EINFÜHRUNG IN DIE FUNKTIONALANALYSIS

Vorlesungsskript, Sommersemester 2023

Christian Clason

Stand vom 14. August 2023

Institut für Mathematik und wissenschaftliches Rechnen
Universität Graz

Inhaltsverzeichnis

I TOPOLOGISCHE GRUNDLAGEN

1 METRISCHE RÄUME 3

2 KOMPAKTE MENGEN 8

II LINEARE OPERATOREN IN NORMIERTEN RÄUMEN

3 NORMIERTE VEKTORRÄUME 16

4 LINEARE OPERATOREN 27

5 DAS PRINZIP DER GLEICHMÄSSIGEN BESCHRÄNKTHEIT 33

6 QUOTIENTENRÄUME 40

III DUALRÄUME UND SCHWACHE KONVERGENZ

7 LINEARE FUNKTIONALE UND DUALRÄUME 46

8 DER SATZ VON HAHN–BANACH 51

9 ADJUNGIERTE OPERATOREN 61

10 REFLEXIVITÄT 68

11 SCHWACHE KONVERGENZ 72

IV HILBERTRÄUME

12 SKALARPRODUKTE UND ORTHOGONALITÄT 80

13 DER SATZ VON RIESZ 92

ÜBERBLICK

Die Funktionalanalysis wurde Anfang des 20. Jahrhunderts entwickelt, um allgemeine Aussagen über die Lösbarkeit von Differentialgleichungen treffen zu können. Statt konkrete Differentialgleichungen wie $f'' + f = g$ für gegebenes g händisch zu lösen, war man an der Frage interessiert, welche Eigenschaften die Differentialgleichung bzw. die rechte Seite g haben muss, damit eine Lösung existiert. Der wesentliche Schritt war dabei, Funktionen als Punkte in einem Vektorraum aufzufassen, auf dem die Abbildung $D : f \mapsto f'' + f$ einen linearen *Differentialoperator* definiert. Zum Vergleich: Ein ähnlicher Übergang von konkreten linearen Gleichungssystemen zu der abstrakten linearen Gleichung $Ax = b$ für eine Matrix A und einen Vektor b bildet den Grundstein der linearen Algebra. Man war nun auf der Suche nach Eigenschaften von D, die analog zur Injektivität und Surjektivität von A oder der Tatsache, dass 0 kein Eigenwert von A ist, die Lösbarkeit von $Df = g$ garantieren. Die wesentliche Schwierigkeit dabei ist, dass viele Kernaussagen der linearen Algebra darauf beruhen, dass die betrachteten Vektorräume endlichdimensional sind (etwa indem der Dimensionssatz verwendet wird). Dies ist aber für Räume von Funktionen in der Regel nicht mehr der Fall, und es wird notwendig, zusammen mit den algebraischen Begriffen auch topologische Begriffe wie Grenzwerte und Kompaktheit zu berücksichtigen. Ein Leitfaden dieser Vorlesung ist es herauszuarbeiten, welche algebraischen, metrischen, topologischen und geometrischen Eigenschaften als Ersatz für die fehlende Endlichdimensionalität dienen können, und wie diese in die einzelnen Resultate eingehen. Dass diese Kombination zu äußerst reichhaltigen Strukturen führt, macht den Reiz der Funktionalanalysis aus und führte dazu, dass sie eine wesentliche Grundlage der modernen angewandten Mathematik bildet, von der Theorie und Numerik von Differentialgleichungen über Optimierung und Wahrscheinlichkeitstheorie bis zu medizinischer Bildgebung und mathematischer Bildverarbeitung.

Dieses Skript folgt [Clason 2019], welches vor allem auf den Werken [Alt 2012; Brokate 2013; Kaballo 2018; Wachsmuth 2013; Werner 2018] basiert; letztere Werke können auch als Ergänzung und Vertiefung dienen.

Teil I

TOPOLOGISCHE GRUNDLAGEN

1 METRISCHE RÄUME

Wir fassen zunächst die grundlegenden topologischen Strukturen zusammen, die in der Funktionalanalysis wichtig sind. Die Kernbegriffe sollten aus der Analysis bekannt sein; eine ausführlichere Darstellung sowie Beweise findet man in den Standard-Lehrbüchern wie [Forster 2017, Kapitel I.1] oder [Rudin 2008, Kapitel 2].

Definition 1.1. Sei X eine Menge. Eine *Metrik* auf X ist eine Abbildung $d : X \times X \to \mathbb{R}$, die für alle $x, y, z \in X$ die folgenden Eigenschaften erfüllt:

(i) $d(x, y) = 0$ genau dann, wenn $x = y$ *(Nichtdegeneriertheit)*;

(ii) $d(x, y) = d(y, x)$ *(Symmetrie)*;

(iii) $d(x, z) \le d(x, y) + d(y, z)$ *(Dreiecksungleichung)*.

In diesem Fall heißt das Paar (X, d) *metrischer Raum*. Ist aus dem Kontext offensichtlich, welche Metrik verwendet wird, bezeichnen wir den metrischen Raum auch kurz mit X.

