle  $x \in H$  und es ergibt sich somit  $\langle Tx, u \rangle = \langle \lambda Ix, u \rangle = \lambda \langle x, u \rangle$ . Da  $T$  selbstadjungiert ist, gilt also  $\langle x, Tu \rangle = \langle x, \bar{\lambda}u \rangle$  und somit auch  $\langle x, (T - \bar{\lambda}I)u \rangle = 0$ .

Sei  $u \neq 0$ . Wir setzen  $x = u$  und erhalten

$$\langle Tu, u \rangle = \lambda \langle u, u \rangle = \langle u, \bar{\lambda}u \rangle = \langle u, Tu \rangle = \langle Tu, u \rangle,$$

also  $\bar{\lambda} = \lambda$ . Daraus folgt, dass  $(T - \lambda I)u = (T - \bar{\lambda}I)u$  ist. Da aber

$$\langle (T - \lambda I)x, u \rangle = 0 = \langle x, (T - \bar{\lambda}I)u \rangle = \langle x, (T - \lambda I)u \rangle$$

für alle  $x \in H$  ist, muss  $(T - \lambda I)u = 0$  sein. Aus der Injektivität von  $(T - \lambda I)$  folgt  $u = 0$ , was wir ausgeschlossen hatten. Also gilt  $H_1^\perp = \{0\}$  und somit ist  $H = H_1$ , woraus folgt, dass  $(T - \lambda I)$  surjektiv ist. Die Beschränktheit von  $(T - \lambda I)^{-1}$  folgt aus der Bedingung, dass  $\|(T - \lambda I)x\| \geq c_\lambda \|x\|$  sein soll. Wir setzen wieder  $y = (T - \lambda I)x$  und es ergibt sich  $\|y\| \geq c_\lambda \|(T - \lambda I)^{-1}y\|$ . Nach kurzem Umformen erhält man  $\|(T - \lambda I)^{-1}\| \leq c_\lambda^{-1}$ .

□

**Satz 3.4.12.** Sei  $T \in \mathcal{L}(H)$ . Dann gilt

$$\|T\| = \sup_{\|x\| \leq 1} \|Tx\| = \sup_{\|x\| \leq 1} \sup_{\|y\| \leq 1} |\langle Tx, y \rangle|.$$

Wenn  $T$  selbstadjungiert ist, gilt

$$\|T\| = \sup_{\|x\| \leq 1} |\langle Tx, x \rangle|.$$

*Beweis.* Wir zeigen nur den selbstadjungierten Fall und definieren  $c = \sup_{\|x\| \leq 1} |\langle Tx, x \rangle|$ . Da  $|\langle Tx, x \rangle| \leq \|Tx\| \|x\| \leq \|T\| \|x\|^2$  gilt, ist  $c \leq \|T\|$ . Nun müssen wir noch zeigen, dass  $\|T\| \leq c$  gilt. Dieser Beweis ist nicht sofort intuitiv klar und bedarf ein wenig Arbeit, am Ende wird sich das gewünschte Resultat aber sehr einfach für uns ergeben. Für  $\alpha > 0$  gilt, wie man durch Umformen sieht,

$$\begin{aligned} \|Tx\|^2 &= \langle Tx, Tx \rangle = \langle T^2x, x \rangle \\ &= \frac{1}{4} (\langle T(\alpha x + \alpha^{-1}Tx), \alpha x + \alpha^{-1}Tx \rangle - \langle T(\alpha x - \alpha^{-1}Tx), \alpha x - \alpha^{-1}Tx \rangle) \\ &\leq \frac{1}{4} (c \|\alpha x + \alpha^{-1}Tx\|^2 + c \|\alpha x - \alpha^{-1}Tx\|^2) \leq \frac{c}{4} (2\|\alpha x\|^2 + 2\|\alpha^{-1}Tx\|^2) \\ &= \frac{c}{2} (\alpha^2 \|x\|^2 + \alpha^{-2} \|Tx\|^2), \end{aligned}$$

wobei wir im vorletzten Schritt die Parallelogrammidentität (Gleichung (1.3.4)) ausgenutzt haben. Wir wählen nun  $\alpha^2 = \frac{\|Tx\|}{\|x\|}$ , wobei o.B.d.A.  $\|Tx\| \neq 0$  und  $\|x\| \neq 0$  seien. Dann ergibt sich

$$\|Tx\|^2 \leq \frac{c}{2} \left( \frac{\|Tx\|}{\|x\|} \|x\|^2 + \frac{\|x\|}{\|Tx\|} \|Tx\|^2 \right) = c \|Tx\| \|x\|.$$

Damit sind wir auch schon fertig, da nun  $\|T\|^2 = \sup_{\|x\| \leq 1} \|Tx\|^2 \leq c \|T\|$  folgt und somit  $\|T\| \leq c$  gilt. □

**Definition 3.4.13.** Für  $T \in \mathcal{L}(H)$  definieren wir

$$m_T = \inf_{\|x\| \leq 1} \langle Tx, x \rangle, \quad M_T = \sup_{\|x\| \leq 1} \langle Tx, x \rangle.$$

Wir nennen  $m_T$  die *untere Grenze* und  $M_T$  die *obere Grenze* von  $T$ . Es gilt dann

$$m_T \|x\|^2 \leq \langle Tx, x \rangle \leq M_T \|x\|^2.$$

**Lemma 3.4.14.** Für selbadjungierte Operatoren  $T \in \mathcal{L}(H)$  gilt  $\sigma(T) \subseteq [m_T, M_T]$ . Insbesondere enthält  $\sigma(T)$  nur reelle Zahlen.

*Beweis.* Sei  $\lambda_0 \in \mathbb{C}$  mit  $\lambda_0 \notin [m_T, M_T]$ . Sei weiterhin  $c = \inf_{\mu \in [m_T, M_T]} |\lambda_0 - \mu|$ . Da  $\lambda_0$  nicht in  $[m_T, M_T]$  liegt, ist  $c > 0$ . Für  $x \in H$  mit  $\|x\| = 1$  gilt

$$\|(T - \lambda_0 I)x\| = \|x\| \|(T - \lambda_0 I)x\| \geq |\langle (T - \lambda_0 I)x, x \rangle| = |\langle Tx, x \rangle - \lambda_0 \langle x, x \rangle| = |\langle Tx, x \rangle - \lambda_0| \geq c,$$

da  $\langle Tx, x \rangle \in [m_T, M_T]$  ist und daraus folgt, dass für beliebiges  $x \in H$  die Ungleichung  $\|(T - \lambda_0 I)x\| \geq c \|x\|$  gilt. Nach Satz 3.4.11 ist somit  $\lambda_0 \in \rho(T)$ .  $\square$

**Beispiel 3.4.15.** Sei  $H = L^2[0, 1]$  und  $Tf(t) = tf(t)$ . Dann ist das Skalarprodukt definiert durch  $\langle Tf, f \rangle = \int_0^1 t |f(t)|^2 \lambda(dt)$  und es gilt offensichtlich  $0 \cdot \langle f, f \rangle \leq \langle Tf, f \rangle \leq 1 \cdot \langle f, f \rangle$ , also folgt nach der eben bewiesenen Folgerung, dass  $\sigma(T) \subseteq [0, 1]$  ist.

**Satz 3.4.16.** Sei  $T \in \mathcal{L}(H)$  selbstadjungiert. Dann sind die Eigenwerte von  $T$  reell. Zu verschiedenen Eigenwerten gehörende Eigenvektoren stehen senkrecht aufeinander.

*Beweis.* Nach Lemma 3.4.14 gilt  $\sigma(T) \subseteq [m_T, M_T]$ . Damit liegen also auch die Eigenwerte in  $[m_T, M_T]$  und sind damit reell. Sei nun  $Tx = \lambda x$  und  $Ty = \mu y$ , wobei  $x, y \neq 0$ ,  $\lambda, \mu \in \mathbb{R}$  und  $\lambda \neq \mu$  sei. Dann gilt

$$\lambda \langle x, y \rangle = \langle \lambda x, y \rangle = \langle Tx, y \rangle = \langle x, T^* y \rangle = \langle x, Ty \rangle = \langle x, \mu y \rangle = \bar{\mu} \langle x, y \rangle = \mu \langle x, y \rangle$$

und damit ist  $(\lambda - \mu) \langle x, y \rangle = 0$ , woraus wegen  $\lambda \neq \mu$  folgt, dass  $\langle x, y \rangle = 0$  sein muss.  $\square$

## 3.5 Kompakte Operatoren im Hilbertraum

### 3.5.1 Kompakte und vollstetige Operatoren in Banachräumen

Seien  $E$  und  $F$  Banachräume. Wir wiederholen und vertiefen zunächst einige topologische Kompaktheitsbegriffe.

**Definition 3.5.1.** Eine Menge  $M \subseteq E$  heißt

- (i) *relativ kompakt*, wenn man aus jeder Folge  $(x_n)_{n \in \mathbb{N}}$  von Elementen aus  $M$  eine in  $E$  konvergente Teilfolge  $(x_{n_k})_{k \in \mathbb{N}}$  auswählen kann.
- (ii) *kompakt*, wenn man aus jeder Folge  $(x_n)_{n \in \mathbb{N}}$  von Elementen aus  $M$  eine in  $E$  konvergente Teilfolge  $(x_{n_k})_{k \in \mathbb{N}}$  auswählen kann, für die  $\lim_{k \rightarrow \infty} x_{n_k} \in M$  gilt.

Eine Folge  $(x_n)_{n \in \mathbb{N}}$  heißt

- (iii) *schwach konvergent* gegen  $x \in E$ , wenn  $\lim_{n \rightarrow \infty} f(x_n) = f(x)$  für alle  $f \in E'$  gilt. Man schreibt dafür  $x_n \rightharpoonup x$ .

Diese Definition der schwachen Konvergenz ist allgemeiner als die in Definition 3.3.1 für Hilberträume, da jedes stetige lineare Funktional auf einem Hilbertraum nach dem zweiten Satz von Riesz durch das Skalarprodukt mit einem festen Element des Hilbertraumes gegeben ist.

**Beispiel 3.5.2.**

- (a) Sei  $E = \mathbb{R}^n$  oder  $E = \mathbb{C}^n$  oder  $E$  endlichdimensional. Dann ist  $M \subseteq E$  relativkompakt genau dann, wenn  $M$  beschränkt ist. Weiterhin ist  $M \subseteq E$  kompakt genau dann, wenn  $M$  beschränkt und abgeschlossen ist. Insbesondere sind  $B_r(x_0) = \{x \in E \mid \|x - x_0\| \leq r\}$  und  $S_r(x_0) = \{x \in E \mid \|x - x_0\| = r\}$  kompakt und  $U_r(x_0) = \{x \in E \mid \|x - x_0\| < r\}$  ist relativkompakt, wenn  $x_0 \in E$  und  $r \geq 0$  gilt.
- (b) Sei  $E$  ein unendlichdimensionaler Hilbertraum. Dann existiert ein Orthonormalsystem  $(x_n)_{n \in \mathbb{N}}$  in  $E$ . Es gilt für alle  $n, m \in \mathbb{N}$  mit  $n \neq m$ :

$$\|x_n - x_m\|^2 = \langle x_n - x_m, x_n - x_m \rangle = \langle x_n, x_n \rangle + \langle x_m, x_m \rangle = 2,$$

da die Vektoren  $x_n$  und  $x_m$  orthonormal sind. Es gilt also für diese  $n, m$  stets  $\|x_n - x_m\| = \sqrt{2}$  und damit ist  $(x_n)$  keine Cauchyfolge. Aus ihr kann man demnach auch keine konvergente Teilfolge auswählen. Da jedoch  $x_n \in B_1(0)$  für alle  $n \in \mathbb{N}$  gilt, ist  $B_1(0)$  nicht kompakt.

**Bemerkung 3.5.3.**

- (i) Jede relativ kompakte Menge  $M$  ist beschränkt. Für einen indirekten Beweis nehme man an,  $M$  wäre nicht beschränkt und finde eine Folge mit  $\|x_n\| \geq \|x_{n-1}\| + 1$ .
- (ii) Jede kompakte Menge  $M$  in  $E$  ist abgeschlossen, denn alle Teilfolgen einer in  $E$  gegen  $x$  konvergenten Folge von Elementen aus  $M$  besitzen ebenfalls den Limes  $x$ .
- (iii) Jede schwach konvergente Folge  $(x_n)$  ist normbeschränkt, also  $\sup_{n \in \mathbb{N}} \|x_n\| < \infty$ .
- (iv) Sei  $T \in \mathcal{L}(E, F)$ . Dann ist  $T$  stetig in der schwachen Konvergenz. Das bedeutet, dass aus  $x_n \rightharpoonup x$  auch  $T(x_n) \rightharpoonup T(x)$  folgt. Zum Beweis sei  $f \in F'$  und  $g(y) = f(T(y))$  für  $y \in E$ .  $g$  ist in  $E'$  als Komposition stetiger linearer Abbildungen selbst stetig und linear. Weil  $x_n \rightharpoonup x$  gilt, folgt  $g(x_n) = f(T(x_n)) \rightarrow g(x) = f(T(x))$ , das heißt  $T(x_n) \rightharpoonup T(x)$ .

**Definition 3.5.4.** Ein linearer Operator  $T : E \rightarrow F$  heißt

- (i) *kompakt*, wenn  $T$  jede beschränkte Menge  $M$  aus  $E$  in eine relativ kompakte Menge  $T(M)$  in  $F$  abbildet.
- (ii) *vollstetig*, wenn aus  $x_n \rightharpoonup x$  in  $E$  stets  $T(x_n) \rightarrow T(x)$  in  $F$  folgt.

Kompakte oder vollstetige Operatoren „verbessern“ also die gegebenen Voraussetzungen. Aus lediglich beschränkten Mengen werden kompakte und aus lediglich schwacher Konvergenz wird die starke.

**Beispiel 3.5.5.**

- (a) Seien  $E, F$  endlichdimensionale Banachräume. Jede lineare Abbildung von  $E$  nach  $F$  ist kompakt und vollstetig.
- (b) Sei  $E$  ein unendlichdimensionaler Hilbertraum und  $T = I$  die identische Abbildung. Dann gilt natürlich  $T(B_1(0)) = B_1(0)$ . Nach Beispiel 3.5.2(b) ist  $T$  damit nicht kompakt. Sei weiter  $(x_n)_n$  ein Orthonormalsystem. Dann gilt  $x_n \rightharpoonup 0$ , aber  $T(x_n) = x_n$  konvergiert nicht stark gegen 0 und besitzt auch keine konvergente Teilfolge. Damit ist  $T = I$  auch nicht vollstetig.

- (c) Komplizierter ist die Situation in einem unendlichdimensionalen Banachraum. Sei  $E = F = \ell^1(\mathbb{N})$ . Man kann zeigen, dass jede schwach konvergente Folge aus  $E$  auch in der Norm konvergiert. Damit ist die identische Abbildung  $I$  vollstetig.  $I$  ist hingegen nicht kompakt. Auf einen Beweis dieser Aussage verzichten wir hier.
- (d) Wir betrachten die Sobolevräume  $H^k(\Omega)$  und  $H_0^k(\Omega)$ , wobei  $\Omega$  eine beschränkte und offene Teilmenge des  $\mathbb{R}^n$  sei. Sei  $T_0 : H_0^{k+1}(\Omega) \rightarrow H_0^k(\Omega)$  die Einbettung. Dann ist  $T_0$  kompakt. Wenn  $\Omega$  einen  $C^1$ -Rand besitzt, dann ist die Einbettung  $T : H^{k+1}(\Omega) \rightarrow H^k(\Omega)$  ebenfalls kompakt. Diese beiden Aussagen bezeichnet man als *Satz von Rellich*.

**Beispiel 3.5.6.** Sei  $E = C^1[0, 1]$  mit der Norm  $\|f\|_1 = \sup_{t \in [0, 1]} |f(t)| + \sup_{t \in [0, 1]} |f'(t)|$  für  $f \in E$ . Sei weiter  $F = C[0, 1]$ . Für  $g \in F$  sei die Norm auf  $F$  definiert durch  $\|g\| = \sup_{t \in [0, 1]} |g(t)|$ . Es sei  $T : E \rightarrow F$  die Einbettung. Dann ist  $T$  kompakt.

*Beweis.* Sei  $f \in E$ . Dann gilt

$$|f(x) - f(x')| = \left| \int_x^{x'} f'(t) dt \right| \leq \sup_{s \in [0, 1]} |f'(s)| \left| \int_x^{x'} dt \right| \leq \|f\|_1 |x - x'|$$

für  $x, x' \in [0, 1]$ . Sei  $M$  beschränkt in  $E$ , also  $\sup_{f \in M} \|f\|_1 = c < \infty$ . Dann gilt  $|f(x) - f(x')| \leq c|x - x'|$ , das heißt  $M$  ist eine gleichgradig stetige Menge.  $T(M) = M$  ist beschränkt in  $F$ . Nach dem Satz von Arzelà-Ascoli kann man aus jeder Folge in dieser Menge eine gleichmäßig konvergente Teilfolge auswählen. Damit ist  $T(M)$  relativ kompakt in  $F$  und damit  $T$  kompakt.  $\square$

**Lemma 3.5.7.** Jede vollstetige lineare Abbildung  $T : E \rightarrow F$  ist stetig, das heißt  $T \in \mathcal{L}(E, F)$ .

**Satz 3.5.8.** Jeder kompakte lineare Operator  $T : E \rightarrow F$  ist vollstetig.

*Beweis.* Angenommen,  $T$  wäre nicht vollstetig. Dann gibt es eine Folge  $x_n \rightarrow x$  in  $E$  derart, dass  $(T(x_n))$  nicht gegen  $T(x)$  konvergiert. Damit existiert ein  $\varepsilon > 0$  und eine Folge  $(T(x_{n_k}))_k$  mit  $\|T(x_{n_k}) - T(x)\| \geq \varepsilon$  für alle  $n_k$ . Da  $(x_n)$  beschränkt ist (o.B.d.A.  $\|x_n\| \leq 1$ ), liegt  $(x_n)$  in  $B_1(0)$ . Die Menge  $\{T(x_n) \mid n \in \mathbb{N}\}$  ist relativ kompakt, da  $T$  kompakt ist. Damit müsste die relativ kompakte Teilmenge  $\{T(x_{n_k}) \mid k \in \mathbb{N}\}$  jedoch eine konvergente Teilfolge besitzen, Widerspruch.  $\square$

Es sei  $\mathcal{K}(E, F)$  die Menge aller kompakten linearen Operatoren von  $E$  nach  $F$ .