Eine Metrik ist die mathematische Formalisierung des intuitiven Begriffs des Abstands. (Beachten Sie, dass wir noch keine algebraische Struktur – und damit die Möglichkeit, die Differenz von Elementen zu betrachten – gefordert haben!) Aus den Eigenschaften folgt direkt, dass eine Metrik stets nicht-negativ ist: Für alle $x, y \in X$ gilt

$$2d(x, y) = d(x, y) + d(x, y) = d(x, y) + d(y, x) \ge d(x, x) = 0.$$

Beispiel 1.2. Kanonische Beispiele für Metriken sind

(i) die *euklidische Metrik*: $X = \mathbb{R}^n$ oder $X = \mathbb{C}^n$ und

$$d(x, y) := \left(\sum_{i=1}^{n} |x_i - y_i|^2\right)^{\frac{1}{2}};$$

(ii) die *Relativmetrik*: ist (X, d) ein metrischer Raum und $A \subset X$, dann ist auch $(A, d|_{A\times A})$ ein metrischer Raum, wobei $d|_{A\times A}$ die Einschränkung von d auf $A \times A$ bezeichnet;

(iii) die *Produktmetrik*: sind (X, d_X) und (Y, d_Y) metrische Räume, dann ist auch $(X \times Y, d_{X\times Y})$ ein metrischer Raum für

$$d_{X\times Y}((x_1, y_1), (x_2, y_2)) := d_X(x_1, x_2) + d_Y(y_1, y_2),$$

und ebenso für

$$d_{X\times Y}((x_1, y_1), (x_2, y_2)) := \max\{d_X(x_1, x_2), d_Y(y_1, y_2)\};$$

(iv) die *diskrete Metrik*: X ist eine beliebige Menge und

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

Im folgenden sei (X, d) stets ein metrischer Raum. Wir definieren nun für $x \in X$ und $r > 0$

(i) die *abgeschlossene Kugel* $B_r(x) := \{y \in X : d(x, y) \leq r\}$,

(ii) die *offene Kugel* $U_r(x) := \{y \in X : d(x, y) < r\}$,

um x mit Radius r. Mit ihrer Hilfe definieren wir nun die folgenden topologischen Grundbegriffe.

Definition 1.3. Eine Menge $U \subset X$ heißt

(i) *offen*, falls für alle $x \in U$ ein $\varepsilon > 0$ mit $U_\varepsilon(x) \subset U$ existiert;

(ii) *Umgebung* von $x \in U$, falls eine offene Menge O mit $x \in O \subset U$ existiert;

(iii) *Umgebung* von $A \subset U$, falls U Umgebung aller $x \in A$ ist.

Eine Menge $C \subset X$ heißt *abgeschlossen*, falls das Komplement $X \setminus C$ offen ist.

Aus der Definition folgt, dass offene und abgeschlossene Kugeln tatsächlich offen respektive abgeschlossen sind. Weiterhin sind der Schnitt endlich vieler offener Mengen sowie die Vereinigung beliebiger (auch unendlich vieler) offener Mengen offen. Wir nennen zwei metrische Räume (X, d_1) und (X, d_2) *äquivalent*, wenn sie die selben offenen Mengen besitzen (etwa die beiden Definitionen in Beispiel 1.2 (iii)). Die Menge aller offenen Teilmengen bezeichnet man als *Topologie* auf X.

Offensichtlich sind sowohl X als auch die leere Menge $\emptyset$ sowohl offen als auch abgeschlossen; andere Mengen können weder offen noch abgeschlossen sein. In diesem Fall können wir aus ihnen offene und abgeschlossene Mengen erzeugen.

Definition 1.4. Für $A \subset X$ definieren wir

(i) das *Innere* $\operatorname{int} A := \bigcup_{\{U \subset A : U \text{ offen}\}} U$;

(ii) den *Abschluss* $\operatorname{cl} A := \bigcap_{\{C \supset A : C \text{ abgeschlossen}\}} C$.

Man verwendet auch oft die Bezeichnungen $A^o := \operatorname{int} A$ und $\overline{A} := \operatorname{cl} A$.

Ein $x \in \operatorname{int} A$ heißt *innerer Punkt* von A; in anderen Worten, $x \in A$ ist ein innerer Punkt genau dann, wenn ein $\varepsilon > 0$ existiert mit $U_\varepsilon(x) \subset A$. Gilt $\operatorname{cl} A = X$, so heißt A *dicht* in X. Existiert eine Menge $A \subset X$, die abzählbar und dicht in X ist, so heißt X *separabel*.

Aus der Definition folgt, dass das Innere von A stets offen und der Abschluss von A stets abgeschlossen ist. Insbesondere ist A offen genau dann, wenn $A = \operatorname{int} A$ gilt, und abgeschlossen genau dann, wenn $A = \operatorname{cl} A$ gilt.

Schließlich nennen wir eine Menge $A \subset X$ *beschränkt*, falls der *Durchmesser*

$$\operatorname{diam}(A) := \sup_{x,y \in A} d(x, y)$$

endlich ist.

Eine Metrik erlaubt es, Konvergenz von Folgen zu definieren.

Definition 1.5. Eine Folge $\{x_n\}_{n\in\mathbb{N}} \subset X$ *konvergiert in* X gegen den *Grenzwert* $x \in X$, falls eine der folgenden äquivalenten Eigenschaften gilt:

(i) Für alle $\varepsilon > 0$ existiert ein $N \in \mathbb{N}$ mit $d(x_n, x) \leq \varepsilon$ für alle $n \geq N$;

(ii) Für jede Umgebung U von x existiert ein $N \in \mathbb{N}$ mit $x_n \in U$ für alle $n \geq N$.

In diesem Fall schreiben wir $x_n \to_{(X,d)} x$ bzw. kurz $x_n \to x$, falls offensichtlich ist, welcher metrische Raum gemeint ist.

Die Äquivalenz folgt dabei direkt aus der Definition von Umgebungen. Aus der Definition folgt auch die Eindeutigkeit des Grenzwertes. Weiterhin gilt für $A \subset X$, dass

$$\operatorname{cl} A = \{x \in X : \text{es existiert } \{x_n\}_{n\in\mathbb{N}} \subset A \text{ mit } x_n \to x\} \tag{1.1}$$

ist. Insbesondere ist A abgeschlossen genau dann, wenn der Grenzwert jeder konvergenten Folge $\{x_n\}_{n\in\mathbb{N}} \subset X$ in A liegt. Weiterhin sind zwei metrische Räume (X, d_1) und (X, d_2) äquivalent genau dann, wenn sie die selben konvergenten Folgen (mit übereinstimmendem Grenzwert) besitzen.

Definition 1.6. Sei $\{x_n\}_{n\in\mathbb{N}} \subset X$ eine Folge.

(i) Ist $\{n_k\}_{k\in\mathbb{N}} \subset \mathbb{N}$ eine streng monoton wachsende Folge, dann ist $\{x_{n_k}\}_{k\in\mathbb{N}}$ eine Folge, genannt *Teilfolge* von $\{x_n\}_{n\in\mathbb{N}}$.

(ii) Hat $\{x_n\}_{n\in\mathbb{N}}$ eine konvergente Teilfolge mit Grenzwert $x \in X$, so heißt x *Häufungspunkt* von $\{x_n\}_{n\in\mathbb{N}}$.

(iii) Existiert für alle $\varepsilon > 0$ ein $N \in \mathbb{N}$ mit $d(x_m, x_n) \leq \varepsilon$ für alle $m, n \geq N$, so heißt $\{x_n\}_{n\in\mathbb{N}}$ *Cauchy-Folge.*

Genau wie für reelle Folgen zeigt man, dass jede Cauchy-Folge maximal einen Häufungspunkt besitzt. Jede konvergente Folge ist also eine Cauchy-Folge; die Umkehrung gilt im Allgemeinen nicht. Metrische Räume, in denen jede Cauchy-Folge konvergiert, heißen *vollständig*; wie wir in Teil II sehen werden, ist dies eine fundamentale Eigenschaft mit weitreichenden Folgen. Beispielsweise sind $\mathbb{R}^n$ oder $\mathbb{C}^n$ sowohl versehen mit der euklidischen als auch mit den Produktmetriken aus Beispiel 1.2 (iii) vollständig. Außerdem bilden abgeschlossene Mengen in vollständigen metrischen Räumen (versehen mit der Relativmetrik) wieder vollständige metrische Räume. Beachten Sie, dass äquivalente metrische Räume zwar die selben konvergenten Folgen, nicht aber die selben Cauchy-Folgen besitzen müssen – Äquivalenz erhält also nicht die Vollständigkeit! (Vollständigkeit ist deshalb eine *topologische* und keine *metrische* Eigenschaft.)

Analog zur Konvergenz von Folgen lässt sich auch die Stetigkeit von Abbildungen auf metrische Räume übertragen.

Definition 1.7. Seien (X, d_X) und (Y, d_Y) metrische Räume. Eine Abbildung $f : X \to Y$ heißt *stetig* in $x \in X$, falls eine der folgenden äquivalenten Eigenschaften gilt:

(i) Für alle $\varepsilon > 0$ existiert ein $\delta > 0$ mit $f(B_\delta(x)) \subset B_\varepsilon(f(x))$ (äquivalent: $f(U_\delta(x)) \subset U_\varepsilon(f(x))$).

(ii) Für jede Umgebung V von $f(x)$ existiert eine Umgebung U von x mit $f(U) \subset V$.

(iii) Für jede Folge $\{x_n\}_{n\in\mathbb{N}} \subset X$ mit $x_n \to x$ in X gilt $f(x_n) \to f(x)$ in Y.

Wir nennen f *stetig* auf X, wenn f stetig in x für alle $x \in X$ ist, und *beschränkt*, wenn $\sup_{x,y\in X} d_Y(f(x), f(y)) < \infty$ ist.

Diese Definition formalisiert die intuitive Vorstellung, dass eine stetige Funktion f Punkte in der Nähe von x auf Punkte in der Nähe von $f(x)$ abbildet. Die Äquivalenz folgt dabei wieder aus den Definitionen von Umgebung und Konvergenz in metrischen Räumen. Eine alternative Charakterisierung, die später nützlich sein wird, ist die folgende.

Satz 1.8. *Seien (X, d_X) und (Y, d_Y) metrische Räume. Eine Abbildung $f : X \to Y$ ist stetig genau dann, wenn für alle offenen Mengen $V \subset Y$ die* Urbilder

$$f^{-1}(V) := \{x \in X : f(x) \in V\}$$

offen sind.

Die Bilder offener Mengen müssen dagegen *nicht* offen sein! Durch Komplementbildung erhält man daraus

Folgerung 1.9. *$f : X \to Y$ ist stetig genau dann, wenn für alle abgeschlossenen Mengen $V \subset Y$ die Urbilder $f^{-1}(V)$ abgeschlossen sind.*

Ist X ein metrischer Raum, so ist der Raum aller stetigen, beschränkten, reellwertigen (oder komplexwertigen) Funktionen

$$C_b(X) := \{f : X \to \mathbb{R} : f \text{ stetig und beschränkt}\}$$

zusammen mit

$$d(f, g) := \sup_{x \in X} |f(x) - g(x)| \tag{1.2}$$

ein vollständiger metrischer Raum. Definition 1.5 entspricht dann genau der gleichmäßigen Konvergenz von Funktionenfolgen.

Metrische Räume stellen nicht den allgemeinsten Rahmen für die oben eingeführten Begriffe dar. Anstatt offene Mengen mit Hilfe offener Kugeln (d. h. über Metriken) zu definieren, kann man diese direkt axiomatisch einführen: Man definiert die Topologie τ auf X als ein System von Teilmengen von X, das X und $\emptyset$ enthält und abgeschlossen bezüglich Vereinigung und endlichen Schnitten ist; das Paar (X, τ) heißt dann *topologischer Raum*. In topologischen Räumen definiert man Konvergenz von Folgen und Stetigkeit direkt über die Eigenschaft (ii) in Definition 1.5 beziehungsweise Definition 1.7; allerdings sind Definition 1.7 (ii) und (iii) nicht mehr unbedingt äquivalent (man bezeichnet Letztere dann als *Folgenstetigkeit*); Details findet man z. B. in [Werner 2009, Kapitel 1] oder [Conway 2014]. Topologische Räume tauchen zum Beispiel auf, wenn man die punktweise (nicht gleichmäßige) Konvergenz von Funktionenfolgen untersuchen möchte, welche im Allgemeinen nicht durch eine Metrik ausgedrückt werden kann.