**Satz 3.5.9.**

- (i)  $\mathcal{K}(E, F)$  ist ein abgeschlossener linearer Teilraum von  $(\mathcal{L}(E, F), \|\cdot\|)$ .
- (ii) Seien  $E_1$  und  $F_1$  Banachräume. Wenn  $B \in \mathcal{L}(E_1, E)$ ,  $A \in \mathcal{L}(F, F_1)$  und  $T \in \mathcal{K}(E, F)$  ist, dann gilt  $ATB \in \mathcal{K}(E_1, F_1)$ .

**Bemerkung 3.5.10.** Sei  $E = F$ . Setze  $\mathcal{K}(E) = \mathcal{K}(E, E)$  und  $\mathcal{L}(E) = \mathcal{L}(E, E)$ . Aus  $A, B \in \mathcal{L}(E)$  und  $T \in \mathcal{K}(E)$  folgt dann  $ATB \in \mathcal{K}(E)$ . Damit ist  $\mathcal{K}(E)$  ein Ideal in der Algebra  $\mathcal{L}(E)$ .

Kompakte Operatoren bleiben „gutartig“, also kompakt, wenn man die Komposition mit „schlechteren“ Operatoren bildet.

**Definition 3.5.11.** Sei  $T \in \mathcal{L}(E, F)$ . Wir definieren  $T^t(f')(x) = f'(T(x))$  für  $x \in E$  und  $f' \in F'$ . Der Operator  $T^t$  heißt der zu  $T$  *transponierte* Operator.

Das Bild  $T^t(f')$  für  $f' \in F'$  liegt in  $E'$ , denn es gilt

$$\|T^t(f')(x)\| = \|f'(T(x))\| \leq \|f'\| \|T(x)\| \leq \|f'\| \|T\| \|x\|,$$

also ist  $T^t(f')$  beschränkt. Damit ist  $T^t$  eine lineare Abbildung von  $F'$  nach  $E'$ . Für ihre Norm gilt  $\|T^t(f')\| \leq \|f'\| \|T\|$ , also  $\|T^t\| \leq \|T\|$  und somit ist  $T^t \in \mathcal{L}(F', E')$ . Es gilt folgende Aussage über die Kompaktheit des transponierten Operators.

**Satz 3.5.12** (Satz von Schauder). *Wenn  $T \in \mathcal{L}(E, F)$  kompakt ist, dann ist  $T^t \in \mathcal{L}(F', E')$  ebenfalls kompakt.*

Der transponierte Operator verallgemeinert gewissermaßen den adjungierten Operator von Hilberträumen auf Banachräume. Wenn nämlich  $E = F$  ein Hilbertraum ist und  $T$  ein stetiger linearer Operator,  $T \in \mathcal{L}(E)$ , dann entspricht  $T^t$  dem adjungierten linearen Operator  $T^*$ . Es gilt nämlich in diesem Fall  $E' = E^-$ , wenn  $E^-$  den Raum  $E$  mit der veränderten Skalarmultiplikation  $(\lambda x)_{E^-} = (\bar{\lambda} x)_E$  darstellt. Wenn man also  $E'$  mit  $E$  aufgrund des Isomorphismus' identifiziert, dann ist  $T^t = T^*$ , ansonsten muss man den entsprechenden Operator auf  $E^-$  betrachten.

**Satz 3.5.13** (Riesz-Schauder-Theorem). *Sei  $E$  ein Banachraum,  $T \in \mathcal{L}(E)$  und  $\lambda \in \mathbb{C}, \lambda \neq 0$ . Dann gilt:*

- (i)  $\text{Range}(T - \lambda I) := \{(T - \lambda I)x \mid x \in E\}$  und  $\text{Kern}(T - \lambda I) := \{x \in E \mid (T - \lambda I)x = 0\}$  sind abgeschlossen.
- (ii)  $E$  ist ein unendlichdimensionaler Banachraum genau dann, wenn es eine Nullfolge  $(\lambda_n)_{n \in \mathbb{N}}$  in  $\mathbb{C}$  gibt mit  $\sigma(T) = \{0, \lambda_n \mid n \in \mathbb{N}\}$ . Dann sind alle  $\lambda \in \sigma(T), \lambda \neq 0$  Eigenwerte endlicher Vielfachheit.
- (iii) (Fredholmsche Alternative) Die Gleichung  $(T - \lambda)x = y$  ist für beliebige  $y \in E$  eindeutig lösbar genau dann, wenn  $(T - \lambda)x = 0$  nur die triviale Lösung  $x = 0$  besitzt, das heißt wenn  $\text{Kern}(T - \lambda I) = \{0\}$  gilt.

Die letzte Aussage lässt sich mit Hilfe des transponierten Operators etwas allgemeiner formulieren.  $(T - \lambda)x = y$  ist für beliebige  $y \in E$  lösbar genau dann, wenn  $y$  durch alle Lösungen  $f \in E'$  von  $(T^t - \lambda)f = 0$  annulliert wird, das heißt dass  $f(y) = 0$  sein muss, falls  $T^t f = \lambda f$  ist.

Sei im folgenden  $\mathcal{F}(E, F) = \{T \in \mathcal{L}(E, F) \mid \dim T(E) < \infty\}$  der Vektorraum aller beschränkten linearen endlichdimensionalen Operatoren von  $E$  nach  $F$ .

### 3.5.2 Kompakte Operatoren in Hilberträumen

Während auf Banachräumen die Kompaktheit die Vollstetigkeit impliziert, muss die Umkehrung nicht gelten. In Hilberträumen mit ihrer reicheren mathematischen Struktur sind diese beiden Eigenschaften jedoch zueinander und auch zu einigen anderen Aussagen äquivalent.

**Satz 3.5.14.** *Sei  $E$  ein Hilbertraum,  $T \in \mathcal{L}(E)$ . Dann sind folgende Aussagen äquivalent:*

- (i)  $T$  ist vollstetig,
- (ii) Wenn  $x_n \rightharpoonup x$  und  $y_n \rightarrow y$  gilt, dann konvergiert  $\langle Tx_n, y_n \rangle$  gegen  $\langle Tx, y \rangle$ ,
- (iii)  $T$  ist kompakt,
- (iv)  $T^*$  ist kompakt.

*Beweis.* (iii) $\Rightarrow$ (i): Klar nach Satz 3.5.8.

(i) $\Rightarrow$ (iii): Sei  $(x_n)_{n \in \mathbb{N}}$  eine beschränkte Folge. Nach Satz 3.3.4 besitzt die Folge eine schwach konvergente Teilfolge  $(x_{n_k})_{k \in \mathbb{N}}$ , sagen wir  $x_{n_k} \rightharpoonup x$ . Da  $T$  vollstetig ist, gilt  $T(x_{n_k}) \rightarrow T(x)$  und damit ist  $T(\{x_n \mid n \in$

$\mathbb{N}$ ) relativ kompakt. Da man die beschränkte Folge beliebig wählen konnte, heißt das, dass  $T$  kompakt ist.

(iii) $\Rightarrow$ (ii) Es gelte  $x_n \rightharpoonup x$  und  $y_n \rightharpoonup y$ . Dann gilt weiter

$$\langle Tx_n, y_n \rangle - \langle Tx, y \rangle = \langle Tx_n - Tx, y_n \rangle + \langle Tx, y_n - y \rangle.$$

Also haben wir

$$|\langle Tx_n, y_n \rangle - \langle Tx, y \rangle| \leq \|Tx_n - Tx\| \|y_n\| + |\langle Tx, y_n - y \rangle|.$$

Angenommen, (ii) wäre falsch. Dann gibt es Teilfolgen  $(x_{n_k})_{k \in \mathbb{N}}$  und  $(y_{n_k})_{k \in \mathbb{N}}$  und ein  $\varepsilon > 0$  mit  $|\langle Tx_{n_k}, y_{n_k} \rangle - \langle Tx, y \rangle| > \varepsilon$  für alle  $n_k \in \mathbb{N}$ . Da  $x_n$  schwach gegen  $x$  konvergiert, ist  $(x_n)_{n \in \mathbb{N}}$  beschränkt. Da außerdem  $T$  nach Voraussetzung kompakt ist, besitzt die Folge  $(Tx_{n_k})_{k \in \mathbb{N}}$  eine konvergente Teilfolge  $(Tx_{n_{k_r}})_{r \in \mathbb{N}}$ , für die etwa  $Tx_{n_{k_r}} \rightarrow \tilde{y}$  gelte. Weil auch  $x_{n_{k_r}} \rightharpoonup x$  gilt, folgt  $Tx_{n_{k_r}} \rightarrow Tx$ , denn  $T$  ist ebenfalls vollstetig. Also muss  $\tilde{y} = Tx$  sein. Es folgt damit

$$|\langle Tx_{n_{k_r}}, y_{n_{k_r}} \rangle - \langle Tx, y \rangle| \leq \underbrace{\|T(x_{n_{k_r}}) - T(x)\|}_{\rightarrow 0} \underbrace{\|y_{n_{k_r}}\|}_{\leq c} + \underbrace{|\langle Tx, y - y_{n_{k_r}} \rangle|}_{\rightarrow 0}.$$

Für  $r \rightarrow \infty$  folgt daraus, dass die rechte Seite Null ist. Dies ist aber ein Widerspruch zu der Annahme, dass die linke Seite größer als  $\varepsilon$  sein muss.

(ii) $\Rightarrow$ (i): Sei  $y_n = Tx_n - Tx$ . Weil  $x_n \rightharpoonup x$  gilt, folgt  $Tx_n \rightharpoonup Tx$ , also  $y_n \rightharpoonup 0$ . Nach Voraussetzung gilt damit  $\langle Tx_n, y_n \rangle \rightarrow \langle Tx, 0 \rangle = 0$ . Es gilt durch Einsetzen von  $y_n$

$$\|Tx_n - Tx\|^2 = \langle Tx_n - Tx, Tx_n - Tx \rangle = \langle Tx_n, Tx_n - Tx \rangle - \langle Tx, Tx_n - Tx \rangle$$

und daraus folgt  $\|Tx_n - Tx\| \rightarrow 0$ , das heißt  $Tx_n \rightarrow Tx$ . Also ist  $T$  vollstetig.

(iv) $\Rightarrow$ (iii): Sei  $T^*$  kompakt. Sei  $(x_n)_{n \in \mathbb{N}}$  eine beschränkte Folge in  $E$ . Da die kompakten Operatoren ein Ideal in  $\mathcal{L}(E)$  bilden, ist  $T^*T$  ebenfalls kompakt. Deshalb besitzt  $(T^*Tx_n)_{n \in \mathbb{N}}$  eine konvergente Teilfolge  $(T^*Tx_{n_k})_{k \in \mathbb{N}}$ , sagen wir  $T^*Tx_{n_k} \rightarrow y$ . Es gilt für  $n, l \in \mathbb{N}$  nach der Cauchy-Schwarz-Ungleichung

$$\begin{aligned} \|Tx_{n_k} - Tx_{n_l}\|^2 &= \langle T(x_{n_k} - x_{n_l}), T(x_{n_k} - x_{n_l}) \rangle = \langle T^*T(x_{n_k} - x_{n_l}), x_{n_k} - x_{n_l} \rangle \\ &\leq \|T^*Tx_{n_k} - T^*Tx_{n_l}\| \|x_{n_k} - x_{n_l}\|. \end{aligned}$$

Da die Folge  $(x_n)_{n \in \mathbb{N}}$  beschränkt ist, folgt hieraus, dass  $(Tx_{n_k})_{k \in \mathbb{N}}$  eine Cauchyfolge ist und damit in  $E$  konvergiert. Also ist  $T$  kompakt.

(iii) $\Rightarrow$ (iv): Sei  $T$  kompakt. Nach der eben bewiesenen Implikation ist  $T^*$  kompakt, wenn man  $T = (T^*)^*$  beachtet.  $\square$

All diese Aussagen bedeuten eine Verbesserung der Eigenschaften. Aussage (ii) wäre immer richtig, wenn eine der Folgen  $(x_n)$  oder  $(y_n)$  stark konvergieren würde. Durch die zweimalige Abschwächung der Konvergenz gilt die Aussage jedoch nicht mehr immer. Wir haben hier nicht bewiesen, dass man aus beschränkten Folgen stets eine konvergente Teilfolge auswählen kann.

Wir werden nun in zwei kurzen Abschnitten Entwicklungen von kompakten Operatoren angeben. Diese sind Verallgemeinerungen von Entwicklungen nach Basen im endlichdimensionalen. Die Unterschiede in den verschiedenen Entwicklungen sind subtil und können in verschiedenen Situationen unterschiedlich nützlich sein.<sup>1</sup>

<sup>1</sup>Moderne Wissenschaft macht nicht das, was sie eigentlich sollte, denn sonst würde die Pharmaindustrie nicht viel mehr Forschung für Anti-aging-Cremes als für schlimme Krankheiten betreiben. Der Kunde zahlt eben besser. Wissenschaftler gehen also dahin, wo es etwas zu holen gibt. Was hat das mit unserem Thema zu tun? Auch hier wird man sich auf bessere Hilberträume zurückziehen, wo es viel mehr Technik gibt, statt dass man Dinge auf schlechten Banachräumen untersucht.



### 3.5.3 Der Hilbert-Schmidtsche Entwicklungssatz für kompakte selbstadjungierte Operatoren

**Satz 3.5.15.**  *$E$  sei ein unendlichdimensionaler, separabler Hilbertraum.  $T$  sei ein selbstadjungierter, kompakter Operator auf  $E$ . Dann existiert eine reelle Nullfolge  $(\lambda_n)_{n \in \mathbb{N}}$  und ein normiertes Orthogonalsystem  $\{e_n \mid n \in \mathbb{N}\}$  in  $E$  mit*

$$Tx = \sum_{n=0}^{\infty} \lambda_n \langle x, e_n \rangle e_n$$

für alle  $x \in E$ . Insbesondere gilt  $Te_k = \lambda_k e_k$  für  $k \in \mathbb{N}$ .

*Beweis.* Sei  $\lambda \neq 0$  ein Eigenwert von  $T$  und  $E_\lambda = \{x \in E \mid Tx = \lambda x\}$  der zugehörige Eigenraum. Mit  $B_{E_\lambda}$  bezeichnen wir die Einheitskugel in  $E_\lambda$ . Offensichtlich gilt  $T(B_{E_\lambda}) = \lambda B_{E_\lambda}$ . Weil  $T$  nach Voraussetzung kompakt ist, ist  $\lambda B_{E_\lambda}$  eine relativ kompakte Menge, also ist auch  $B_{E_\lambda}$  relativ kompakt und damit muss nach Beispiel 3.5.2  $E_\lambda$  endlichdimensional sein.

Seien  $\lambda_1, \lambda_2$  Eigenwerte von  $T$  mit  $\lambda_1, \lambda_2 \neq 0$  und  $\lambda_1 \neq \lambda_2$ . Als Übungsaufgabe zeigt man, dass  $E_{\lambda_1} \perp E_{\lambda_2}$  gilt. Wir wählen nun Orthogonalbasen für  $E_{\lambda_n}$  für höchstens abzählbar unendlich viele Eigenwerte  $\lambda_n \neq 0$  von  $T$ . Sei  $\{e_n \mid n \in N\}$  die Menge aller dieser orthogonalen Vektoren, wobei  $N = \{1, \dots, N\}$  endlich oder  $N = \mathbb{N}$ , also abzählbar sei. Weiter sei  $H_1$  der von  $\{e_n \mid n \in N\}$  erzeugte lineare Teilraum von  $E$ . Es gilt für  $r \in \mathbb{N}$

$$T \left( \sum_{n=1}^r \alpha_n e_n \right) = \sum_{n=1}^r \lambda_n \alpha_n e_n \in H_1,$$

da  $\alpha_n e_n \in H_1$  ist. Hieraus folgt  $T(H_1) \subseteq H_1$ .

Im nächsten Schritt wollen wir zeigen, dass  $T(H_1^\perp) \subseteq H_1^\perp$  ist. Seien hierzu  $x \in H_1$  und  $y \in H_1^\perp$ . Wir benutzen die Selbstadjungiertheit von  $T$  und  $Tx \in H_1$  und erhalten

$$0 = \langle y, Tx \rangle = \langle Ty, x \rangle$$

für alle  $x \in H_1$ . Deshalb muss  $Ty \in H_1^\perp$  sein. Wir definieren nun einen Operator  $T_1$  durch  $T_1 y = Ty$  für alle  $y \in H_1^\perp$ . Damit ist dann  $T_1$  ein linearer Operator auf dem Hilbertraum  $H_1^\perp$ , und die Kompaktheit und Selbstadjungiertheit vererbt sich von  $T$  auf  $T_1$ .

Nun wollen wir zeigen, dass  $T_1$  auf  $H_1^\perp$ , also auf seinem gesamten Definitionsbereich, der Nulloperator ist. Wäre  $T_1 \neq 0$ , so wäre  $\|T_1\| \neq 0$  und nach Lemma 3.5.16 besitzt  $T_1$  dann einen Eigenwert  $\mu \neq 0$ . Sei  $T_1 z = \mu z$  für einen Eigenvektor  $z$ , für den  $\|z\| > 1$  gilt. O.B.d.A. ist  $\mu = \lambda_k$  für ein  $k \in \mathbb{N}$ , also gilt auch  $Tz = \mu z$  und damit ist  $z \in H_1$  nach Konstruktion des Raumes  $H_1$ . Andererseits gilt  $z \in H_1^\perp$  und damit  $\langle z, z \rangle = 0$ , also  $z = 0$ , was nicht sein kann. Also ist  $T_1$  tatsächlich der Nulloperator auf  $H_1^\perp$ . Sei  $\{f_k \mid k \in I\}$  mit einer Indexmenge  $I$  ein vollständiges normiertes Orthogonalsystem für  $H_1^\perp$ . Die Menge  $\{e_n, f_k \mid n \in \mathbb{N}, k \in I\}$  ist dann ein vollständiges normiertes Orthogonalsystem für den gesamten Raum  $E$ .

Sei nun  $x \in E$ . Es gilt die Darstellung

$$x = \sum_{n \in \mathbb{N}} \langle x, e_n \rangle e_n + \sum_{k \in I} \langle x, f_k \rangle f_k.$$

Damit ergibt sich

$$Tx = \sum_{n \in \mathbb{N}} \langle x, e_n \rangle Te_n + \sum_{k \in I} \langle x, f_k \rangle Tf_k = \sum_{n \in N} \lambda_n \langle x, e_n \rangle e_n,$$

denn es gilt  $Tf_k = T_1 f_k = 0$  und  $Te_n = \lambda_n e_n$ . Wenn  $N$  endlich ist, ergänzen wir nun  $\{e_n \mid n \in N\}$  durch Hinzufügen von abzählbar vielen  $f_k$  zu einem normierten Orthogonalsystem  $\{e_n \mid n \in \mathbb{N}\}$ . Weil  $T$  kompakt ist, ist  $(\lambda_n)_{n \in \mathbb{N}}$  eine Nullfolge, wie man als Übungsaufgabe zeigt, und aus der Selbstadjungiertheit von  $T$  folgt, dass alle  $\lambda_n$  reell sein müssen.  $\square$

Schmidt bewies diesen Satz in seiner Dissertation. Der entscheidende Schritt ist das folgende Lemma, das wir im vorangegangenen Beweis bereits benutzt haben.<sup>2</sup>

**Lemma 3.5.16.** Wenn  $T$  ein kompakter und selbstadjungierter Operator ist, dann ist  $\|T\|$  oder  $-\|T\|$  ein Eigenwert von  $T$ .

*Beweis.* Da  $T = T^*$  ist, gilt nach Satz 3.4.12, dass  $\|T\| = \sup_{\|x\|=1} |\langle Tx, x \rangle|$ . Deshalb existiert eine Folge  $(x_n)_{n \in \mathbb{N}}$  mit  $\|x_n\| = 1$  für alle  $n \in \mathbb{N}$ , für die  $\langle Tx_n, x_n \rangle \rightarrow \alpha$  gilt, wobei  $\alpha$  reell ist mit  $|\alpha| = \|T\|$ . Da  $T$  kompakt ist, hat die Folge  $Tx_n$  eine konvergente Teilfolge  $(Tx_{n_k})_{k \in \mathbb{N}}$ , für die etwa  $Tx_{n_k} \rightarrow y$  gelte. Wir haben dann

$$\begin{aligned} \|(T - \alpha)x_{n_k}\|^2 &= \langle (T - \alpha)x_{n_k}, (T - \alpha)x_{n_k} \rangle = \|Tx_{n_k}\|^2 + \alpha^2 \|x_{n_k}\|^2 - 2\alpha \langle Tx_{n_k}, x_{n_k} \rangle \\ &\leq \|T\|^2 + \alpha^2 - 2\alpha \langle Tx_{n_k}, x_{n_k} \rangle \end{aligned}$$

und es gilt

$$\|T\|^2 + \alpha^2 - 2\alpha \langle Tx_{n_k}, x_{n_k} \rangle \rightarrow \|T\|^2 + \alpha^2 - 2\alpha\alpha = 0.$$

O.B.d.A. sei  $T \neq 0$ , also  $\|T\| = |\alpha| \neq 0$ . Wir haben eben gezeigt, dass  $(T - \alpha)x_{n_k} \rightarrow 0$  gilt. Daraus folgt  $\alpha x_{n_k} = -(T - \alpha)x_{n_k} + Tx_{n_k} \rightarrow 0 + y$ . Also folgt  $x_{n_k} \rightarrow \frac{y}{\alpha}$ . Da  $T$  beschränkt ist, erhalten wir daraus  $Tx_{n_k} \rightarrow T\frac{y}{\alpha}$  und weiterhin  $(T - \alpha)x_{n_k} \rightarrow (T - \alpha)\frac{y}{\alpha} = 0$ . Also muss  $(T - \alpha)y = 0$  und damit  $Ty = \alpha y$  sein. Damit ist  $\|T\|$  oder  $-\|T\|$  ein Eigenwert zum Eigenvektor  $y$ , denn weil  $x_{n_k} \rightarrow \frac{y}{\alpha}$  und  $\|x_{n_k}\| = 1$  gilt, folgt  $\|\frac{y}{\alpha}\| = 1$ , also ist  $y \neq 0$ .  $\square$

### 3.5.4 Die kanonische Form (Schmidt-Darstellung) kompakter Operatoren im Hilbertraum

**Satz 3.5.17** (Schmidt-Darstellung). *E sei ein unendlichdimensionaler separabler Hilbertraum und  $T$  ein kompakter Operator auf  $E$ . Dann existieren eine monoton fallende, nichtnegative Zahlenfolge  $(s_n)_{n \in \mathbb{N}_0}$  und Orthonormalsysteme  $\{e_n \mid n \in \mathbb{N}_0\}$  und  $\{f_n \mid n \in \mathbb{N}_0\}$ , so dass für alle  $x \in E$  gilt:*

$$Tx = \sum_{n=0}^{\infty} s_n \langle x, e_n \rangle f_n.$$

*Insbesondere gilt  $Te_m = s_m f_m$ .*

*Beweis.* Da  $T$  kompakt und  $T^*$  beschränkt ist, ist  $T^*T$  kompakt, da die kompakten Operatoren ein Ideal in der Operatoralgebra bilden. Es gilt  $(T^*T)^* = T^*(T^*)^* = T^*T$ , also ist  $T^*T$  selbstadjungiert. Nach Satz 3.5.3, den wir auf den Operator  $T^*T$  anwenden, existiert eine reelle Nullfolge  $(\lambda_n)_{n \in \mathbb{N}_0}$  und ein normiertes Orthogonalsystem  $\{e_n \mid n \in \mathbb{N}_0\}$  mit  $T^*Tx = \sum_{n=0}^{\infty} \lambda_n \langle x, e_n \rangle e_n$  für alle  $x \in H$ . Wir haben die Gleichheit

$$\|Te_n\|^2 = \langle Te_n, Te_n \rangle = \langle T^*Te_n, e_n \rangle = \langle \lambda_n e_n, e_n \rangle = \lambda_n \langle e_n, e_n \rangle = \lambda_n,$$

also gilt  $\lambda_n \geq 0$  für alle  $n \in \mathbb{N}_0$ . O.B.d.A. ist  $(\lambda_n)_{n \in \mathbb{N}}$  monoton fallend, ansonsten erreicht man dies durch einfaches Umnummerieren. Wir definieren  $s_n := \sqrt{\lambda_n}$  und  $f_n := \frac{1}{\sqrt{\lambda_n}} Te_n$ , falls  $\lambda_n \neq 0$  ist. Sind  $\lambda_n \neq 0$  und  $\lambda_m \neq 0$ , so gilt

$$\langle f_n, f_m \rangle = \frac{1}{\sqrt{\lambda_n \lambda_m}} \langle Te_n, Te_m \rangle = \frac{1}{\sqrt{\lambda_n \lambda_m}} \langle T^*Te_n, e_m \rangle = \frac{\lambda_n}{\sqrt{\lambda_n \lambda_m}} \langle e_n, e_m \rangle = \delta_{n,m}.$$

<sup>2</sup>Man sagt, dass man bei der Forschung im Leistungszentrum sitzen muss, da die Probleme, die man in Publikationen liest, entweder zu dem Zeitpunkt bereits gelöst sind oder sie sind so schwer, dass man sie auch nicht mehr herausbekommt. Bei Riesz war das ganz anders, er war in Ungarn und hatte lediglich Korrespondenz per Post, er hat alles alleine herausbekommen. Man kann sich also auch in einem Zimmer einschließen und einfach am Schreibtisch alleine alle Probleme lösen, das geht auch. Merke aber: Nicht jeder ist ein Riesz! Alleine sollte man es also trotz alledem nicht versuchen.

Die Menge der Vektoren  $f_n$  mit  $\lambda_n \neq 0$  ist also ein normiertes Orthogonalsystem. Durch Hinzunahme von Elementen  $f_k/\|f_k\|$ , für die  $Tf_k = 0$  ist, ergänzen wir das System zu einem normierten Orthogonalsystem  $\{f_n \mid n \in \mathbb{N}_0\}$ . Wenn für ein  $y \in H$  die Bedingung  $y \perp e_n$  für alle  $n \in \mathbb{N}_0$  erfüllt ist, dann ist

$$\|Ty\|^2 = \langle Ty, Ty \rangle = \langle T^*Ty, y \rangle = \sum_{n=0}^{\infty} \lambda_n \langle y, e_n \rangle \langle e_n, y \rangle = 0.$$

Dann muss natürlich  $Ty = 0$  sein. Für ein beliebiges  $x \in H$  gilt zudem

$$x = x - \underbrace{\sum_{n=0}^{\infty} \langle x, e_n \rangle e_n}_{=:z} + \sum_{n=0}^{\infty} \langle x, e_n \rangle e_n.$$

Es gilt  $z \perp e_k$  für alle  $k \in \mathbb{N}_0$ , denn

$$\langle z, e_k \rangle = \langle x, e_k \rangle - \sum_{n=0}^{\infty} \langle x, e_n \rangle \underbrace{\langle e_n, e_k \rangle}_{=\delta_{nk}} = \langle x, e_k \rangle - \langle x, e_k \rangle = 0.$$

Daraus folgt nach obigen Betrachtungen also  $Tz = 0$ . Damit ist

$$Tx = Tz + T \left( \sum_{n=0}^{\infty} \langle x, e_n \rangle e_n \right) = \sum_{n=0}^{\infty} \langle x, e_n \rangle Te_n = \sum_{n=0}^{\infty} s_n \langle x, e_n \rangle f_n.$$

□

Mit Hilfe der Schmidt-Darstellung eines kompakten Operators können wir nun beweisen, dass sich jeder kompakte Operator durch eine Folge von endlichdimensionalen Operatoren approximieren lässt.

**Korollar 3.5.18.** Sei  $T$  kompakt. Dann existiert eine Folge  $(T_n)_{n \in \mathbb{N}}$  von endlichdimensionalen Operatoren  $T_n \in \mathcal{F}(E) = \{A \in \mathcal{L}(H) \mid \dim A(E) < \infty\}$  mit  $\lim_{n \rightarrow \infty} \|T - T_n\| = 0$ .

*Beweis.* Sei  $Tx = \sum_{k=0}^{\infty} s_k \langle x, e_k \rangle f_k$  die Darstellung von  $T$  nach Satz 3.5.17. Für  $x \in H$  sei

$$T_n x := \sum_{k=0}^n s_k \langle x, e_k \rangle f_k.$$

$T_n$  ist offensichtlich stetig und es gilt  $T_n(E) \subseteq \text{span}\{f_0, f_1, \dots, f_n\}$ , also ist  $T_n \in \mathcal{F}(E)$ . Es gilt

$$\|(T - T_n)x\|^2 = \left\| \sum_{k=n+1}^{\infty} s_k \langle x, e_k \rangle f_k \right\|^2 = \sum_{k=n+1}^{\infty} \|s_k \langle x, e_k \rangle f_k\|^2 = \sum_{k=n+1}^{\infty} s_k^2 |\langle x, e_k \rangle|^2 \leq s_{n+1}^2 \|x\|^2,$$

wobei wir im letzten Schritt sowohl die Besselsche Ungleichung benutzt haben, als auch, dass  $s_n$  monoton fallend ist. Im Schritt davor benutzten wir den Satz des Pythagoras, der sich ohne weiteres auf unendliche Summen erweitern lässt. Es folgt also  $\|(T - T_n)x\| \leq s_{n+1} \|x\|$ , das heißt es ist  $\|T_n - T\| \leq s_{n+1}$  und damit  $\lim_{n \rightarrow \infty} \|T - T_n\| = 0$ . □

Bemerkt sei noch, dass die Behauptung für Banachräume im Allgemeinen nicht gilt, wie das Beispiel eines Banachraumes ohne Approximationseigenschaft von Enflo aus dem Jahr 1973 zeigt.

### 3.5.5 Singuläre Zahlen und Schattenklassen

**Definition 3.5.19.** Sei  $T$  ein kompakter Operator auf dem unendlichdimensionalen, separablen Hilbertraum  $H$ . Die Zahlen  $s_n$  aus Satz 3.5.17 heißen die *singulären Zahlen* von  $T$ .

**Satz 3.5.20.** Für die singulären Zahlen  $s_n$  gilt

$$s_n = \inf \{ \|T - S\| \mid S \in \mathcal{F}(H), \dim S(H) \leq n \}.$$

Aus diesem Satz folgt, dass die Folge  $(s_n)_{n \in \mathbb{N}_0}$  durch  $T$  eindeutig bestimmt ist. Wir schreiben zwecks Übersichtlichkeit  $s_n = s_n(T)$ .

**Bemerkung 3.5.21.** Allgemein gelten für die singulären Zahlen folgende Eigenschaften.

- (i)  $(s_n(T))_{n \in \mathbb{N}_0}$  ist eine monoton fallende nichtnegative Nullfolge.
- (ii) Es ist stets  $s_0(T) = \|T\|$ . Weiter ist  $s_n(T) = 0$ , falls  $\dim T(H) \leq n$  ist.
- (iii)  $s_n(ATB) \leq \|A\| \|B\| s_n(T)$  für alle  $A, B \in \mathcal{L}$ .
- (iv)  $s_n(T) = s_n(T^*)$  und  $s_n(\lambda T) = |\lambda| s_n(T)$  für  $\lambda \in \mathbb{C}$  mit  $\lambda \neq 0$ .
- (v)  $s_{n+m}(T + S) \leq s_n(T) + s_m(S)$ .
- (vi)  $s_{2n}(TS) \leq s_n(T) s_n(S)$ .

**Definition 3.5.22** (Schattenklasse). Die Menge aller Operatoren  $T \in \mathcal{K}(H)$ , deren singuläre Zahlen  $(s_n(T))_{n \in \mathbb{N}_0} \in \ell_p(\mathbb{N}_0)$  erfüllen, heißt *p-te Schattenklasse* auf  $H$ , in Zeichen  $\mathcal{S}_p(H)$ .

**Satz 3.5.23.** Für  $T \in \mathcal{S}_p(H)$ ,  $p \in [1, \infty)$  sei

$$\nu_p(T) = \left( \sum_{n=0}^{\infty} s_n(T)^p \right)^{1/p}$$

die Hilbert-Schmidt-Norm von  $T$ . Dann ist  $(\mathcal{S}_p(H), \nu_p(\cdot))$  ein Banachraum.

Die Operatoren aus  $\mathcal{S}_2(H)$  heißen *Hilbert-Schmidt-Operatoren*. Die Operatoren aus  $\mathcal{S}_1(H)$  heißen *nukleare Operatoren* oder *Operatoren der Spurklasse*. Aus Bemerkung 3.5.21 (iii) folgt, dass  $ATB \in \mathcal{S}_p(H)$  ist, wenn  $T \in \mathcal{S}_p(H)$  und  $A, B \in \mathcal{L}(H)$  sind.

**Satz 3.5.24.** Für  $T \in \mathcal{L}(H)$  sind die folgenden Aussagen äquivalent:

- (i)  $T \in \mathcal{S}_2(H)$ ,
- (ii) Es gilt  $\sum_{n=0}^{\infty} \|Te_n\|^2 < \infty$  für jedes vollständige normierte Orthogonalsystem  $\{e_n \mid n \in \mathbb{N}_0\}$ ,
- (iii) Es gilt  $\sum_{n=0}^{\infty} \|Te_n\|^2 < \infty$  für ein vollständiges normiertes Orthogonalsystem  $\{e_n \mid n \in \mathbb{N}_0\}$ .

### 3.5.6 Integraloperatoren mit quadratisch integrierbarem Kern

Sei  $M$  eine offene Menge im  $\mathbb{R}^d$ .  $K(t, s)$  sei eine Funktion aus  $L^2(M \times M, \lambda)$ , das heißt

$$\int_M \int_M |K(t, s)|^2 \lambda(dt, ds) < \infty.$$

Sei  $(T_K f)(t) = \int_M K(t, s) f(s) \lambda(ds)$  für  $f \in E = L^2(M, \lambda)$ . Beispielsweise könnte  $f = \frac{H}{|t-s|^\alpha}$  sein, wobei  $\alpha$  so gewählt werden muss, dass  $f$  quadratisch integrierbar ist.<sup>3</sup> Aus Beispiel 2.2.8 wissen wir, dass  $T_K \in \mathcal{L}(E)$  gilt, und es ist

$$\|T_K\| \leq \int_M \int_M |K(t, s)|^2 \lambda(dt, ds). \quad (3.5.1)$$

Damit ergibt sich folgendes Resultat.

**Satz 3.5.25.**

(i)  $T_K$  ist kompakt.

(ii)  $T_K$  ist ein Hilbert-Schmidt-Operator und  $\nu_2(T_K) = \left( \int_M \int_M |K(t, s)|^2 \lambda(dt, ds) \right)^{1/2}$ .