2 KOMPAKTE MENGEN

Eine fundamentale metrische Eigenschaft ist die Kompaktheit; salopp gesprochen werden wir sehen, dass stetige Funktionen auf kompakten Mengen ähnlich gute Eigenschaften besitzen wie Funktionen auf endlichen Mengen.

Sei im folgenden wieder (X, d) ein metrischer Raum. Wir definieren zunächst mehrere verwandte Kompaktheitsbegriffe.

Definition 2.1. Eine Menge $K \subset X$ heißt

(i) *kompakt*, falls jede offene Überdeckung von K eine endliche Teilüberdeckung besitzt, d. h. falls für jede Familie $\{U_i\}_{i\in I}$ von Mengen mit $U_i \subset K$ offen und $K \subset \bigcup_{i\in I} U_i$ eine *endliche* Teilmenge $J \subset I$ existiert mit $K \subset \bigcup_{j\in J} U_j$;

(ii) *folgenkompakt*, falls jede Folge $\{x_n\}_{n\in\mathbb{N}} \subset K$ eine konvergente Teilfolge $\{x_{n_k}\}_{k\in\mathbb{N}}$ mit Grenzwert $x \in K$ besitzt;

(iii) *präkompakt* (oder *totalbeschränkt*), falls für alle $\varepsilon > 0$ eine endliche Überdeckung mit offenen Kugeln mit Radius ε existiert, d. h. $N = N(\varepsilon) \in \mathbb{N}$ und $x_1, \dots, x_N \in K$ existieren, so dass $K \subset \bigcup_{n=1}^{N} U_\varepsilon(x_i)$ gilt.

Ist (K, d) ein metrischer Raum und K kompakt, spricht man auch von einem *kompakten Raum*.

Definition (i) ist technisch, deutet aber bereits an, weshalb Kompaktheit ein guter Ersatz für Endlichkeit ist: es genügt, eine Eigenschaft durch Betrachtung endlich vieler offener Umgebungen zu verifizieren. Definition (ii) ist die in der Praxis nützlichste Eigenschaft, da sie erlaubt, aus *jeder* Folge einen Häufungspunkt zu extrahieren. Die alternative Bezeichnung „totalbeschränkt“ in (iii) erklärt sich dadurch, dass jede totalbeschränkte Menge beschränkt ist wegen $\operatorname{diam}(K) \leq 2\varepsilon N(\varepsilon) < \infty$ für beliebiges $\varepsilon > 0$.

In metrischen Räumen sind alle drei Eigenschaften äquivalent.[1]

Satz 2.2. *Für $K \subset X$ sind äquivalent:*

(i) K ist kompakt,

[1] Dies ist nicht mehr unbedingt der Fall in topologischen Räumen; siehe z. B. [Werner 2009, Seite 29].

(ii) *K ist folgenkompakt,*

(iii) *K ist vollständig (bezüglich der Relativmetrik) und präkompakt.*

Beweis. *(i) $\Rightarrow$ (ii):* Wir führen einen Widerspruchsbeweis. Angenommen, es existiert eine Folge $\{x_n\}_{n\in\mathbb{N}} \subset K$ ohne Häufungspunkt. Für alle $x \in K$ existiert dann ein $r_x > 0$, so dass $U_{r_x}(x)$ nur endlich viele Folgenglieder x_n enthält (sonst könnten wir eine Teilfolge bilden, die nach Definition 1.5 (ii) gegen x konvergiert, im Widerspruch zur Annahme). Die Familie $\{U_{r_x}(x)\}_{x\in K}$ bildet aber eine offene Überdeckung von K, und aus der Kompaktheit von K folgt die Existenz einer endlichen Teilüberdeckung $\{U_{r_{\tilde{x}_i}}(\tilde{x}_i)\}_{i=1,\dots,N}$ von K. Da jede dieser Mengen nur endlich viele Folgenglieder enthält, gilt das auch für ihre (endliche) Vereinigung. Also kann K selber nur endlich viele Folgenglieder enthalten, im Widerspruch zu $\{x_n\}_{n\in\mathbb{N}} \subset K$.

(ii) $\Rightarrow$ (iii): Da jede Cauchy-Folge mit Häufungspunkt konvergiert und nach Annahme jede Folge einen Häufungspunkt besitzt, ist $(K, d|_{K\times K})$ nach Definition vollständig. Angenommen, K ist nicht präkompakt. Dann existiert ein $\varepsilon > 0$, so dass K nicht mit endlich vielen ε-Kugeln überdeckt werden kann. Wir können also rekursiv eine Folge $\{x_n\}_{n\in\mathbb{N}} \subset K$ konstruieren, in dem wir $x_1 \in K$ beliebig und

$$x_{n+1} \in K \setminus \bigcup_{i=1}^{n} U_\varepsilon(x_i), \qquad n \in \mathbb{N},$$

auswählen. (Diese Konstruktion ist möglich, da die Menge auf der rechten Seite nach Annahme nie leer wird.) Damit enthält jede ε-Kugel höchstens ein Folgeglied, so dass $\{x_n\}_{n\in\mathbb{N}}$ keinen Häufungspunkt haben kann im Widerspruch zur Folgenkompaktheit.