*Beweis.*  $E = L^2(M, \lambda)$  ist separabel und unendlichdimensional. Sei  $\{\varphi_n(t) \mid n \in \mathbb{N}\}$  ein vollständiges Orthonormalsystem für  $E$ . Dann ist  $\{\varphi_n(t)\varphi_m(s) \mid n, m \in \mathbb{N}\}$  ein vollständiges Orthonormalsystem für  $L^2(M \times M, \lambda)$ . Da der Kern  $K \in L^2(M \times M, \lambda)$  ist, kann man ihn fourierentwickeln:

$$K = \sum_{n, m=1}^{\infty} a_{nm} \varphi_n(t) \overline{\varphi_m(s)}.$$

Sei

$$K_N = \sum_{n, m=1}^N a_{nm} \varphi_n(t) \overline{\varphi_m(s)}.$$

Dann konvergiert  $K_N$  für  $N \rightarrow \infty$  in der  $L^2$ -Norm gegen  $K$ . Betrachte die Wirkung von  $K_N$  auf ein Element des Hilbertraumes:

$$\begin{aligned} (T_{K_N} f)(t) &= \int_M \left( \sum_{n, m=1}^N a_{nm} \varphi_n(t) \overline{\varphi_m(s)} \right) f(s) \lambda(ds) = \sum_{n, m=1}^N a_{nm} \varphi_n(t) \int_M \overline{\varphi_m(s)} f(s) \lambda(ds) \\ &= \sum_{m, n=1}^N a_{nm} \langle f, \varphi_m \rangle \varphi_n(t). \end{aligned}$$

Damit ist  $T_{K_N}$  ein stetiger linearer Operator, dessen Wertebereich endlichdimensional ist, denn er wird lediglich von Linearkombinationen von  $\varphi_1, \dots, \varphi_N$  aufgespannt. Damit ist  $T_{K_N} \in \mathcal{F}(E) \subseteq \mathcal{K}(E)$ , also  $T_{K_N}$  kompakt. Es gilt weiter

$$\|T_K - T_{K_N}\| = \|T_{K-K_N}\| \leq \int_M \int_M |K(t, s) - K_N(t, s)|^2 \lambda(dt, ds) \rightarrow 0,$$

wobei wir Gleichung (3.5.1) auf  $K - K_N$  angewendet haben.  $K - K_N$  ist der Abschnitt von  $N$  bis  $\infty$  der Fourierreihe von  $K$ . Daher konvergiert der obige Ausdruck gegen 0 für  $N \rightarrow \infty$ , also

$$\lim_{N \rightarrow \infty} \|T_K - T_{K_N}\| = 0.$$

Da  $T_{K_N} \in \mathcal{K}(E)$ , ist  $T \in \mathcal{K}(E)$ . Wir müssen noch die Hilbert-Schmidt-Norm ausrechnen. Dafür betrachten wir zunächst die Wirkung von  $T_K$  auf ein Element  $\varphi_j$  des Orthonormalsystems.

$$(T_K \varphi_j)(t) = \int_M \left( \sum_{n, m=1}^{\infty} a_{nm} \varphi_n(t) \overline{\varphi_m(s)} \right) \varphi_j(s) \lambda(ds) = \sum_{n, m=1}^{\infty} a_{nm} \varphi_n(t) \langle \varphi_j, \varphi_m \rangle = \sum_{n=1}^{\infty} a_{nj} \varphi_n(t).$$

<sup>3</sup>In der Physik verhält es sich bei allen relevanten Fällen so, dass  $\alpha$  gerade den Wert hat, dass die Funktion nicht mehr in  $L^2$  liegt. Dass eine physikalische nie in die mathematische Theorie hineinpasst, scheint ein allgemeines Theorem zu sein. Es gibt eine sehr schöne mathematische Streutheorie, die nicht einmal für das einfachste sinnvolle physikalische Potential, das Coulombpotential, funktioniert.

Da  $(\varphi_n)$  vollständig in  $E$  ist, folgt aus der Parsevalschen Gleichung

$$\|T_K \varphi_j\|^2 = \sum_{n=1}^{\infty} |a_{nj}|^2,$$

und damit gilt

$$\sum_{j=1}^{\infty} \|T_K \varphi_j\|^2 = \sum_{j=1}^{\infty} \sum_{n=1}^{\infty} |a_{nj}|^2 = \int_M \int_M |K(t, s)|^2 \lambda(dt, ds),$$

weil  $(\varphi_n(t) \overline{\varphi_m(s)})$  vollständig in  $L^2(M \times M, \lambda)$  war.  $\square$

Über Hilbert-Schmidt-Operatoren kann man eine ganze Menge aussagen und die Resultate sind wesentlich einfacher zu handhaben als beispielsweise die für nukleare Operatoren. Deshalb tauchen die Operatoren aus  $\mathcal{S}_2$  häufig in der Literatur auf.<sup>4</sup>

---

<sup>4</sup>Ein Anwärter auf den Dokortitel hatte bei seiner Dissertationsverteidigung den Begriff des Hilbert-Schmidt-Operators benutzt, konnte aber auf die anschließend gestellte Frage, was diese Operatoren denn seien, keine Auskunft geben. Es folgte eine heftige Diskussion, ob man den Dokortitel zuerkennen sollte oder nicht. Man sollte also lieber keine mathematischen Vokabeln verwenden, von denen man nicht weiß, was sie bedeuten. Hätte der Doktorand einfach das mathematische Kalkül ohne den Namen verwendet, wäre es nicht zu dieser Frage gekommen oder er hätte zumindest sagen können, noch nie etwas von Hilbert-Schmidt-Operatoren gehört zu haben (was nach Verwendung des Begriffs schwierig wird).

## Kapitel 4

# Das Spektraltheorem für beschränkte selbstadjungierte Operatoren

Dieses Kapitel wird technisch etwas schwieriger.<sup>1</sup> Um eine Motivation zu bekommen, betrachten wir  $E = \mathbb{C}^n$  mit dem üblichen Skalarprodukt. Sei weiter  $A$  eine hermitesche  $n \times n$ -Matrix, das heißt  $A = (a_{ij})$  und  $a_{ij} = \overline{a_{ji}}$ . Wenn wir  $Tx = Ax$  setzen, dann ist  $T$  ein selbstadjungierter Operator auf  $\mathbb{C}^n$ . Es existiert dann eine Orthonormalbasis  $\{e_1, \dots, e_n\}$  aus Eigenvektoren von  $T$  (Hauptachsentransformation hermitescher Matrizen). Für diese gelte  $Te_j = \lambda_j e_j$ . Für  $x = \sum_{j=1}^n \langle x, e_j \rangle e_j$  haben wir dann

$$Tx = \sum_{j=1}^n \langle x, e_j \rangle Te_j = \sum_{j=1}^n \lambda_j \langle x, e_j \rangle e_j,$$

was die Hilbert-Schmidt-Darstellung im Endlichen ist.  $P_j$  sei der Projektionsoperator auf  $\mathbb{C} \cdot e_j$ , also  $P_j x = \langle x, e_j \rangle e_j$ . Wir haben dann  $Tx = \sum_{j=1}^n \lambda_j P_j x$  für alle  $x \in E$ . Damit haben wir einen Operator  $T$ , der sich schreiben lässt als

$$T = \sum_{j=1}^n \lambda_j P_j, \tag{4.0.1}$$

wobei  $P_1, \dots, P_n$  Projektionsoperatoren sind, die die Eigenschaften

$$P_i P_j = 0, \quad i \neq j \tag{4.0.2}$$

und

$$\sum_{j=1}^n P_j = I \tag{4.0.3}$$

erfüllen. Wir wollen die Darstellung (4.0.1) mit den Eigenschaften (4.0.2) und (4.0.3) auf beschränkte, selbstadjungierte Operatoren im Hilbertraum verallgemeinern. Dabei wird man (4.0.1) nicht wirklich in dieser Form erwarten können, da die Eigenwerte eines Operators auch einfach 0 sein können, obwohl der Operator nicht der Nulloperator ist (siehe Volterrasche Integraloperatoren). Weiterhin könnte das

---

<sup>1</sup>Aber: Wenn man etwa liest, dass viele Leute Optionen und Derivate nicht verstehen, und dass diese Konzepte für normale Menschen nicht mehr durchschaubar seien, dann fragt man sich, was überhaupt los ist. Die Relativitätstheorie und Quantenmechanik und das Spektraltheorem sind wesentlich schwieriger zu begreifen. Wenn Leute Derivate nicht verstehen, dann liegt das nicht an der Schwierigkeit des Begriffs, sondern an der allgemeinen Unfähigkeit allerorten. Das Konzept der verschiedenen Finanzpakete ist ein Klacks gegen das, was wir hier machen.

Spektrum überabzählbar sein, womit man auch Probleme bekäme. Eher lässt sich die Darstellung von  $T$  in der Form

$$T = \int_a^b \lambda dE(\lambda) \quad (4.0.4)$$

erwarten, wobei  $E(\lambda)$  eine Spektralschar von Projektionsoperatoren sein soll. Wir werden diese intuitiv vielversprechende Formel jedoch formal definieren und erklären müssen. Zuerst machen wir uns aber noch klar, welche Vorteile wir von einem solchen Kalkül hätten.

Zunächst handelt es sich um ein weiteres Strukturtheorem für beschränkte, selbstadjungierte Operatoren. Es ist möglich, kompliziertere Operatoren  $T$  durch ein Integral über einfachere Operatoren, nämlich Projektionsoperatoren, darzustellen. Damit erhält man möglicherweise mehr Informationen über das Verhalten von  $T$  oder kann stärkere mathematische Techniken anwenden. Der zweite Grund für die Definition von Spektralscharen ist, dass damit Funktionen von  $T$  möglich werden, indem man für eine Funktion  $f$  definiert

$$f(T) = \int f(\lambda) dE(\lambda).$$

Diese Funktionen besitzen den sogenannten Funktionalkalkül, man kann auf sie also die gewöhnlichen mathematischen Operationen wie  $+$ ,  $-$  und  $\cdot$  anwenden. Funktionen von Operatoren sind außerordentlich wichtig. Beispiele sind die Funktion  $e^{itH}$  für einen Operator  $H$ , die als Zeitentwicklung des Systems in der Physik auftaucht, sowie die Resolvente  $(T - \lambda I)^{-1}$ , die wir bereits in Abschnitt 3.4 untersucht haben. Meistens kann man die Werte dieser Funktion natürlich nicht explizit ausrechnen.

Damit kommen wir nun zur Definition der Spektralschar, dem Begriff, der die Definition von (4.0.4) ermöglicht.

## 4.1 Spektralscharen

Sei  $H$  ein Hilbertraum.

**Definition 4.1.1.** Eine *Spektralschar* auf  $H$  ist eine Familie  $(E(\lambda) \mid \lambda \in \mathbb{R})$  von Projektionsoperatoren auf  $H$  mit folgenden Eigenschaften:

- (i)  $E(\lambda_1) \leq E(\lambda_2)$  für  $\lambda_1 \leq \lambda_2$  (Monotonie).
- (ii)  $\lim_{\lambda \downarrow \lambda_0} E(\lambda)x = E(\lambda_0)x$  für jedes  $\lambda_0 \in \mathbb{R}$  und  $x \in H$  (Rechtsstetigkeit in der starken Konvergenz).
- (iii)  $\lim_{\lambda \rightarrow -\infty} E(\lambda)x = 0$  und  $\lim_{\lambda \rightarrow +\infty} E(\lambda)x = x$ .

Diese Definition ist ähnlich wie die der Stieltjes-Integrale aus der Analysis und die der Verteilungsfunktionen aus der Wahrscheinlichkeitstheorie.

**Definition 4.1.2.** Man sagt, eine Spektralschar  $(E(\lambda) \mid \lambda \in \mathbb{R})$  hat einen *beschränkten Träger*, wenn es  $a, b \in \mathbb{R}$  gibt mit  $E(a) = 0$  und  $E(b) = I$ .

**Bemerkung 4.1.3.**

- (i) Die ältere Terminologie für Spektralscharen ist *Zerlegung der Eins*. Wir werden später sehen, warum diese Bezeichnung sinnvoll ist.
- (ii) Das Axiom (i) in Definition 4.1.1 bedeutet  $\langle E(\lambda_1)x, x \rangle \leq \langle E(\lambda_2)x, x \rangle$  für alle  $x \in H$ , was äquivalent zu  $E(\lambda_1)H \subseteq E(\lambda_2)H$  ist.
- (iii) Wenn  $E(a) = 0$  ist, so folgt aus 4.1.1 (i), dass  $E(\lambda) = 0$  ist für alle  $\lambda \leq a$ . Ebenfalls folgt aus  $E(b) = I$ , dass  $E(\lambda) = I$  ist für alle  $\lambda \geq b$ .



Wir wollen diese Sachverhalte an einem Beispiel verdeutlichen.

**Beispiel 4.1.4.** Sei  $(\lambda_n)_{n \in \mathbb{N}}$  eine Folge von paarweise verschiedenen reellen Zahlen. Weiter sei  $(P_n)_{n \in \mathbb{N}}$  eine Folge von Projektionsoperatoren auf  $H$  mit  $P_n P_k = 0$  für  $n \neq k$ , für die  $\sum_{n=1}^{\infty} P_n x = x$  für alle  $x \in H$  gilt. Wir definieren  $E(\lambda)x = \sum_{\lambda_n \leq \lambda} P_n x$ , wobei die Summe leer ist und damit den Wert 0 hat, wenn keine  $\lambda_n \leq \lambda$  existieren. Dann ist  $(E(\lambda) \mid \lambda \in \mathbb{R})$  eine Spektralschar.

Lassen wir die Beschränkung fallen, dass die  $\lambda_n$  paarweise verschieden sein müssen, dann kann man sich das Beispiel in einer Art diskretem Spezialfall veranschaulichen. Sei  $\lambda_1 < \lambda_2 < \lambda_3$  und  $\lambda_n = \lambda_3$  für  $n > 3$ . Weiter seien  $P_1, P_2, P_3$  Projektionsoperatoren mit  $P_i P_j = 0$  für  $i \neq j$  und  $P_1 + P_2 + P_3 = I$ . Auch hier handelt es sich dann um eine Spektralschar, die man sich skizzenhaft so vorstellen kann (natürlich kann man in Wirklichkeit keine Projektionsoperatoren zeichnen):

*Beweis.* Wir zeigen, dass  $E(\lambda)$  ein Projektionsoperator ist. Es gilt zunächst

$$E(\lambda)^2 = \left( \sum_{\lambda_j \leq \lambda} P_j \right) \left( \sum_{\lambda_l \leq \lambda} P_l \right) = \sum_{\lambda_j, \lambda_l \leq \lambda} P_j P_l = \sum_{\lambda_j, \lambda_l \leq \lambda} P_j \delta_{jl} = \sum_{\lambda_j \leq \lambda} P_j = E(\lambda).$$

Weiterhin haben wir

$$E(\lambda)^* = \left( \sum_{\lambda_j \leq \lambda} P_j \right)^* = \sum_{\lambda_j \leq \lambda} P_j^* = E(\lambda).$$

Nach Satz 3.2.1 ist damit  $E(\lambda)$  ein Projektionsoperator.

Nun zeigen wir die drei Axiome der Definition einer Spektralschar, siehe Definition 4.1.1.

(i): Weil  $P_n \geq 0$  ist, ist  $E(\lambda)$  monoton.

(ii): Wir zeigen  $\lim_{\lambda \downarrow \lambda_0} E(\lambda)x = E(\lambda_0)x$  für  $x \in H$ . Aufgrund des Satzes des Pythagoras gilt für  $x = \sum_{n=1}^{\infty} P_n x$  die Gleichung  $\|x\|^2 = \sum_{k=1}^{\infty} \|P_k x\|^2$ , da die  $P_n x$  senkrecht aufeinander stehen. Sei  $\varepsilon > 0$ . Dann existiert ein  $N_\varepsilon$  mit  $\sum_{n \geq N_\varepsilon} \|P_n x\|^2 < \varepsilon$ . Wähle  $\lambda_\varepsilon > \lambda_0$  so, dass alle Zahlen  $\lambda_j$  mit  $j \leq N_\varepsilon$  nicht im Intervall  $(\lambda_0, \lambda_\varepsilon]$  liegen. Damit gilt nach dem Satz des Pythagoras für  $\lambda_\varepsilon \geq \lambda > \lambda_0$

$$\|(E(\lambda) - E(\lambda_0))x\|^2 = \left\| \sum_{\lambda_0 < \lambda_n \leq \lambda} P_n x \right\|^2 = \sum_{\lambda_0 < \lambda_n \leq \lambda} \|P_n x\|^2 \leq \sum_{n \geq N_\varepsilon} \|P_n x\|^2 < \varepsilon.$$

Das funktioniert nur, weil wir die Summe über  $\lambda_n \leq \lambda$  und nicht  $\lambda_n < \lambda$  bilden. Damit folgt also  $\lim_{\lambda \downarrow \lambda_0} E(\lambda)x = E(\lambda_0)x$ .

(iii): Analog zu (ii). □

**Beispiel 4.1.5.** Sei  $H = L^2([a, b])$  mit einem beliebigen, positiven Borelmaß auf  $[a, b]$ . Wir definieren  $(E(\lambda)f)(t) = e_\lambda(t)f(t)$  für  $f \in H$ , wobei

$$e_\lambda(t) = \mathbb{1}_{(-\infty, \lambda]}(t) = \begin{cases} 1, & t \leq \lambda \\ 0, & t > \lambda, \end{cases}$$

sei. Dabei ist  $\mathbb{1}_A$  die charakteristische Funktion einer Menge  $A$ . Dann ist  $(E(\lambda) \mid \lambda \in \mathbb{R})$  eine Spektralschar und wie man leicht sieht, ist  $E(\lambda) = 0$ , wenn  $\lambda < a$  und  $E(\lambda) = I$ , wenn  $\lambda > b$  ist, also hat  $(E(\lambda) \mid \lambda \in \mathbb{R})$  einen beschränkten Träger.

### 4.1.1 Eigenschaften einer Spektralschar

- (i) Sei zunächst  $\lambda_0 \in \mathbb{R}$  fest. Nach Axiom (i) wächst  $E(\lambda)$  mit  $\lambda$ , also gilt für eine monoton gegen  $\lambda_0$  wachsende Folge  $(\lambda_n)_{n \in \mathbb{N}}$ , dass  $(E(\lambda_n))_{n \in \mathbb{N}}$  eine monoton wachsende, beschränkte Folge selbstadjungierter Operatoren ist. Nach Satz 3.3.9 konvergiert diese Folge in der starken Konvergenz, also

existiert der Grenzwert

$$E(\uparrow \lambda_0) := \lim_{\lambda \uparrow \lambda_0} E(\lambda)x,$$

welcher wiederum ein selbstadjungierter Operator ist. Es gilt  $(E(\lambda_n))^2 = E(\lambda_n)$ , also folgt im Grenzübergang auch  $(E(\uparrow \lambda_0))^2 = E(\uparrow \lambda_0)$ , und damit ist  $E(\lambda_0 \uparrow)$  ein Projektionsoperator. Für  $\lambda_1 < \lambda_0 \leq \lambda_2$  gilt weiter  $E(\lambda_1) \leq E(\uparrow \lambda_0) \leq E(\lambda_2)$ .

Zur Anwendung von Spektralscharen auf ganze Intervalle benötigen wir nun nur noch den Begriff der Stetigkeit. Eine Spektralschar  $(E(\lambda) \mid \lambda \in \mathbb{R})$  heißt stetig in  $\lambda_0$ , wenn  $E(\uparrow \lambda_0) = E(\lambda_0)$  ist, wenn sie also nicht nur rechts- sondern auch linksseitig stetig ist. Hier gilt insbesondere  $E(\lambda_0) - E(\uparrow \lambda_0) \geq 0$  und damit ist diese Differenz ebenfalls wieder ein Projektionsoperator.

- (ii) Wir definieren für Intervalle  $I$  mit den Grenzen  $\lambda, \mu \in \mathbb{R}$ , wobei  $\lambda \leq \mu$  gelte, den Projektionsoperator  $E(I)$  durch

$$\begin{aligned} E([\lambda, \mu]) &:= E(\mu) - E(\lambda \uparrow) \\ E((\lambda, \mu]) &:= E(\mu) - E(\lambda) \\ E([\lambda, \mu)) &:= E(\mu \uparrow) - E(\lambda \uparrow) \\ E((\lambda, \mu)) &:= E(\mu \uparrow) - E(\lambda). \end{aligned}$$

$E(I)$  ist wieder ein Projektionsoperator, und wenn  $I_1$  und  $I_2$  zwei disjunkte Intervalle der obigen Form sind, so ist  $E(I_1)E(I_2) = 0$ .

### 4.1.2 Spektralmaße

In diesem Abschnitt sei  $H$  wieder ein Hilbertraum.  $\mathfrak{P}(H)$  sei die Menge aller Projektionsoperatoren auf  $H$  und  $\mathcal{B}(\mathbb{R})$  die Borel- $\sigma$ -Algebra auf  $\mathbb{R}$ . Das aus der Maßtheorie bekannte Lebesgue-Maß ist eine Abbildung  $\mathcal{B}(\mathbb{R}) \rightarrow \mathbb{R}$ . Ohne Schwierigkeiten können wir einer Borelmenge jedoch auch einen Projektionsoperator und keine reelle Zahl zuordnen. Dies führt uns auf die Definition eines Spektralmaßes.

**Definition 4.1.6** (Spektralmaß). Ein Spektralmaß auf  $H$  ist eine Abbildung  $E : \mathcal{B}(\mathbb{R}) \rightarrow \mathfrak{P}(H)$  mit folgenden Eigenschaften:

- (i)  $E(\mathbb{R}) = I$  und  $E(\emptyset) = 0$ .  
(ii) Wenn  $A_n \in \mathcal{B}(\mathbb{R})$  eine Folge von paarweise disjunkten Mengen ist, dann gilt für  $x \in H$ :

$$E\left(\bigcup_{n \in \mathbb{N}} A_n\right)x = \sum_{n \in \mathbb{N}} E(A_n)x.$$

Man beachte bei (ii), dass es sich bei  $E(A)$  für eine Borelmenge  $A$  um einen Projektionsoperator, also eine Abbildung  $H \rightarrow H$  handelt, und deshalb die elementweise Forderung der Gleichheit sinnvoll ist.

**Bemerkung 4.1.7.** Wenn  $E$  ein Spektralmaß und  $x \in H$  ist, so ist die Abbildung  $\mu_x : \mathcal{B}(\mathbb{R}) \rightarrow \mathbb{R}$ , definiert durch  $\mu_x(M) = \langle E(M)x, x \rangle$  für  $M \in \mathcal{B}(\mathbb{R})$ , ein positives Borelmaß auf  $\mathbb{R}$ .

Zwischen Spektralmaßen und Spektralscharen besteht eine enge Verbindung, die der folgende Satz charakterisiert.

**Satz 4.1.8.**

- (i) Wenn  $E$  ein Spektralmaß auf  $H$  ist, dann definiert  $E(\lambda) := E((-\infty, \lambda])$  für  $\lambda \in \mathbb{R}$  eine Spektralschar  $(E(\lambda) \mid \lambda \in \mathbb{R})$  auf  $H$ .

- (ii) Sei  $(E(\lambda) \mid \lambda \in \mathbb{R})$  eine Spektralschar auf  $H$ . Dann existiert ein eindeutig bestimmtes Spektralmaß  $E$  auf  $H$  mit  $E(\lambda) = E((-\infty, \lambda])$  für alle  $\lambda \in \mathbb{R}$ .

Der Beweis der ersten Aussage ist relativ leicht, für die zweite muss man jedoch etwas länger überlegen. Dies wollen wir an dieser Stelle nicht tun, man lese den Beweis etwa in einem Buch nach. stattdessen betrachten wir die Eigenschaften eines Spektralmaßes etwas genauer. Die folgenden Behauptungen ergeben sich direkt aus den entsprechenden Aussagen über Projektionsoperatoren im Abschnitt 3.2.

**Bemerkung 4.1.9.** Sei  $E$  ein Spektralmaß auf  $H$ .

- (i) Wenn  $A_1, A_2 \in \mathcal{B}(\mathbb{R})$  sind und  $A_1 \subseteq A_2$  ist, dann gilt auch  $E(A_1) \subseteq E(A_2)$  und damit

$$E(A_1)E(A_2) = E(A_2)E(A_1) = E(A_1).$$

- (ii) Wenn  $A_1, A_2 \in \mathcal{B}(\mathbb{R})$  gilt und  $A_1$  und  $A_2$  disjunkt sind, so ist  $E(A_1 \cup A_2) = E(A_1) + E(A_2)$  und damit  $E(A_1)E(A_2) = 0$ . Weiterhin gilt  $E(A_1)H \perp E(A_2)H$ , die Räume  $E(A_1)H$  und  $E(A_2)H$  sind also orthogonal.

- (iii) Wenn  $A_1, A_2 \in \mathcal{B}(\mathbb{R})$  sind, dann gilt  $E(A_1 \cap A_2) = E(A_1)E(A_2) = E(A_2)E(A_1)$ .

**Beispiel 4.1.10.**

- (a) Betrachte die Spektralschar  $E(\lambda) = \sum_{\lambda_j \leq \lambda} P_j$  aus Beispiel 4.1.4. Wir definieren für  $M \in \mathcal{B}(\mathbb{R})$  ein Spektralmaß  $E$  durch  $E(M) = \sum_{\lambda_j \in M} P_j$ . Man sieht hieraus leicht, dass  $E$  das zu  $(E(\lambda) \mid \lambda \in \mathbb{R})$  gehörende Spektralmaß ist.
- (b) Sei eine Spektralschar gegeben durch  $(E(\lambda)f)(t) = e_\lambda(t)f(t) = \mathbb{1}_{(-\infty, \lambda]}(t)f(t)$ . Wir definieren für  $M \in \mathcal{B}(\mathbb{R})$  die Abbildung  $(E(M)f)(t) = \mathbb{1}_M(t)f(t)$ . Damit ist  $E$  das zu der Spektralschar  $(E(\lambda) \mid \lambda \in \mathbb{R})$  gehörende Spektralmaß.

### 4.1.3 Integration bezüglich einer Spektralschar

Im folgenden sei  $(E(\lambda) \mid \lambda \in \mathbb{R})$  eine Spektralschar mit beschränktem Träger, das heißt es gibt  $a_0, b \in \mathbb{R}$  mit  $E(a_0) = 0$  und  $E(b) = I$ . Sei weiter  $a < a_0$  fest und  $f(\lambda)$  eine Funktion auf  $[a, b]$ . Wir wollen im folgenden das Integral  $\int_a^b f(\lambda) dE(\lambda)$  definieren, also bezüglich einer Spektralschar integrieren. Damit werden wir dem Ziel dieses Kapitels ein großes Stück näher kommen.

Sei zunächst  $f(\lambda)$  eine stetige Funktion und o.B.d.A. reellwertig. Wenn  $f$  nämlich komplexwertig ist, also  $f = f_1 + if_2$  mit zwei reellwertigen Funktionen  $f_1$  und  $f_2$ , dann definieren wir

$$\int_a^b f(\lambda) E(d\lambda) = \int_a^b f_1(\lambda) E(d\lambda) + i \int_a^b f_2(\lambda) E(d\lambda).^2$$

Die folgenden Überlegungen kennt man bereits aus der Definition des Riemann-Integrals. Sei  $\mathfrak{Z}$  eine Zerlegung  $a = t_0 < t_1 < \dots < t_n = b$  des Intervalls  $[a, b]$  mit der Feinheit  $\delta_{\mathfrak{Z}} = \max_{j=0, \dots, n-1} |t_{j+1} - t_j|$ . Für Zwischenwerte  $\mathfrak{z}_j \in [t_j, t_{j+1}]$  definieren wir die Zwischensumme

$$S(f, \mathfrak{Z}) := \sum_{j=0}^{n-1} f(\mathfrak{z}_j)(E(t_{j+1}) - E(t_j)).$$

Diese ist natürlich ein Projektionsoperator. Mit Hilfe des Satzes von Cantor<sup>3</sup> kann man zeigen, dass ein beschränkter Operator  $T_f$  auf  $H$  existiert, für den  $\lim_{\delta_{\mathfrak{Z}} \rightarrow 0} \|T_f - S(f, \mathfrak{Z})\| = 0$  bei beliebiger Wahl der Zwischenpunkte  $\mathfrak{z}_j$  gilt. Wir schreiben

$$T_f = \int_a^b f(\lambda) E(d\lambda),$$

<sup>2</sup>Das sollte man alles ja schon einmal in der Schule gesehen haben. Normalerweise wär die Schulbildung ja okay, wenn das alles so gut umgesetzt werden würde.

<sup>3</sup>Eine stetige Funktion auf einem kompakten Intervall ist dort gleichmäßig stetig.

womit das Integral also der Grenzwert von  $\lim_{\delta_3 \rightarrow 0} S(f, \mathfrak{Z})$  in der Operatornorm ist. Nachdem wir nun bestimmt haben, wie wir ein solches Integral im stetigen Fall von  $f$  verstehen wollen, ist klar, dass man dies ebenso für alle  $E$ -messbaren Funktionen  $f(\lambda)$  tun kann. Wir wollen präzisieren, was wir darunter verstehen, weil der Begriff sich nicht mit dem aus der Maßtheorie bekannten deckt.

**Definition 4.1.11.** Eine reellwertige Funktion  $f(\lambda)$  heißt *E-messbar*, wenn  $f$  bezüglich jedes Maßes  $\mu_x(\cdot) = \langle E(\cdot)x, x \rangle$  mit  $x \in H$  messbar ist.

Dies ist unter anderem erfüllt, wenn  $f(\lambda)$  eine Borelfunktion ist.<sup>4</sup> Um unser Integral etwas genauer untersuchen zu können, benötigen wir sogenannte Sesquilinearformen. Diese kann man als Verallgemeinerungen des Skalarproduktes auffassen. Zwischen beschränkten Operatoren und Sesquilinearformen gibt es weiterhin einige Verknüpfungen, die wir in einem Satz zusammentragen möchten.

**Definition 4.1.12** (Sesquilinearform). Eine *Sesquilinearform*  $B$  auf  $H$  ist eine Abbildung  $B : H \times H \rightarrow \mathbb{C}$ , so dass  $B$  in der ersten Variablen linear und in der zweiten antilinear (das heißt konjugiert linear) ist.  $B$  heißt *stetig* oder *beschränkt*, wenn es ein  $c > 0$  gibt, so dass  $|B(x, y)| \leq c \|x\| \|y\|$  für alle  $x, y \in H$  ist.

**Satz 4.1.13.**

- (i) Sei  $T \in \mathcal{L}(H)$  und  $B_T(x, y) := \langle Tx, y \rangle$  für  $x, y \in H$ . Dann ist  $B_T$  eine beschränkte Sesquilinearform.
- (ii) Sei  $B$  eine beschränkte Sesquilinearform auf  $H$ . Dann existiert ein eindeutig bestimmter, beschränkter Operator  $T \in \mathcal{L}(H)$  mit  $B(x, y) = \langle Tx, y \rangle$  für alle  $x, y \in H$ .
- (iii)  $B$  sei eine Sesquilinearform auf  $H$ . Weiterhin existiere ein  $c > 0$  mit  $|B(x, x)| \leq c \|x\|^2$  für alle  $x \in H$ . Dann gilt sogar  $|B(x, y)| \leq 4c \|x\| \|y\|$  für alle  $x, y \in H$ , das heißt  $B$  ist beschränkt.

*Beweis.* Die ersten beiden Aussagen beweisen wir hier nicht. Die dritte lässt sich mit der Polarisationsidentität beweisen. Es gilt nämlich

$$\begin{aligned} |4B(x, y)| &= |B(x+y, x+y) - B(x-y, x-y) + iB(x+iy, x+iy) - iB(x-iy, x-iy)| \\ &\leq c(\|x+y\|^2 + \|x-y\|^2 + \|x+iy\|^2 + \|x-iy\|^2). \end{aligned}$$

Wenn nun  $\|x\| \leq 1$  und  $\|y\| \leq 1$  ist, dann folgt  $|B(x, y)| \leq 4c$ . Für beliebige  $x, y \in H$  folgt die Ungleichung durch einen auch in diesem Skript schon oft angewendeten Trick, nämlich das Normieren der Vektoren.  $\square$

Es sei  $u(\lambda)$  eine durch die Konstante  $C$  beschränkte Borelfunktion auf  $[a, b]$ , also  $|u(\lambda)| \leq C$  für alle  $\lambda \in [a, b]$ . Durch die Polarisationsidentität erhalten wir für das komplexe Maß  $\lambda E(M)x, y$  auf der Borel- $\sigma$ -Algebra  $\mathcal{B}(\mathbb{R})$  die Gleichung

$$\begin{aligned} \langle E(M)x, y \rangle &= \frac{1}{4}(\langle E(M)(x+y), x+y \rangle - \langle E(M)(x-y), x-y \rangle \\ &\quad + i\langle E(M)(x+iy), x+iy \rangle - i\langle E(M)(x-iy), x-iy \rangle). \end{aligned}$$

In der Maßtheorie werden gewöhnlich nur positive Maße betrachtet, ein komplexes (signiertes) Maß ist aber leicht als Linearkombination positiver Maße zu definieren. In obiger Gleichung lässt sich das komplexe Maß durch die Polarisationsidentität ebenfalls gerade in die Summe aus vier positiven Maßen zerlegen. Dabei ist  $\langle E(\cdot)z, z \rangle$  stets positiv, da es sich bei  $E(\cdot)$  um ein Spektralmaß handelt. In der Maßtheorie integriert man über komplexe Maße, in dem man auch die Integrale in vier Teile zerlegt, also

$$\int d\mu = \int d\mu_1 - \int d\mu_2 + i \int d\mu_3 - i \int d\mu_4,$$

<sup>4</sup>Borelfunktionen sind alles, was Sie aufmalen können.

wenn  $\mu = \mu_1 - \mu_2 + i\mu_3 - i\mu_4$  ein komplexes Maß ist, das aus vier positiven Mäßen zusammengesetzt ist. Das bedeutet für unseren Fall konkret

$$\begin{aligned} \int_a^b u(\lambda) \langle E(d\lambda)x, y \rangle &:= \frac{1}{4} \int_a^b u(\lambda) \langle E(d\lambda)(x+y), x+y \rangle - \frac{1}{4} \int_a^b u(\lambda) \langle E(d\lambda)(x-y), x-y \rangle \\ &\quad + \frac{i}{4} \int_a^b u(\lambda) \langle E(d\lambda)(x+iy), x+iy \rangle - \frac{i}{4} \int_a^b u(\lambda) \langle E(d\lambda)(x-iy), x-iy \rangle. \end{aligned}$$

Wir definieren nun für  $x, y \in H$  eine Abbildung

$$B(x, y) := \int_a^b u(\lambda) \langle E(d\lambda)x, y \rangle.$$

Diese ist eine Sesquilinearform auf  $H$ . Aus komplexen Integralen kann man aufgrund der auftretenden negativen Glieder in obigen Gleichungen nur schwer Faktoren herauschätzen, aber nach Satz 4.1.13 (iii) reicht es, wenn wir lediglich  $B(x, x)$  untersuchen. Dafür gilt, da es sich bei  $\langle E(\lambda)x, x \rangle$  um ein positives Maß handelt,

$$|B(x, x)| = \left| \int_a^b u(\lambda) \langle E(d\lambda)x, x \rangle \right| \leq C \left| \int_a^b 1 \langle E(d\lambda)x, x \rangle \right| = C \|x\|^2,$$

denn wegen  $E(b) = I$  und  $E(a) = 0$  gilt  $\langle E(b)x, x \rangle - \langle E(a)x, x \rangle = \langle x, x \rangle = \|x\|^2$ . Dies reicht nach Satz 4.1.13 (i) für die Stetigkeit von  $B$  und nach Satz 4.1.13 (ii) existiert ein  $T_u \in \mathcal{L}(H)$ , so dass sich  $B(x, y)$  sowohl als Integral als auch durch  $T_u$  in der Form

$$B(x, y) = \int_a^b u(\lambda) \langle E(d\lambda)x, y \rangle = \langle T_u x, y \rangle$$

ausdrücken lässt. Symbolisch schreibt man deshalb auch

$$T_u = \int_a^b u(\lambda) E(d\lambda).$$

Um sich die Wirkung dieses Operators besser einzuprägen, können wir nach Definition auch

$$\langle T_u x, y \rangle = \left\langle \left( \int_a^b u(\lambda) E(d\lambda) \right) x, y \right\rangle = \int_a^b u(\lambda) \langle E(d\lambda)x, y \rangle$$

schreiben. Man kann also gewissermaßen das Skalarprodukt in das Integral hineinziehen. Natürlich sind das alles nur symbolische Schreibweisen, die uns durch die richtig gewählten Definitionen gestattet sind.

**Bemerkung 4.1.14.**

- (i) Haben wir  $a_0 < b$  und ist  $(E(\lambda) \mid \lambda \in \mathbb{R})$  eine Spektralschar mit  $E(b) = I$  und  $E(a) = 0$  für alle  $a < a_0$ , dann hängt das Integral überhaupt nicht von den Funktionswerten von  $u(\lambda)$  für  $\lambda < a_0$  ab. Dies führt dann zu Problemen, wenn auf  $a_0$  gerade eine „Sprungstelle“ der Projektionsoperatoren vorhanden ist, wenn die Spektralschar also dort unstetig von links ist. Beispielsweise könnte  $E(a_0) = E(b) = I$  sein, die Spektralschar also konstant verlaufen. Dann wäre der Wert des Integrals einfach Null. In diesem Falle möchten wir den Sprung bei  $a_0$  gerne noch im Integral berücksichtigen. Man schreibt dann

$$\int_a^b u(\lambda) E(d\lambda) = \int_{\uparrow a_0}^b u(\lambda) E(d\lambda),$$

wobei es auf den konkreten Wert von  $a$  nicht ankommt, sofern  $a < a_0$  ist. Die Schreibweise  $\uparrow a_0$  steht dabei für den linksseitig gebildeten Grenzwert des Integrals.