(iii) $\Rightarrow$ (i): Dies ist die schwierigste Richtung. Wir führen wieder einen Widerspruchsbeweis. Es sei $\{U_i\}_{i\in I}$ eine offene Überdeckung von K. Wir definieren nun das Mengensystem $\mathcal{B}$ aller Mengen, die nur mit unendlich vielen U_i überdeckt werden können, und zeigen, dass die Annahme $K \in \mathcal{B}$ zu einem Widerspruch führt. Dafür konstruieren wir per Induktion offene Kugeln $B_n := U_{2^{-n}}(x_n)$ mit $B_n \cap B_{n-1} \neq \emptyset$ und $B_n \in \mathcal{B}$. Für $n = 1$ verwenden wir die Präkompaktheit von K, um endlich viele offene Kugeln mit Radius $\varepsilon = \frac{1}{2}$ zu wählen, deren Vereinigung K überdeckt. Nach Annahme muss es darunter mindestens eine Kugel in $\mathcal{B}$ geben (denn sonst wäre $K \notin \mathcal{B}$); wir bezeichnen diese mit $B_1 = U_{\frac{1}{2}}(x_1)$. Sei nun $B_{n-1} \in \mathcal{B}$ entsprechend gewählt. Dann existieren wiederum endlich viele offene Kugeln mit Radius 2^{-n}, deren Vereinigung K überdeckt und mit B_{n-1} nichtleeren Durchschnitt habi. Wegen $B_{n-1} \in \mathcal{B}$ ist nun mindestens eine dieser Kugeln in $\mathcal{B}$, die wir mit $B_n = U_{2^{-n}}(x_n)$ bezeichnen. Wir zeigen nun, dass dadurch eine Cauchy-Folge $\{x_n\}_{n\in\mathbb{N}} \subset K$ definiert wird.Für alle $n \in \mathbb{N}$ existiert nach Annahme ein $x \in B_n \cap B_{n+1}$, woraus folgt

$$d(x_n, x_{n+1}) \leq d(x_n, x) + d(x, x_{n+1}) < 2^{-n} + 2^{-(n+1)} < 2^{-n+1}$$

und damit

$$d(x_n, x_m) < 2^{-n+2} \qquad \text{für alle } n, m \text{ mit } m \geq n.$$

Da $(K, d|_{K\times K})$ vollständig ist, konvergiert $x_n \longrightarrow x \in K$. Aus der Überdeckungseigenschaft folgt, dass $x \in U_j$ für ein $j \in I$ ist. Da U_j offen ist, existiert ein $\varepsilon > 0$ mit $x \in U_\varepsilon(x) \subset U_j$; für n groß genug gilt dann

$$B_n = U_{2^{-n}}(x_n) \subset U_\varepsilon(x) \subset U_j,$$

im Widerspruch zu $B_n \in \mathcal{B}$. □

Die Richtung (iii) ⇒ (i) zeigt insbesondere, dass aus jeder unendlichen präkompakten Menge eine Cauchy-Folge extrahiert werden kann.

Eigenschaft (iii) kann man als stärkere Version der Abgeschlossenheit und Beschränktheit interpretieren. Tatsächlich folgen beide Eigenschaften aus der Kompaktheit.

Folgerung 2.3. *Ist $K \subset X$ kompakt, dann ist K beschränkt und abgeschlossen.*

Beweis. Die Menge K ist nach Satz 2.2 (iii) präkompakt und daher beschränkt. Weiterhin hat nach Satz 2.2 (ii) jede Folge einen Häufungspunkt in K; also liegt jeder Grenzwert einer konvergenten Folge (der dann ja der einzige Häufungspunkt ist) in K, und damit ist K abgeschlossen. □

Für $X = \mathbb{R}^n$ gilt auch die Umkehrung. Wir verwenden dafür das folgende Lemma.

Lemma 2.4. *Ist $K \subset X$ kompakt und $C \subset K$ abgeschlossen, dann ist auch C kompakt.*

Beweis. Sei $\{U_i\}_{i\in I}$ eine offene Überdeckung von C. Da C abgeschlossen ist, ist $X \setminus C$ offen, und damit ist $\{U_i\}_{i\in I} \cup (X \setminus C)$ eine offene Überdeckung von K. Da K kompakt ist, existiert eine endliche Teilüberdeckung $\{U_n\}_{n=1,\dots,N} \cup (X \setminus C)$ von K und damit auch von $C \subset K$. Wegen $C \cap (X \setminus C) = \emptyset$ ist $X \setminus C$ sicher kein notwendiger Teil einer Überdeckung von C, also ist insbesondere $\{U_n\}_{n=1,\dots,N}$ eine endliche Teilüberdeckung von C. Damit ist C kompakt. □

Um das gewünschte Resultat zu zeigen, müssen wir $\mathbb{R}^n$ mit einer Standardmetrik, etwa Beispiel 1.2 (i), versehen.

Satz 2.5 (Heine–Borel). *Eine Teilmenge von $\mathbb{R}^n$ ist bezüglich der euklidischen Metrik kompakt genau dann, wenn sie beschränkt und abgeschlossen ist.*

Beweis. Wegen Folgerung 2.3 ist nur zu zeigen, dass beschränkte und abgeschlossene Teilmengen des $\mathbb{R}^n$ kompakt sind. Sei daher $C \subset \mathbb{R}^n$ beschränkt und abgeschlossen. Wir verwenden, dass eine Menge bezüglich der euklidischen Metrik genau dann beschränkt ist, wenn ihre Elemente komponentenweise beschränkt sind. Es existiert also ein $M > 0$, so dass