(ii) Sei nun  $u(\lambda) = \mathbb{1}_{(-\infty, \mu]}(\lambda)$  eine charakteristische Funktion für einen Wert  $\mu \in \mathbb{R}$ . Wir haben dann

$$\left\langle \left( \int_{\uparrow a_0}^b u(\lambda) E(d\lambda) \right) x, y \right\rangle = \int_{\uparrow a_0}^b \mathbb{1}_{(-\infty, \mu]}(\lambda) \langle E(d\lambda)x, y \rangle = \langle E(\mu)x, y \rangle.$$

Bis jetzt hatten wir Spektralscharen vorgegeben und daraus Operatoren konstruiert. Was aber das eigentliche Ziel dieses Kapitels und auch der praxisrelevante Teil des Problems ist, ist die Konstruktion eines Operators aus einer Spektralschar. Wir wollen also den Rest des Kapitels den umgekehrten Weg gehen und einen selbstadjungierten Operator  $T$  vorgeben. Dann wollen wir eine Spektralschar finden, so dass  $T = \int \lambda E(d\lambda)$  ist. Dass dies stets möglich ist, wird der Inhalt des Spektraltheorems sein.

Man beachte, dass es uns bis jetzt noch nicht möglich ist, Funktionen  $f(T)$  von Operatoren sinnvoll zu definieren, wie wir es in der Einführung dieses Kapitels uns kurz überlegt hatten. Dort hatten wir gemutmaßt, dass

$$f(T) = \int f(\lambda) E(d\lambda)$$

eine sinnvolle Definition wäre. Die rechte Seite dieses Termes verstehen wir inzwischen, sofern eine Spektralschar gegeben ist, und wir wissen auch, dass das Integral für  $E$ -messbare Funktionen  $f$  gegen einen Operator aus  $\mathcal{L}(H)$  konvergiert. Wie wir allerdings die linke Seite verstehen dürfen, ist uns noch nicht klar. Eine formale Definition genügt unseren Ansprüchen nicht, da damit überhaupt nichts über den Zusammenhang zwischen dem Operator und der Funktion ausgesagt ist. In diesem Kontext ist ebenfalls noch nicht geklärt, welche Spektralschar man überhaupt verwenden muss, damit eine sinnvolle Relation zwischen  $f(T)$  und  $T$  entsteht.

Betrachten wir eine Schwierigkeit etwas detaillierter. Wir könnten zunächst mit unseren bisherigen Ergebnissen für jedes  $\mu \in \mathbb{R}$  den Ausdruck

$$E(\mu) = \int \mathbb{1}_{(-\infty, \mu]}(\lambda) E(d\lambda)$$

fordern, wenn die Spektralschar  $(E(\lambda) \mid \lambda \in \mathbb{R})$  unseren Ansprüchen genügen soll. Dies ist nur dann im Sinne unserer Überlegungen, wenn  $E(\mu) = \mathbb{1}_{(-\infty, \mu]}(T)$  gelten würde. Was aber die charakteristische Funktion, verknüpft mit einem Operator, sein soll, ist schwierig zu verstehen. Einen solchen Ausdruck kann man nicht auf einfache Weise definieren.  $\mathbb{1}$  ist nicht einmal stetig, es gibt also keine konvergente Reihe oder dergleichen. Im Buch von Riesz/Notch wird das Spektraltheorem jedoch auf eine sehr einsichtige Weise eingeführt. Zunächst werden die nötigen Ansätze für Polynome entwickelt, dann werden stetige Funktionen mit deren Hilfe approximiert. Als letzter Schritt folgt die Verallgemeinerung für Funktionen mit abzählbar vielen Unstetigkeitsstellen. Auf diese Weise wollen wir hier ebenfalls vorgehen.

## 4.2 Funktionalkalkül für beschränkte selbstadjungierte Operatoren

Sei  $T \in \mathcal{L}(H)$  selbstadjungiert, also  $T = T^*$ . Das Ziel dieses Abschnittes ist es, Funktionen von  $T$  zu definieren und Eigenschaften für diese zu beweisen. Sei im folgenden

$$\mathbb{C}[x] = \left\{ p(x) = \sum_{n=0}^k \alpha_n x^n \mid \alpha_n \in \mathbb{C}, k \in \mathbb{N} \right\}$$

die Algebra aller Polynome über den komplexen Zahlen. In dieser Algebra ist  $x^0 := 1$  definiert. Auf dieselbe Weise können wir für unseren Operator  $T$  die Algebra

$$\mathbb{C}[T] = \left\{ p(T) = \sum_{n=0}^k \alpha_n T^n \mid \alpha_n \in \mathbb{C}, k \in \mathbb{N} \right\}$$

definieren, wobei hier  $T^0 := I$  gesetzt wird. Wir beweisen auf dem Weg zu unseren schrittweisen Definitionen zunächst zwei Hilfssätze.

**Lemma 4.2.1** (Spektraler Abbildungssatz für Polynome). Für  $p \in \mathbb{C}[x]$  gilt

$$\sigma(p(T)) = p(\sigma(T)) := \{p(\lambda) \mid \lambda \in \sigma(T)\}.$$

*Beweis.* Sei  $\lambda \in \sigma(T)$ . Das Polynom  $p(x) - p(\lambda) \cdot 1$  hat für  $x = \lambda$  eine Nullstelle. Daher gibt es ein  $q(x)$  mit

$$p(x) - p(\lambda) \cdot 1 = (x - \lambda)q(x) = q(x)(x - \lambda).$$

Da Polynomgleichungen eine Entsprechung in der Algebra  $\mathbb{C}[T]$  besitzen, können wir für den Operator  $T$  ebenfalls die Gleichung

$$p(T) - p(\lambda)I = (T - \lambda I)q(T) = q(T)(T - \lambda I)$$

aufstellen. Angenommen, es wäre  $p(\lambda) \in \rho(p(T))$ . Dann würde die Resolvente  $(p(T) - p(\lambda)I)^{-1} \in \mathcal{L}(H)$  existieren. Wir multiplizieren obige Gleichung mit dieser Resolvente, zuerst von rechts:

$$I = (T - \lambda I)q(T)(p(T) - p(\lambda)I)^{-1}$$

und ebenfalls von links:

$$I = (p(T) - p(\lambda)I)^{-1}q(T)(T - \lambda I).$$

Wir setzen zur Abkürzung  $A := q(T)(p(T) - p(\lambda)I)^{-1}$  und  $B := (p(T) - p(\lambda)I)^{-1}q(T)(T - \lambda I)$ . Es gilt  $A, B \in \mathcal{L}(H)$  und mit unseren Abkürzungen lauten die obigen Gleichungen  $I = (T - \lambda I)A = B(T - \lambda I)$ . Damit folgt

$$B = B((T - \lambda I)A) = (B(T - \lambda I))A = A,$$

also  $B = A = (T - \lambda I)^{-1}$  im Widerspruch zu  $\lambda \in \sigma(T)$ . Damit muss die Annahme falsch und  $p(\lambda) \in \sigma(p(T))$  sein.

Sei umgekehrt  $\mu \in \sigma(p(T))$ . Zu zeigen ist, dass es ein  $\lambda \in \sigma(T)$  gibt mit  $\mu = p(\lambda)$ . Wir zerlegen das Polynom  $p$  in seine Linearfaktoren:

$$p(x) - \mu = a_n(x - \beta_1) \cdots (x - \beta_n)$$

und machen dasselbe mit dem Polynom in  $T$ :

$$p(T) - \mu I = a_n(T - \beta_1 I) \cdots (T - \beta_n I).$$

Wären alle  $\beta_i \in \rho(T)$ , dann hätte die rechte Seite ein beschränktes Inverses und damit auch die linke Seite, was im Widerspruch zu  $\mu \in \sigma(p(T))$  stünde. Also gibt es ein  $\beta_i \in \sigma(T)$ , für das dann  $p(\beta_i) - \mu = 0$  gilt. Daraus folgt  $\mu = p(\beta_i)$ .  $\square$

Im Limes erhält man durch Approximation denselben Satz für alle stetigen Funktionen. Die Aussage gilt sogar für *sämtliche* Funktionen, sofern man die Menge  $f(\sigma(T))$  noch abschließt. Dies ist auch nach unseren bisherigen Begriffsbildungen zumindest intuitiv klar, denn beim Spektrum handelt es sich notwendig um eine abgeschlossene Menge. Es folgt das zweite angekündigte Lemma, eine topologische Aussage.

**Lemma 4.2.2** (Erweiterungssatz von Tietze). Sei  $M$  eine abgeschlossene, beschränkte Teilmenge von  $\mathbb{R}$ , und es sei etwa  $M \subseteq [a, b]$ . Jede stetige Funktion  $f$  auf  $M$  lässt sich zu einer stetigen Funktion  $\tilde{f}$  auf  $[a, b]$  erweitern.

*Beweis.* Den Beweis findet man in gängigen Topologiebüchern, etwa in Franz: Topologie Band I.  $\square$

Aus diesem Lemma können wir eine für uns wichtige Folgerung ziehen.

**Korollar 4.2.3.** Sei  $M$  eine beschränkte und abgeschlossene Teilmenge in  $\mathbb{R}$ . Die Menge  $\mathbb{C}[x]$  aller Polynome auf  $M$  ist dicht im Banachraum  $(\mathcal{C}(M), \|\cdot\|)$ , wobei  $\|f\| = \sup_{t \in M} |f(t)|$ .

*Beweis.* Erweitere  $f \in \mathcal{C}(M)$  zu  $\tilde{f} \in \mathcal{C}[a, b]$  und approximiere  $\tilde{f}$  nach dem Satz von Weierstraß in der gleichmäßigen Konvergenz durch Polynome.  $\square$

Nun werden wir eine enge Verknüpfung zwischen den beiden Algebren  $\mathbb{C}[x]$  und  $\mathbb{C}[T]$  herstellen, indem wir einen  $*$ -Homomorphismus feststellen. Dieser lässt sich für das Spektrum  $\sigma(T)$  jedes Operators  $T$  eindeutig zu einem stetigen  $*$ -Homomorphismus auf  $(\mathcal{C}(\sigma(T)), \|\cdot\|)$  erweitern. Daraus folgen dann sehr viele angenehme Eigenschaften unseres Funktionalkalküls.

**Lemma 4.2.4.** Die Abbildung  $\varphi_0(p(x)) = p(T)$  von  $\mathbb{C}[x]$  auf  $\mathbb{C}[T] \subseteq \mathcal{L}(H)$  ist ein  $*$ -Homomorphismus der  $*$ -Algebra  $\mathbb{C}[x]$  in die  $*$ -Unteralgebra  $\mathbb{C}[T]$  von  $\mathcal{L}(H)$ .

*Beweis.* Man ermittelt die Identitäten  $\varphi_0(p_1 p_2) = \varphi_0(p_1) \varphi_0(p_2)$  und  $\varphi_0(\bar{p}) = \varphi_0(p)^*$  durch Nachrechnen. Beispielsweise gilt

$$\varphi_0(\bar{p}) = \sum_{n=0}^k \overline{\alpha_n} T^n = \left( \sum_{n=0}^k \alpha_n T^n \right)^* = \varphi_0(p)^*.$$

Hier haben wir explizit die Selbstadjungiertheit von  $T$  benötigt.  $\square$

**Satz 4.2.5** (Funktionalkalkül). *Es existiert genau ein stetiger  $*$ -Homomorphismus  $\varphi$  von  $(\mathcal{C}(\sigma(T)), \|\cdot\|)$  in die  $C^*$ -Algebra  $\mathcal{L}(H)$ , für den  $\varphi(p) = \varphi_0(p)$  für alle Polynome  $p \in \mathbb{C}[x]$  gilt. Dieser heißt der stetige Funktionalkalkül von  $T$ . Es ist  $\varphi(1) = I$  und  $\varphi(x) = T$ .*

*Beweis.* Zunächst beweisen wir die Existenz eines solchen Homomorphismus. Sei  $p \in \mathbb{C}[x]$  ein Polynom. Es gilt für selbstadjungierte Operatoren  $S$  stets  $\|S\|^2 = \|S^* S\|$ , also auch

$$\|\varphi_0(p)\|^2 = \|(\varphi_0(p))^* \varphi_0(p)\| = \|\varphi_0(\bar{p}p)\| = r(\varphi_0(\bar{p}p)) = \sup\{|\lambda| \mid \lambda \in \sigma(\varphi_0(\bar{p}p))\},$$

denn  $\varphi_0(\bar{p}p)$  ist ein selbstadjungierter, beschränkter Operator, für den ein Spektralradius existiert. Wir haben weiterhin die  $*$ -Homomorphismeigenschaft von  $\varphi_0$  ausgenutzt. Es gilt weiterhin  $\varphi_0(\bar{p}p) = (\bar{p}p)(T)$ , also

$$\sigma(\varphi_0(\bar{p}p)) = \sigma((\bar{p}p)(T)) = \bar{p}p(\sigma(T)) = \{\bar{p}p(\mu) \mid \mu \in \sigma(T)\}$$

nach Lemma 4.2.1. Damit haben wir

$$\|\varphi_0(p)\|^2 = \sup\{|\lambda| \mid \lambda \in \sigma(\varphi_0(\bar{p}p))\} = \|\bar{p}p\| = \|p\|^2,$$

wobei für  $\|\bar{p}p\|$  die Norm über das Spektrum gebildet wurde. Damit erhält der Homomorphismus die Norm, das heißt es gilt  $\|\varphi_0(p)\| = \|p(T)\| = \|p(\lambda)\| = \sup_{\lambda \in \sigma(T)} |p(\lambda)|$ . Da nach Folgerung 4.2.3  $\mathbb{C}[x]$  dicht in  $\mathcal{C}(\sigma(T))$  liegt, lässt sich der Operator  $\varphi_0 : \mathbb{C}[x] \rightarrow \mathbb{C}[T]$  zu einem stetigen  $*$ -Homomorphismus  $\varphi$  von  $\mathcal{C}(\sigma(T))$  nach  $\mathcal{L}(H)$  fortsetzen.

Wir zeigen nun die Eindeutigkeit dieses Homomorphismus. Seien  $\varphi, \tilde{\varphi}$  zwei solche stetigen  $*$ -Homomorphismen. Dann ist  $\varphi(p) = \tilde{\varphi}(p)$  für alle  $p \in \mathbb{C}[x]$ . Da  $\mathbb{C}[x]$  dicht in  $\mathcal{C}(\sigma(T))$  liegt, folgt aufgrund der Stetigkeit  $\varphi(f) = \tilde{\varphi}(f)$  für alle  $f \in \mathcal{C}(\sigma(T))$ .  $\square$

**Bemerkung 4.2.6.** Es folgt aus diesem Satz, dass zu jeder stetigen Funktion  $f$  auf  $\sigma(T)$  der Operator  $\varphi(f)$  existiert. Man schreibt dafür auch  $f(T) := \varphi(f)$ .



Das Wichtigste an dem eben bewiesenen Satz ist die Homomorphieeigenschaft. Alles, was für Funktionen gilt, ist damit auch gültig für Operatoren. Beispielsweise gilt  $e^{it\lambda}e^{is\lambda} = e^{i(t+s)\lambda}$ , woraus für einen Operator  $T$  auch  $e^{itT}e^{isT} = e^{i(t+s)T}$  folgt. Dieses Ergebnis ist umso erstaunlicher, als dass es für beliebig komplizierte Operatoren absolut nicht intuitiv ist, insbesondere nach unseren vorherigen Überlegungen zu Definitionen von Funktionen eines Operators über Integrale. Insbesondere müssen diese Integrale die für sie eher untypische multiplikative Homomorphieeigenschaft erfüllen.

Konkret besitzen wir folgende Relationen durch die \*-Homomorphieeigenschaft:

$$(\bar{f})(T) = f(T)^*, \quad (f_1 + f_2)(T) = f_1(T) + f_2(T), \quad (f_1 \cdot f_2)(T) = f_1(T) \cdot f_2(T), \quad (\lambda f)(T) = \lambda f(T).$$

Dies ist nun tatsächlich der sogenannte *Funktionalkalkül*. Er besitzt folgende weitere Eigenschaften.

**Satz 4.2.7.**

- (i)  $\|f(T)\| = \|f\|$  für  $f \in \mathcal{C}(\sigma(T))$ .
- (ii) Ist  $f(\lambda) \geq 0$  für alle  $\lambda \in \sigma(T)$ , so folgt  $f(T) \geq 0$ .
- (iii)  $f(T)$  ist normal für alle  $f \in \mathcal{C}(\sigma(T))$ .
- (iv)  $\sigma(f(T)) = f(\sigma(T))$  für alle  $f \in \mathcal{C}(\sigma(T))$ .

*Beweis.* Wir geben lediglich Beweisskizzen. (i): Aus dem Beweis von Satz 4.2.5 wissen wir, dass  $\|\varphi_0(p)\| = \|p(T)\| = \|p(\lambda)\|$  für Polynome  $p \in \mathbb{C}[x]$  gilt. Stetige Fortsetzung und Bildung von Limiten auf beiden Seiten zeigt  $\|f(T)\| = \|f(\lambda)\|$  für  $f \in \mathcal{C}(\sigma(T))$ .

(ii): Da  $f \geq 0$  auf  $\sigma(T)$  ist, ist  $g = \sqrt{f}$  eine stetige Funktion auf  $\sigma(T)$  und insbesondere reellwertig. Damit folgt

$$f(T) = (g^2)(T) = g(T)g(T) = g(T)^*g(T) \geq 0.$$

(iii): Wir haben durch Anwendung der Homomorphieeigenschaften

$$f(T)f(T)^* = f(T)\bar{f}(T) = (f\bar{f})(T) = (\bar{f}f)(T) = \bar{f}(T)f(T) = f(T)^*f(T).$$

(iv): folgt aus Lemma 4.2.1 mit einigen technischen Schritten. □

Es handelt sich bei der Aussage in (i) um eine relativ starke, denn wenn nur  $f = 0$  auf  $\sigma(T)$  ist, ist  $f(T)$  automatisch der Nulloperator. Dabei kann  $f$  außerhalb von  $\sigma(T)$  beliebige Werte annehmen.

Nun folgt endlich das Spektraltheorem. Es zeigt, dass unsere verschiedenen Definitionen von  $f(T)$  dasselbe Ergebnis liefern und garantiert weiterhin, dass man jeden selbstadjungierten, beschränkten Operator nach einer Spektralschar entwickeln kann. Diese Spektralschar lebt lediglich auf dem Intervall, das durch die untere und obere Grenze des Operators,  $m_T$  und  $M_T$ , gegeben ist.