$$C \subset \prod_{i=1}^{n} [-M, M] =: K$$

gilt. Aus der Analysis ist bekannt, dass das abgeschlossene Intervall $[-M, M]$ kompakt ist. Wir zeigen nun, dass daraus die (Folgen-)Kompaktheit des n-dimensionalen Quaders K folgt, indem wir ein Diagonalfolgenargument anwenden. Dies erfordert eine spezielle Notation. Sei $\{x_k\}_{k\in\mathbb{N}} \subset K$ eine Folge, wobei wir für $x_k \in \mathbb{R}^n$ schreiben $x_k = (x_k^1, \dots, x_k^n)$. Dann gilt $\{x_k^1\}_{k\in\mathbb{N}} \subset [-M, M]$, es existiert also eine konvergente Teilfolge, die wir mit $\{x_k^1\}_{k\in\mathbb{N}_1}$ für eine unendliche Teilmenge $\mathbb{N}_1 \subset \mathbb{N}$ bezeichnen, deren Grenzwert $x^1 \in [-M, M]$ liegt. Wir betrachten nun die Folge $\{x_k^2\}_{k\in\mathbb{N}_1} \subset [-M, M]$, die wiederum eine Teilfolge $\{x_k^2\}_{k\in\mathbb{N}_2}$ für $\mathbb{N}_2 \subset \mathbb{N}_1$ mit Grenzwert $x^2 \in [-M, M]$ besitzt. Schlussendlich erhalten wir daraus ein $\mathbb{N}_n \subset \dots \subset \mathbb{N}_1 \subset \mathbb{N}$ so dass $\{x_k^n\}_{k\in\mathbb{N}_n}$ konvergiert gegen $x^n \in [-M, M]$. Da jede Teilfolge einer konvergenten Folge gegen den selben Grenzwert konvergiert, folgt daraus insbesondere die Existenz einer Teilfolge $\{x_k\}_{k\in\mathbb{N}_n}$ mit Grenzwert $x := (x^1, \dots, x^n) \in \prod_{i=1}^n [-M, M]$, zunächst bezüglich der komponentenweisen Konvergenz. Da Folgen im $\mathbb{R}^n$ bezüglich der euklidischen Metrik genau dann konvergieren, wenn sie komponentenweise konvergieren, folgt damit die Kompaktheit von K. Die Kompaktheit von C folgt nun mit Lemma 2.4. □

Der Satz von Heine–Borel beruht also wesentlich auf der Äquivalenz von metrischer und komponentenweiser Konvergenz (bzw. Beschränktheit), und ist daher in unendlichdimensionalen metrischen Räumen im Allgemeinen falsch. (Wir werden später ein Gegenbeispiel sehen.) Darin äußert sich eine der wesentlichen Komplikationen in der Funktionalanalysis gegenüber der linearen Algebra.[2]

Oft wird Satz 2.5 direkt für Folgenkompaktheit formuliert.

Folgerung 2.6 (Satz von Bolzano–Weierstraß). *Jede (komponentenweise) beschränkte Folge in $\mathbb{R}^n$ besitzt eine (komponentenweise) konvergente Teilfolge.*

Ist in Satz 2.2 (iii) die Menge K selber nicht vollständig, so gilt immer noch eine schwächere Version dieses Satzes. Für den Beweis verwenden wir die folgende Variante von Lemma 2.4.

Lemma 2.7. *Ist $K \subset X$ präkompakt und $C \subset K$, dann ist auch C präkompakt.*

[2] Auch im $\mathbb{R}^n$ lassen sich Metriken konstruieren, für die eine dieser Äquivalenzen und damit auch die Aussage nicht gilt.

Beweis. Sei $\varepsilon > 0$ beliebig. Da K präkompakt ist, existiert eine Überdeckung mit endlich vielen Kugeln $U_{\varepsilon/2}(\tilde{x}_i)$ mit $\tilde{x}_i \in K$, $i = 1, \dots, \tilde{N}$. Diese Kugeln können wegen $C \subset K \subset \bigcup_{i=1,\dots,\tilde{N}} U_{\varepsilon/2}(\tilde{x}_i)$ nicht alle leeren Schnitt mit C haben. Falls nötig durch Umnummerieren finden wir daher

$$x_i \in U_{\varepsilon/2}(\tilde{x}_i) \cap C, \qquad i = 1, \dots, N \le \tilde{N}.$$

Sei nun $x \in C \subset K$ beliebig. Dann gilt $x \in U_{\varepsilon/2}(\tilde{x}_j)$ für ein $1 \le j \le N$, woraus folgt

$$d(x, x_j) \le d(x, \tilde{x}_j) + d(\tilde{x}_j, x_j) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

d. h. $x \in U_\varepsilon(x_j)$. Also ist $\{U_\varepsilon(x_j)\}_{j=1,\dots,N}$ die gesuchte endliche Überdeckung von C. □

Satz 2.8. *Sei (X, d) vollständig und $A \subset X$. Dann sind äquivalent:*

(i) A ist präkompakt;

(ii) A ist relativkompakt*, d. h.* $\operatorname{cl} A$ *ist kompakt;*

(iii) jede Folge in A hat eine konvergente Teilfolge (deren Grenzwert nicht in A liegen muss).

Beweis. *(iii) ⇒ (ii):* Sei $\{x_n\}_{n\in\mathbb{N}}$ eine Folge in $\operatorname{cl} A$. Wegen (1.1) existiert dann für jedes x_n eine Folge $\{x_{n,k}\}_{k\in\mathbb{N}} \subset A$ mit $x_{n,k} \to x_n$ für $k \to \infty$, d. h. für alle $\varepsilon > 0$ existiert ein $N_n \in \mathbb{N}$ und $\tilde{x}_n := x_{n,N_n}$ mit $d(x_n, \tilde{x}_n) \le \varepsilon/2$. Betrachte nun die Folge $\{\tilde{x}_n\}_{n\in\mathbb{N}} \subset A$, die nach Annahme eine konvergente Teilfolge $\{\tilde{x}_{n_k}\}_{k\in\mathbb{N}} \subset A$ mit Grenzwert $x \in \operatorname{cl} A$ besitzt (da der Grenzwert jeder konvergenten Folge in A stets in $\operatorname{cl} A$ liegt). Für $\varepsilon > 0$ existiert dann ein $N \in \mathbb{N}$ mit $d(x, \tilde{x}_{n_k}) \le \varepsilon/2$ für alle $k \ge N$, woraus folgt

$$d(x, x_{n_k}) \le d(x, \tilde{x}_{n_k}) + d(\tilde{x}_{n_k}, x_{n_k}) \le \varepsilon \qquad \text{für alle } k \ge N,$$

d. h. die Teilfolge $\{x_{n_k}\}_{k\in\mathbb{N}}$ konvergiert gegen $x \in \operatorname{cl} A$ und damit ist $\operatorname{cl} A$ kompakt.