**Satz 4.2.8** (Spektraltheorem). *Sei  $T$  ein beschränkter, selbstadjungierter Operator auf dem Hilbertraum  $H$ .  $m = m_T$  und  $M = M_T$  seien die untere beziehungsweise obere Grenze von  $T$ . Dann gilt*

(i) *Es existiert eine Spektralschar  $(E(\lambda) \mid \lambda \in \mathbb{R})$  mit folgenden Eigenschaften:*

- 1)  $E(\lambda) = 0$  für  $\lambda < m$  und  $E(\lambda) = I$  für  $\lambda \geq M$ .
- 2)  $u(T) = \int_m^M u(\lambda) E(d\lambda)$  für alle  $u \in \mathcal{C}(\sigma(T))$ .
- 3) Insbesondere ist  $T = \int_m^M \lambda E(d\lambda)$ .

(ii) *Die Spektralschar  $(E(\lambda) \mid \lambda \in \mathbb{R})$  ist durch 1) und 3) eindeutig bestimmt.*

*Beweis.* Im Beweis werden wir zunächst eine Spektralschar  $(E(\lambda) \mid \lambda \in \mathbb{R})$  konstruieren, danach die zweite Aussage des Satzes verifizieren und zum Schluss die Eindeutigkeit beweisen. Wir beginnen mit der Funktion

$$g_{\lambda,n}(t) = \begin{cases} 1 & t \leq \lambda \\ 0 & t \geq \lambda + \frac{1}{n} \\ 1 - n(t - \lambda) & \lambda < t < \lambda + \frac{1}{n}. \end{cases}$$

Diese ist offensichtlich stetig und auf ganz  $\mathbb{R}$  gilt  $1 \geq g_{\lambda,n} \geq g_{\lambda,n+1} \geq 0$ . Nach dem Funktionenkalkül wird die Ordnungsrelation auch bei Operatoren erhalten, also folgt entsprechend  $I \geq g_{\lambda,n}(T) \geq g_{\lambda,n+1}(T) \geq 0$  und da wir  $T$  selbstadjungiert gewählt haben, ist  $(g_{\lambda,n})_{n \in \mathbb{N}}$  eine monoton fallende, beschränkte Folge selbstadjungierter Operatoren. Deshalb existiert nach Satz 3.3.9 der starke Grenzwert  $E(\lambda)x = \lim_{n \rightarrow \infty} g_{\lambda,n}(T)x$ . Dies definiert uns eine Spektralschar, deren Axiome im folgenden nachgewiesen werden.

Nach Satz 3.3.9 folgt ebenfalls, dass auch  $E(\lambda)$  selbstadjungiert ist, da alle  $g_{\lambda,n}$  es sind. Wir wenden nun erneut Resultate, die wir für  $g_{\lambda,n}$  auf  $\mathbb{R}$  erhalten, auf Operatoren an und zeigen damit, dass  $E(\lambda)$  sogar ein Projektionsoperator ist. Es gilt nämlich für alle  $t \in \mathbb{R}$  die Ungleichung  $g_{\lambda,2n}(t) \leq (g_{\lambda,n}(t))^2 \leq g_{\lambda,n}(t)$ . Also folgt ebenfalls  $g_{\lambda,2n}(T) \leq (g_{\lambda,n}(T))^2 \leq g_{\lambda,n}(T)$ . Der Grenzübergang für  $n \rightarrow \infty$  ergibt dann  $E(\lambda) \leq (E(\lambda))^2 \leq E(\lambda)$  und damit ist offensichtlich  $E(\lambda)$  ein Projektionsoperator.

Ist  $\lambda < m$ , so ist nach Definition  $g_{\lambda,n}(t) = 0$  für alle  $t \in \sigma(T)$ , wenn  $n$  hinreichend groß ist, denn es gilt ja  $\sigma(T) \subseteq [m, M]$ . Damit ist dann auch  $g_{\lambda,n}(T) = 0$  und damit  $E(\lambda) = 0$ . Analog folgert man, dass  $E(\lambda) = I$  ist, wenn  $\lambda \geq M$  und  $n$  genügend groß ist. Für  $\mu < \lambda$  folgt aus der Monotonie von  $(g_{\lambda,n})_n$ , dass  $g_{\mu,n}(t) - g_{\lambda,k}(t) \geq 0$  ist, wenn  $\frac{1}{k} - \frac{1}{n} \leq \mu - \lambda$  ist. Auch hier folgern wir dieselbe Aussage für Operatoren und damit, dass  $g_{\mu,n}(T) \geq g_{\lambda,k}(T)$  und somit  $E(\mu) \geq E(\lambda)$  ist. Damit ist die Monotonie gezeigt und wir müssen uns nun nur noch der Rechtsstetigkeit zuwenden.

Betrachte die Gleichung

$$\begin{aligned} \|(E(\lambda + \varepsilon) - E(\lambda))x\|^2 &= \langle (E(\lambda + \varepsilon) - E(\lambda))x, (E(\lambda + \varepsilon) - E(\lambda))x \rangle \\ &\leq \langle (E(\lambda + \varepsilon) - E(\lambda))x, x \rangle \\ &\leq \langle (g_{\lambda+\varepsilon,n}(T) - E(\lambda))x, x \rangle \\ &= \langle (g_{\lambda+\varepsilon,n} - g_{\lambda,n}(T))x, x \rangle + \langle (g_{\lambda,n}(T) - E(\lambda))x, x \rangle \end{aligned}$$

Im folgenden werden wir zeigen, dass dieser Ausdruck beliebig klein wird und  $E(\lambda)$  damit rechtsstetig ist. Sei also  $\delta > 0$ . Weil  $\lim_{n \rightarrow \infty} (g_{\lambda,n}(T) - E(\lambda))x = 0$  ist, gibt es ein  $n_0 \in \mathbb{N}$ , so dass  $\|(g_{\lambda,n}(T) - E(\lambda))x\| < \delta$  für alle  $n > n_0$  ist. Man sieht leicht, dass man ein  $\varepsilon > 0$  so wählen kann, dass  $n\varepsilon < 1$  und  $\|g_{\lambda+\varepsilon,n} - g_{\lambda,n}\| < n\varepsilon$  gilt. Also gilt

$$|\langle (g_{\lambda+\varepsilon,n}(T) - g_{\lambda,n}(T))x, x \rangle + \langle (g_{\lambda,n}(T) - E(\lambda))x, x \rangle| \leq \varepsilon n \|x\|^2 + \delta$$

für  $n > n_0$ . Wenn man nun ein  $\varepsilon_0$  klein genug wählt, so ist  $n_0\varepsilon_0 \|x\|^2 < \delta$  und damit haben wir

$$\|(E(\lambda + \varepsilon) - E(\lambda))x\|^2 \leq 2\delta,$$

wobei wir  $\delta > 0$  beliebig klein wählen können. Damit haben wir also gezeigt, dass die konstruierte Schar  $(E(\lambda) \mid \lambda \in \mathbb{R})$  rechtsstetig ist und somit alle Eigenschaften einer Spektralschar erfüllt. Der erste Schritt des Beweises ist damit beendet.

Im zweiten Schritt zeigen wir

$$u(T) = \int_{m \uparrow}^M u(\lambda) E(d\lambda)$$

für  $u \in C(\sigma(T))$ , dass also das durch die konstruierte Spektralschar gegebene Integral gegen den durch den Funktionenkalkül bestimmten Operator  $u(T)$  konvergiert. Sei  $u(t)$  o.B.d.A. reellwertig. Auf dem Intervall  $[m, M]$  bilden wir die Zerlegung  $\mathfrak{z}$  mit den Stützpunkten  $t_0 = m < t_1 < \dots < t_n = M$  mit der

Feinheit  $\delta_{\mathfrak{z}} = \max_{i=1, \dots, n} |t_i - t_{i-1}|$ . Sei nun  $k < \frac{1}{2}\delta_{\mathfrak{z}}$ . Wir wählen Zwischenwerte  $\mathfrak{z}_i \in [t_{i-1}, t_i]$  und setzen  $m_i := \inf_{t \in [t_{i-1}, t_i]} u(t)$  und  $M_i := \sup_{t \in [t_{i-1}, t_i]} u(t)$ , sowie  $\delta(u) := \sup_{i=1, \dots, n} |M_i - m_i|$ . Es gilt

$$m_i(g_{t_i, k} - g_{t_{i-1}, k}) \leq u(t)(g_{t_i, k} - g_{t_{i-1}, k}) \leq M_i(g_{t_i, k} - g_{t_{i-1}, k})$$

für die oben definierten Funktionen und damit auch für die entsprechenden Operatoren, so dass wir für den Grenzübergang  $k \rightarrow \infty$  die Abschätzung

$$m_i(E(t_i) - E(t_{i-1})) \leq u(T)(E(t_i) - E(t_{i-1})) \leq M_i(E(t_i) - E(t_{i-1})).$$

erhalten. Es gilt nun durch Teleskopsummenbildung

$$u(T) - S(u, \mathfrak{z}) = u(T) - \sum_{i=1}^n u_i(\mathfrak{z}_i)(E(t_i) - E(t_{i-1})) = \sum_{i=1}^n (u(T) - u(\mathfrak{z}_i))(E(t_i) - E(t_{i-1})),$$

denn

$$\sum_{i=1}^n u(T)(E(t_i) - E(t_{i-1})) = u(T)(E(t_n) - E(t_0)) = u(T)(I - 0) = u(T).$$

Wir können diese Summe abschätzen und erhalten erneut eine Teleskopsumme:

$$\begin{aligned} \sum_{i=1}^n (u(T) - u(\mathfrak{z}_i))(E(t_i) - E(t_{i-1})) &\leq \sum_{i=1}^n (M_i - m_i)(E(t_i) - E(t_{i-1})) \\ &\leq \sum_{i=1}^n \delta(u)(E(t_i) - E(t_{i-1})) \\ &= \delta(u)I. \end{aligned}$$

Analog zeigt man  $S(u, \mathfrak{z}) - u(T) \leq \delta(u)I$  und aus diesen beiden Ungleichungen erhalten wir insgesamt die Abschätzung  $\|u(T) - S(u, \mathfrak{z})\| \leq \delta(u)$ .

Wenn wir nun ein  $\varepsilon > 0$  wählen, gibt es nach dem Satz von Cantor ein  $\delta > 0$ , so dass  $|u(\lambda_i) - u(\mu_i)| < \varepsilon$  gilt, falls nur  $|\lambda_i - \mu_i| < \delta$  ist, wobei  $\mu_i, \lambda_i \in [m, M]$  sein sollen. Wir haben aber soeben gezeigt, dass aus  $|t_i - t_{i-1}| \leq \delta$  folgt, dass  $|M_i - m_i| < \delta(u)$  ist. Also wählen wir  $\delta_{\mathfrak{z}} \leq \delta$  und haben dann  $\delta(u) < \varepsilon$ , woraus folgt, dass  $\|u(T) - S(u, \mathfrak{z})\| \leq \varepsilon$  ist, wenn wir die Feinheit kleiner als  $\delta$  wählen. Somit ist die zweite Aussage des Spektralsatzes bewiesen.

Wir wollen zum Schluss noch die Eindeutigkeit der Spektralschar zeigen. Sei hierzu zunächst  $(E(\lambda) \mid \lambda \in \mathbb{R})$  eine beliebige Spektralschar, die die Eigenschaften (i) und (iii) erfüllt. Dann gilt für diese

$$T = \int_{m \uparrow}^M \lambda E(d\lambda) = \lim_{\delta_{\mathfrak{z}} \rightarrow 0} \sum_{i=1}^n \mathfrak{z}_i \underbrace{(E(t_i) - E(t_{i-1}))}_{\Delta E_i}$$

und

$$\left( \sum_{i=1}^n \mathfrak{z}_i \Delta E_i \right)^2 = \sum_{i, j=1}^n \mathfrak{z}_i \mathfrak{z}_j \underbrace{\Delta E_i \Delta E_j}_{=0, \text{ wenn } i \neq j} = \sum_{i=1}^n (\mathfrak{z}_i)^2 \Delta E_i.$$

Analog erhält man für  $r \in \mathbb{N}$  den Ausdruck

$$\left( \sum_{i=1}^n \mathfrak{z}_i \Delta E_i \right)^r = \sum_{i=1}^n (\mathfrak{z}_i)^r \Delta E_i$$

als Zwischensumme für  $(\int \lambda E(d\lambda))^r$  und somit

$$T^r = \left( \int_{m \uparrow}^M \lambda E(d\lambda) \right)^r = \int_{m \uparrow}^M \lambda^r E(d\lambda).$$

Also folgt für Polynome  $p(T)$  die Gleichung  $p(T) = \int_{m\uparrow}^M p(\lambda) E(d\lambda)$ . Nach dieser kurzen Betrachtung ist der Rest nicht schwer. Seien  $E$  und  $\tilde{E}$  zwei Spektralscharen mit den Eigenschaften (i) und (iii) aus dem Theorem. Nach dem eben Gezeigten gilt

$$p(T) = \int_{m\uparrow}^M p(\lambda) E(d\lambda) = \int_{m\uparrow}^M p(\lambda) \tilde{E}(d\lambda).$$

Somit gilt auch

$$\langle p(T)x, x \rangle = \int_{m\uparrow}^M p(\lambda) \langle E(d\lambda)x, x \rangle = \int_{m\uparrow}^M p(\lambda) \langle \tilde{E}(d\lambda)x, x \rangle,$$

also folgt die Gleichheit beider Scharen für stetige Funktionen auf  $[m - \varepsilon, M]$  und  $\langle E(\lambda)x, x \rangle$ . Weiterhin sind  $\langle \tilde{E}(\lambda)x, x \rangle$  als Maße gleich auf diesem Intervall. Wie im ersten Schritt gezeigt, gilt  $E(\lambda) = 0 = \tilde{E}(\lambda)$  für  $\lambda < m$  und  $E(\lambda) = I = \tilde{E}(\lambda)$  für  $\lambda \geq M$ . Somit müssen die Maße aber auf ganz  $\mathbb{R}$  übereinstimmen und mittels der Polarisierungsidentität folgt  $E(\lambda) = \tilde{E}(\lambda)$ .  $\square$

**Bemerkung 4.2.9.** Die Abbildung  $\mathcal{C}(\sigma(T)) \rightarrow \mathcal{L}(H)$ , die durch  $u(\lambda) \mapsto u(T)$  gegeben ist, ist ein  $*$ -Homomorphismus. Deshalb folgt nach Satz 4.2.8 2) (ii) die Identität

$$\int_{\uparrow m}^M u_1(\lambda) E(d\lambda) + \int_{\uparrow m}^M u_2(\lambda) E(d\lambda) = \int_{\uparrow m}^M (u_1 + u_2)(\lambda) E(d\lambda)$$

und weiterhin

$$\left( \int_{\uparrow m}^M u_1(\lambda) E(d\lambda) \right) \left( \int_{\uparrow m}^M u_2(\lambda) E(d\lambda) \right) = \int_{\uparrow m}^M (u_1 u_2)(\lambda) E(d\lambda)$$

sowie

$$\left( \int_{\uparrow m}^M u(\lambda) E(d\lambda) \right)^* = \int_{\uparrow m}^M \bar{u}(\lambda) E(d\lambda).$$

Weiterhin gilt, und dies ist entscheidend für die Eindeutigkeit,

$$T = \int_{\uparrow m}^M \lambda E(d\lambda)$$

sowie

$$I = \int_{\uparrow m}^M 1 E(d\lambda).$$

In einigen Büchern werden diese beiden Aussagen als Eindeutigkeitsaussagen für die Spektralschar formuliert.

Eine naheliegende philosophische Frage ist, wie man sich erklären kann, dass etwa

$$e^{i(t+s)H} = e^{itH} e^{isH}$$

für einen Operator  $H$  ist. Mathematisch ist dieser Sachverhalt hochgradig nichttrivial und entspricht dennoch der physikalischen Realität, wenn  $U(t) := e^{itH}$  etwa der Zeitentwicklungsoperator eines quantenmechanischen Systems ist. Man erwartet dann aus physikalischer Sicht

$$U(t+s) = U(t)U(s),$$

denn ob man zuerst die Zeit  $t$  und dann  $s$  abwartet, oder andersherum, oder die Gesamtzeit  $t+s$  vergehen lässt, muss für das System egal sein. Wenn dies nicht stimmen würde, könnte man die gesamte Physik vergessen.

**Beispiel 4.2.10.** Sei  $H = L^2([a, b], \mu)$  der betrachtete Hilbertraum. Wir definieren eine Spektralschar durch die Wirkung  $(E(\lambda)f)(t) = e_\lambda(t)f(t)$  auf Elemente  $f \in H$ , siehe Beispiel 4.1.5. Man kann zeigen, dass

$$\left( \int_{\uparrow a}^b u(\lambda) E(d\lambda) \right) f(t) = u(t)f(t)$$

ist. Damit wäre  $T = \int \lambda E(d\lambda)$  der Multiplikationsoperator mit  $t$ , also  $(Tf)(t) = tf(t)$ . Dann ist es klar, dass man zwei Integrale multiplizieren kann und

$$\left( \int u_1(\lambda) E(d\lambda) \right) \left( \int u_2(\lambda) E(d\lambda) \right) f = u_1(u_2f) = (u_1u_2)f = \left( \int u_1(\lambda)u_2(\lambda) E(d\lambda) \right) f$$

erhält. Der Vergleich mit den Verteilungsfunktionen aus der Wahrscheinlichkeitstheorie ist also nicht besonders gut gewählt, da  $\int dE$  keine Zahl, sondern ein Operator ist.

Nun kommt eine Art philosophische Aussage: Alle wichtigen Informationen über einen Operator finden sich in seiner Spektralschar. Dies stimmt nur eingeschränkt, wie wir am Beispiel der Voltterraschen Integraloperatoren sahen, dennoch ist der folgende Satz sehr mächtig.