(ii) ⇒ (i): Ist $\operatorname{cl} A$ kompakt, so ist $\operatorname{cl} A$ insbesondere präkompakt. Nach Lemma 2.7 gilt dies dann auch für $A \subset \operatorname{cl} A$.

(i) ⇒ (iii): Sei $\{x_n\}_{n\in\mathbb{N}} \subset A$ eine Folge. Ist A präkompakt, so nach Lemma 2.7 auch die Teilmenge aller Folgenglieder. Aufgrund der Bemerkung nach Satz 2.2 enthält daher $\{x_n\}_{n\in\mathbb{N}}$ eine Cauchy-Teilfolge, die aufgrund der Vollständigkeit von (X, d) konvergiert. □

Wir betrachten nun stetige Funktionen auf kompakten Mengen.

Satz 2.9. *Seien (X, d_X) und (Y, d_Y) metrische Räume und sei $f : X \to Y$ stetig. Ist $K \subset X$ kompakt, so ist auch $f(K) \subset Y$ kompakt.*

Beweis. Sei $\{U_i\}_{i\in I}$ eine offene Überdeckung von $f(K)$. Da f stetig ist, sind nach Satz 1.8 die Urbilder $\{f^{-1}(U_i)\}_{i\in I}$ wieder offen und müssen eine Überdeckung von K bilden (sonst gäbe es ein $x \in K$ mit $f(x) \notin f(K)$, was nach Definition nicht möglich ist). Aus der Kompaktheit von K folgt nun die Existenz einer endlichen Teilüberdeckung $\{f^{-1}(U_i)\}_{i\in J}$, und $\{U_i\}_{i\in J}$ ist die gesuchte endliche Teilüberdeckung von $f(K)$. □

Als Folgerung erhält man die in der Optimierung nützliche Tatsache, dass stetige reellwertige Funktionen auf kompakten Mengen stets ihr Maximum und Minimum annehmen.

Folgerung 2.10 (Satz von Weierstraß). *Sei (K, d) ein kompakter Raum und $f : K \to \mathbb{R}$ stetig. Dann existieren $a, b \in K$ mit $f(a) \leq f(x) \leq f(b)$ für alle $x \in K$.*

Beweis. Nach Satz 2.9 ist $f(K) \subset \mathbb{R}$ kompakt und daher beschränkt und abgeschlossen. Wegen der Beschränktheit sind $\alpha := \inf f(K)$ und $\beta := \sup f(K)$ endlich. Aus den Eigenschaften des Infimums und Supremums sowie der Definition des Bildes folgt dann die Existenz von Folgen $\{x_n\}_{n\in\mathbb{N}} \subset K$ mit $f(x_n) \to \alpha$ und $\{y_n\}_{n\in\mathbb{N}} \subset K$ mit $f(y_n) \to \beta$, und die Abgeschlossenheit von $f(K)$ liefert $\alpha, \beta \in f(K)$, woraus die Behauptung folgt. □

Insbesondere sind stetige Funktionen auf kompakten Mengen stets beschränkt; in kompakten metrischen Räumen (K, d) gilt daher

$$C_b(K) = C(K) := \{f : K \to \mathbb{R} : f \text{ stetig}\},$$

welcher versehen mit der Metrik (1.2) ein vollständiger metrischer Raum ist.

Bisher ist noch nicht klar, ob überhaupt kompakte Teilmengen in unendlichdimensionalen Räumen existieren. Dies werden wir jetzt für $C(K)$ zeigen. Wir benötigen dafür das folgende Lemma.

Lemma 2.11. *Jeder kompakte Raum ist separabel.*

Beweis. Sei (K, d) kompakt. Wir müssen zeigen, dass eine abzählbare dichte Teilmenge existiert. Dafür verwenden wir, dass K präkompakt ist, für alle $\varepsilon > 0$ also eine endliche Überdeckung mit offenen ε-Kugeln existiert. Bezeichne für $n \in \mathbb{N}$ die Menge aller Mittelpunkte dieser Kugeln für $\varepsilon = \frac{1}{n}$ mit P_n. Da alle P_n endlich sind, ist $P := \bigcup_{n\in\mathbb{N}} P_n \subset K$ abzählbar; dies ist unser Kandidat. Zuerst folgt aus $P \subset K$ zusammen mit der Abgeschlossenheit von K als kompakter Raum auch $\operatorname{cl} P \subset K$. Sei umgekehrt $x \in K$ beliebig. Aus der Überdeckungseigenschaft erhalten wir dann für alle $n \in \mathbb{N}$ ein $x_n \in P_n$ mit $x \in U_{1/n}(x_n)$, d. h. $d(x_n, x) < \frac{1}{n}$. Daraus folgt sofort $\{x_n\}_{n\in\mathbb{N}} \subset P$ sowie $x_n \to x$, d. h. $x \in \operatorname{cl} P$ und damit $K \subset \operatorname{cl} P$. Also ist P dicht in K. □

Wir geben nun eine Charakterisierung der Präkompaktheit in $C(K)$ an.

Satz 2.12 (Arzelà–Ascoli). *Sei (K, d) ein kompakter Raum und $A \subset C(K)$. Ist A*

(i) punktweise beschränkt, d. h. für alle $x \in K$ ist die Menge $\{f(x) : f \in A\} \subset \mathbb{R}$ beschränkt,

(ii) gleichgradig stetig, *d. h. für alle $\varepsilon > 0$ existiert ein $\delta > 0$ so dass für alle $f \in A$ und $x, y \in K$ gilt $|f(x) - f(y)| \leq \varepsilon$ falls $d(x, y) \leq \delta$,*

so ist A präkompakt.