**Satz 4.2.11.** *Sei  $T \in \mathcal{L}$  ein beschränkter, selbstadjungierter Operator mit der eindeutig bestimmten Spektralschar  $(E(\lambda) \mid \lambda \in \mathbb{R})$ . Jede nichtreelle Zahl liegt in  $\rho(T)$ . Für ein  $\lambda_0 \in \mathbb{R}$  gibt es zwei Kriterien, um zu bestimmen, ob  $\lambda_0 \in \sigma(T)$  ist. Es ist*

- (i)  $\lambda_0 \in \rho(T)$  genau dann, wenn es ein  $\varepsilon > 0$  gibt, so dass  $E(\lambda) = E(\lambda_0)$  für alle  $\lambda \in (\lambda_0 - \varepsilon, \lambda_0 + \varepsilon)$  ist, wenn  $\lambda_0$  also in einem Intervall liegt, auf dem die Spektralschar konstant ist.
- (ii)  $\lambda_0 \in \sigma_p(T)$ , also  $\lambda_0$  ein Eigenwert von  $T$  genau dann, wenn  $E(\lambda_0) \neq E(\uparrow\lambda_0)$  ist, wenn die Spektralschar also in  $\lambda_0$  unstetig ist. Der Eigenraum ist dabei so groß wie der Unterraum

$$(E(\lambda_0) - E(\uparrow\lambda_0))H = \{x \in H \mid (T - \lambda_0 I)x = 0\}.$$

*Beweis.* Zunächst beweisen wir (i). Sei  $\lambda_0 \in \rho(T)$ . Dann gibt es ein  $\varepsilon > 0$ , so dass das Intervall  $(\lambda_0 - 2\varepsilon, \lambda_0 + 2\varepsilon]$  eine Teilmenge von  $\rho(T)$  ist, denn die Resolventenmenge ist ja offen. Wir wählen weiter ein  $n > \frac{1}{\varepsilon}$ . Dann sind die Funktionen  $g_{\lambda,n}$  (siehe Beweis des Satzes 4.2.8) mit  $\lambda_0 - \varepsilon \leq \lambda \leq \lambda_0 + \varepsilon$  alle gleich auf  $\sigma(T)$ , also haben wir für  $n \rightarrow \infty$  auch die Gleichheit des Grenzwertes  $E(\lambda) = \lim_{n \rightarrow \infty} g_{\lambda,n}(T)$  für alle diese  $\lambda$ . Somit ist  $E(\lambda)$  in der Umgebung  $|\lambda - \lambda_0| < \varepsilon$  konstant, also  $E(\lambda) = E(\lambda_0)$ . Nehmen wir dies umgekehrt als Voraussetzung, so gilt für alle stetigen  $f$  die Zerlegung

$$f(T) = \int_{m\uparrow}^M f(\lambda) E(d\lambda) = \int_a^{\lambda_0 - \varepsilon} f(\lambda) E(d\lambda) + \int_{\lambda_0 + \varepsilon}^b f(\lambda) E(d\lambda)$$

mit  $a$  und  $b$  so gewählt, dass  $a < m, b \geq M$  und  $a < \lambda_0 - \varepsilon < \lambda_0 + \varepsilon < b$  erfüllt ist. Sei

$$g(t) = \begin{cases} \frac{1}{t - \lambda_0} & |t - \lambda_0| \geq \varepsilon \\ \frac{t - \lambda_0}{\varepsilon^2} & |t - \lambda_0| < \varepsilon. \end{cases}$$

Wie man leicht sieht, ist  $g$  stetig. Sei weiter  $f(t) = (t - \lambda_0)g(t)$ . Dann gilt für  $f(T)$ :

$$f(T) = \int_a^{\lambda_0 - \varepsilon} 1 E(d\lambda) + \int_{\lambda_0 + \varepsilon}^b 1 E(d\lambda) = E(\lambda_0 - \varepsilon) - E(a) + E(b) - E(\lambda_0 + \varepsilon) = E(b) - E(a) = I.$$

Also ist wegen  $I = f(T) = (T - \lambda_0 I)g(T)$  der Operator  $g(T) = (T - \lambda_0 I)^{-1}$  die Resolvente von  $T$  im Punkte  $\lambda_0$  und  $\lambda_0$  liegt in der Resolventenmenge von  $T$ .

Wir zeigen nun (ii). Weil definitionsgemäß für alle  $x \in H$  die Konvergenz  $E(\uparrow\lambda_0)x = \lim_{\lambda \uparrow \lambda_0} E(\lambda)x$  erfüllt ist, gilt  $E(\uparrow\lambda_0) \leq E(\lambda)$  und damit ist  $E(\lambda_0) - E(\uparrow\lambda_0)$  ein Projektionsoperator. Dieser projiziert auf einen Unterraum  $H_0 = (E(\lambda_0) - E(\uparrow\lambda_0))H$ . Wir wählen ein  $x \in H_0$  und zeigen nun, dass  $x$  im Kern

von  $(T - \lambda_0 I)$  enthalten ist. Indem wir danach die umgekehrte Inklusion zeigen, ist der Beweis beendet. Für  $\lambda < \lambda_0$  gilt

$$E(\lambda)x = E(\lambda)(E(\lambda_0) - E(\uparrow\lambda_0))x = E(\lambda)x - E(\lambda)x = 0.$$

Hingegen gilt für  $\lambda \geq \lambda_0$  die Gleichung  $E(\lambda)x = E(\lambda_0)x - E(\uparrow\lambda_0)x = x$ . Desweiteren haben wir

$$\langle Tx, y \rangle = \int_{m\uparrow}^M \lambda \langle E(d\lambda)x, y \rangle = \lambda_0 \langle x, y \rangle$$

für alle  $y \in H$ . Also ist  $Tx = \lambda_0 x$ , das heißt  $H_0 \subseteq \text{Kern}(T - \lambda_0 I)$ . Sei nun umgekehrt  $x \in \text{Kern}(T - \lambda_0 I)$ , das heißt  $Tx = \lambda_0 x$ . Dann ist

$$0 = \|(T - \lambda_0)x\|^2 = \langle (T - \lambda_0)^2 x, x \rangle = \int_{m\uparrow}^M (\lambda - \lambda_0)^2 \langle E(d\lambda)x, x \rangle$$

und daher folgt  $\langle E(\lambda)x, x \rangle = 0$  für  $\lambda < \lambda_0$  und  $\langle E(\lambda)x, x \rangle = \|x\|^2$  für  $\lambda \geq \lambda_0$ , wobei daraus dann  $E(\lambda)x = 0$  beziehungsweise  $E(\lambda)x = x$  folgt. Insgesamt ergibt sich also  $(E(\lambda_0) - E(\uparrow\lambda_0))x = x - 0 = x$  und damit  $x \in H_0$ .  $\square$

**Bemerkung 4.2.12.** Es gilt nach (i):

$$u(T) = \int_{m\uparrow}^M u(\lambda) E(d\lambda) = \int_{\sigma(T)} u(\lambda) E(d\lambda).$$

In Quantenmechanikbüchern steht häufig als Spektralzerlegung eines Operators der Ausdruck

$$T = \int_{\sigma(T)} \lambda E(d\lambda).$$

Das stimmt so auf jeden Fall, es ist jedoch nicht so gut erklärbar wie unsere definierten Ausdrücke. Man denke dabei beispielsweise an den Fall, dass  $\sigma(T)$  eine Cantormenge ist.

Als Anwendung für obigen Satz sei folgendes Beispiel gebracht:

**Beispiel 4.2.13.** Sei  $T$  ein kompakter, selbstadjungierter Operator. Nach dem Hilbert-Schmidtschen Entwicklungssatz (Satz 3.5.15) gibt es ein Orthonormalsystem  $(e_k)$ , so dass  $Te_k = \lambda_k e_k$  gilt, wobei  $\lambda_k$  eine reelle Nullfolge ist. Wir haben dann für die Spektralschar den Ausdruck

$$E(\lambda)x = \sum_{k:\lambda_k \leq \lambda} \langle x, e_k \rangle e_k.$$

Es folgen nun abschließend einige wichtige Anwendungen des Spektraltheorems. Zunächst wollen wir die Existenz der Quadratwurzel für einen selbstadjungierten, positiven Operator beweisen.

**Satz 4.2.14.** Sei  $T$  ein selbstadjungierter, stetiger und positiver Operator. Dann existiert genau ein positiver Operator  $S$  mit  $S^2 = T$  und  $ST = TS$ . Für jeden Operator  $C \in \mathcal{L}(H)$ , für den  $CT = TC$  gilt, gilt dann auch  $CS = SC$ .

*Beweis.* Wir zeigen zunächst die Existenz eines solchen Operators  $S$ . Weil  $T$  positiv ist, ist auch das Spektrum auf der positiven Halbachse zu finden, also  $\sigma(T) \subseteq [0, M]$ . Damit ist  $f(\lambda) = \sqrt{\lambda}$  eine stetige Funktion auf  $\sigma(T)$ . Wir definieren nun den Operator  $S$  durch  $S := f(T)$  und erhalten

$$\langle Sx, x \rangle = \int_0^M \sqrt{\lambda} \langle E(d\lambda)x, x \rangle \geq 0$$

sowie  $S^2 = f(T)f(T) = f^2(T) = T$ . Weiterhin gilt

$$TS = Tf(T) = \int \lambda f(\lambda) E(d\lambda) = \int f(\lambda)\lambda E(d\lambda) = f(T)T = ST,$$

so dass  $S$  und  $T$  also vertauschen. Sei nun  $C$  ein weiterer beschränkter Operator mit  $CT = TC$ . Dann gilt für alle Polynome  $p_n \in \mathbb{C}[t]$  die Gleichung  $Cp_n(T) = p_n(T)C$  und weil die Menge komplexer Polynome in der Menge der stetigen Funktionen dicht liegt, folgt  $Cf(T) = f(T)C$ .

Die Eindeutigkeit ist nun nicht mehr schwer zu zeigen. Sei  $S' \geq 0$  ein weiterer Operator mit  $S'^2 = T$ . Dann gilt  $S'T = (S')^3 = (S')^2S' = TS'$ , also ist  $S'S = SS'$ . Desweiteren gilt

$$(S - S')S(S - S') + (S - S')S'(S - S') = (S^2 - S'^2)(S - S') = 0.$$

Da die beiden Summanden beide positiv sind, es handelt sich ja um positive Operatoren, müssen beide Terme ebenfalls 0 sein und damit folgt durch Kommutieren und Subtraktion  $(S - S')^3 = 0$ , also auch  $(S - S')^4 = 0$ . Es ergibt sich  $0 = \|(S - S')^4\| = \|(S - S')^2\|^2 = \|(S - S')\|^4 = 0$  und damit  $S = S'$ .  $\square$

Ein weiteres Thema von Interesse ist die Polarzerlegung beschränkter Operatoren. Von den komplexen Zahlen ist bekannt, dass sich jedes  $z \in \mathbb{C}$  auch in Polardarstellung schreiben lässt als  $z = re^{i\varphi}$ , wobei  $\varphi \in [0, 2\pi)$  und  $r = |z|$  ist. Im folgenden soll gezeigt werden, wie wir eine solche Polarzerlegung auch für Operatoren erhalten. Hierzu definieren wir zunächst den Begriff der partiellen Isometrie.

**Definition 4.2.15** (partielle Isometrie). Sei  $H$  ein Hilbertraum und  $U \in \mathcal{L}(H)$ . Der Operator  $U$  heißt eine *partielle Isometrie*, wenn die Einschränkung von  $U$  auf das orthogonale Komplement des Kerns von  $U$  isometrisch ist, wenn also  $(U|_{\text{Kern}(U)^\perp})^* U|_{\text{Kern}(U)^\perp} = I$  gilt.

Ein beschränkter Operator  $U$  ist genau dann eine partielle Isometrie, wenn  $UU^*U = U$  ist, was äquivalent zu  $U^*UU^* = U^*$  ist. Damit ist  $U$  genau dann eine partielle Isometrie, wenn  $U^*$  eine solche ist. Beispielsweise ist das Produkt von Operatoren mit bestimmten Eigenschaften eine partielle Isometrie:

**Beispiel 4.2.16.** Sei  $W$  ein unitärer und  $P$  ein Projektionsoperator. Dann ist  $WP$  eine partielle Isometrie, denn es gilt

$$WP(WP)^*WP = WPP^*W^*WP = WP.$$

Wenn wir einen Operator polarzerlegen wollen, dann benötigen wir auf jeden Fall den Betrag des Operators. Für einen beschränkten Operator  $T$  ist das Produkt  $T^*T \geq 0$ , also können wir für den Betrag  $|T| = (T^*T)^{1/2}$  definieren. Es gilt dann folgender Satz.

**Satz 4.2.17.** Sei  $T \in \mathcal{L}(H)$ . Dann existiert genau eine partielle Isometrie  $V$ , so dass  $T = V|T|$  ist und  $\text{Kern}T = \text{Kern}V = \text{Kern}|T|$  gilt.

# Index

- $C^*$ -Algebra, 52
- $C^*$ -Bedingung, 52
- Ableitung
  - verallgemeinerte, 13
- absolutstetig, 15
- adjungierter Operator, 51
- \*-Algebra, 52
- Äquivalenz
  - von Normen, 10
- Bairesches Kategoriethem, 30
- Banach-Schauder, Satz von, 39
- Banach-Steinhaus, Satz von, 37
- Banachalgebra, 52
- Beschränktheit
  - einer linearen Abbildung, 31
  - einer Menge, 31
- Besselsche Ungleichung, 23
- Cauchy-Schwarz-Ungleichung, 16
- Cauchyfolge, 7
- dicht, 7, 29
  - nirgends, 29
- direkte orthogonale Summe, 19
- diskrete Metrik, 6
- Dualraum, 36
- Eigenvektor, 62
- Eigenwert, 62
- Einheitskugel, 31
- Einheitssphäre, 31
- endlichdimensionaler Operator, 71
- Euklidische Metrik, 6
- Fouriertransformation, 35
- Funktional
  - lineares, 21
  - stetiges, 21
  - sublinear, 42
- Funktionalkalkül, 88
- Funktionalkalkül, 89
- Graph, 40
  - Satz vom abgeschlossenen, 41
- Grenze
  - eines Operators, 68
  - untere, 68
- Hölder-Ungleichung, 12
- Hahn-Banach, Satz von, 45
- Halbnorm, 9, 42
- Halbordnung, 41
- Hellinger-Toeplitz
  - Satz von, 52
- Hilbert-Schmidt-Norm, 76
- Hilbert-Schmidt-Operator, 34, 76
- Hilbert-Schmidtscher Entwicklungssatz, 73
- Hilbertraum, 17
- Hilbertraumisomorphismus, 19
- Hilbertscher Folgenraum, 19
- induzierte Metrik, 9
- Isometrie
  - partielle, 95
- isometrischer Isomorphismus, 36
- isometrischer Operator, 54
- Isomorphie
  - von normierten Räumen, 36
- Isomorphismus
  - eines Hilbertraums, 19
- kanonische Form, 74
- Kategorie, 29
- Kategoriethem, 30
- kompakt, 8, 68
  - Operator, 69
  - relativ, 68
- Konvergenz
  - im  $p$ -ten Mittel, 12
  - schwach, 59
  - von Operatoren, 60
- Konvexität, 6
- lineare Abbildung
  - beschränkt, 31
  - stetig, 31
- lineares Funktional, 21
- lokalkompakt, 8
- maximales Element, 42



- Mazur-Orlicz, Satz von, 39
- Menge
- kompakt, 68
  - relativ kompakt, 68
- messbare Funktion, 84
- Metrik, 5
- diskrete, 6
  - Euklidische, 6
  - induziert, 9
- Metrischer Raum, 5
- kompakt, 8
  - Vervollständigung, 7
  - Vollständigkeit, 7
- Minkowski-Funktional, 46
- Minkowski-Ungleichung, 12
- nirgends dicht, 29
- Norm, 9
- äquivalent, 10
  - Hilbert-Schmidt-, 76
- normaler Operator, 53
- nuklearer Operator, 76
- obere Grenze, 68
- obere Schranke, 42
- Operator
- adjungiert, 51
  - der Spurklasse, 76
  - endlichdimensional, 71
  - Hilbert-Schmidt-, 76
  - isometrisch, 54
  - kanonische Form, 74
  - kompakt, 69
  - normal, 53
  - nuklearer, 76
  - positiv, 54
  - selbstadjungiert, 53
  - transponiert, 70
  - unitär, 54
  - vollstetig, 69
- orthogonal, 17
- orthogonales Komplement, 20
- Orthogonalsystem, 22
- Orthonormalbasis, 24
- Orthonormalsystem, 22
- vollständiges, 24
- Parallelogrammidentität, 18
- Parsevalsche Gleichung, 24
- partielle Isometrie, 95
- Polarisationsidentität, 17
- positiver Operator, 54
- Projektionsoperator, 21, 57
- Punktspektrum, 63
- Pythagoras, Satz des, 23
- relativ kompakt, 68
- Resolvente, 62
- Resolventenidentitäten, 64
- Resolventenmenge, 62
- Riesz, 1. Satz von, 21
- Riesz, 2. Satz von, 21
- Riesz-Fischer, Satz von, 12
- Riesz-Schauder-Theorem, 71
- Satz
- vom abgeschlossenen Graphen, 41
  - von Hahn-Banach, 45
  - von Hellinger-Toeplitz, 52
  - des Pythagoras, 23
  - von Riesz (1), 21
  - von Riesz (2), 21
  - von Riesz-Fischer, 12
  - von Riesz-Schauder, 71
  - von Schauder, 71
- Satz von
- Banach-Schauder, 39
  - Banach-Steinhaus, 37
  - Mazur-Orlicz, 39
- Schattenklasse, 76
- Schauder
- Satz von, 71
- Schmidt-Darstellung, 74
- schwach konvergent, 68
- schwache Konvergenz, 59
- von Operatoren, 60
- Schwartzraum, 56
- selbstadjungierter Operator, 53
- Sesquilinearform, 61
- Sesquilinearität, 16
- $\sigma$ -kompakt, 8
- singuläre Zahlen, 76
- Sobolev-Lemma, 16
- Sobolev-Slobodecki-Raum, 15
- Sobolevraum, 13
- Spektralmaß, 82
- Spektralradius, 65
- Spektralschar, 80
- beschränkter Träger, 80
- Spektrum, 62
- Punktspektrum, 63
  - stetiges, 63
- Spurklasse, 76
- starke Konvergenz
- von Operatoren, 60
- stetiges Spektrum, 63
- Stetigkeit, 31
- sublineares Funktional, 42

Summe

    direkte orthogonale, 19

Teilraum, 7

Totalordnung, 42

Träger, 13

    einer Spektralschar, 80

transponierter Operator, 70

unitärer Operator, 54

untere Grenze, 68

Unterraum, 20

verallgemeinerte Ableitung, 13

Vollständigkeit, 24

    eines metrischen Raumes, 7

vollstetig, 69

Volterrascher Integraloperator, 64

Weyl-Lemma, 16