Beweis. Wir verwenden Satz 2.8 (iii) und konstruieren für eine gegebene Folge $\{f_n\}_{n\in\mathbb{N}} \subset A$ eine konvergente Teilfolge über ein Diagonalfolgenargument. Nach Lemma 2.11 existiert eine abzählbare dichte Teilmenge $\{x_1, x_2, \dots\} =: X \subset K$. Wir setzen $f_n^0 := f_n$ und betrachten die Folge $\{f_n^0(x_1)\}_{n\in\mathbb{N}} \subset \mathbb{R}$. Diese ist nach Annahme (i) beschränkt und hat daher nach dem Satz von Bolzano–Weierstraß (Folgerung 2.6) eine konvergente Teilfolge, die wir mit $\{f_n^1(x_1)\}_{n\in\mathbb{N}}$ bezeichnen. So fortfahrend finden wir also für jedes $j \in \mathbb{N}$ eine Teilfolge $\{f_n^j\}_{n\in\mathbb{N}}$, so dass $\{f_n^j(x_k)\}_{n\in\mathbb{N}}$ für alle $k \leq j$ konvergiert. Aus dieser Folge von Teilfolgen bilden wir nun die Diagonalfolge durch $f_n^* := f_n^n \in A$. Diese ist eine Teilfolge von $\{f_n\}_{n\in\mathbb{N}}$ und für alle $j \leq n$ auch von $\{f_n^j\}_{n\in\mathbb{N}}$. Also konvergiert $\{f_n^*(x_j)\}_{n\in\mathbb{N}}$ für alle $j \in \mathbb{N}$, d. h. $\{f_n^*\}_{n\in\mathbb{N}}$ konvergiert punktweise auf einer abzählbaren dichten Teilmenge.

Nun verwenden wir die gleichgradige Stetigkeit, um zu zeigen, dass daraus bereits die gleichmäßige Konvergenz folgt. Da $C(K)$ vollständig ist, genügt es zu zeigen, dass $\{f_n^*\}_{n\in\mathbb{N}}$ bezüglich der Metrik (1.2) eine Cauchy-Folge ist. Sei dafür $\varepsilon > 0$ beliebig und wähle $\delta > 0$ nach der Definition der gleichgradigen Stetigkeit. Da K präkompakt ist, existiert eine Überdeckung von K mit endlich vielen offenen Kugeln $U_1, \dots, U_p$ mit Radius $\frac{\delta}{2}$. Da X dicht in K liegt, muss jede dieser Kugeln mindestens einen Punkt aus X enthalten; um die Notation übersichtlich zu halten, gehen wir davon aus, dass $x_i \in U_i$ für alle $i = 1, \dots, p$ gilt. Aus der Konvergenz der endlich vielen Folgen $\{f_n^*(x_i)\}_{n\in\mathbb{N}}$ folgt die Existenz eines $N \in \mathbb{N}$ mit

$$|f_n^*(x_i) - f_m^*(x_i)| \leq \varepsilon \qquad \text{für alle } m, n \geq N \text{ und } i = 1, \dots, p.$$

Sei nun $x \in K$ beliebig. Dann existiert ein $j \in \{1, \dots, p\}$ mit $x, x_j \in U_j$, d. h. $d(x, x_j) < \delta$, und aus der gleichgradigen Stetigkeit von $\{f_n^*\}_{n\in\mathbb{N}} \subset A$ folgt

$$|f_n^*(x_j) - f_n^*(x)| \leq \varepsilon \qquad \text{für alle } n \in \mathbb{N}.$$

Zusammen gilt für alle $m, n \geq N$

$$|f_n^*(x) - f_m^*(x)| \leq |f_n^*(x) - f_n^*(x_j)| + |f_n^*(x_j) - f_m^*(x_j)| + |f_m^*(x_j) - f_m^*(x)| \leq 3\varepsilon.$$

Bilden wir das Supremum über alle $x \in K$, folgt daraus $d(f_n^*, f_m^*) \leq 3\varepsilon$ für alle $n, m \geq N$, d. h. die Teilfolge $\{f_n^*\}_{n\in\mathbb{N}}$ ist eine Cauchy-Folge und damit konvergent. □

Tatsächlich gilt auch die Umkehrung; siehe z. B. [Kaballo 2018, Satz 2.2].

Teil II

LINEARE OPERATOREN IN NORMIERTEN RÄUMEN

3 NORMIERTE VEKTORRÄUME

Wir kombinieren nun die in Teil I betrachteten topologischen bzw. metrischen Eigenschaften mit der algebraischen Struktur eines Vektorraums. Wie wir in den nächsten Kapiteln sehen werden, hat insbesondere die Vollständigkeit weitreichende Folgen.

Zur Erinnerung: Ein Vektorraum X über einem Körper $\mathbb{K}$ ist eine nichtleere Menge, die abgeschlossen ist bezüglich der Addition von Elementen aus X (den Vektoren) und Multiplikation mit Elementen aus $\mathbb{K}$ (den Skalaren, insbesondere der multiplikativen Identität 1 mit $1x = x$ für alle $x \in X$) sowie Assoziativ- und Distributivgesetze erfüllt. Wir werden uns hier auf die Fälle $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ beschränken.

Definition 3.1. Sei X ein Vektorraum über $\mathbb{K}$. Eine *Norm* auf X ist eine Abbildung $\|\cdot\|_X : X \to \mathbb{R}^+ := [0, \infty)$, die für alle $x, y \in X$ und $\lambda \in \mathbb{K}$ die folgenden Eigenschaften erfüllt:

(i) $\|x\|_X = 0$ genau dann, wenn $x = 0 \in X$ *(Nichtdegeneriertheit)*;

(ii) $\|\lambda x\|_X = |\lambda|\|x\|_X$ *(Homogenität)*;

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

In diesem Fall heißt das Paar $(X, \|\cdot\|_X)$ *normierter Vektorraum*. Ist die Norm aus dem Kontext offensichtlich, bezeichnen wir den normierten Vektorraum auch kurz mit X. Ist umgekehrt der Vektorraum offensichtlich, schreiben wir für die Norm kurz $\|\cdot\|$.

Zwei Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ auf X heißen *äquivalent*, falls Konstanten $c, C > 0$ existieren mit

$$c\|x\|_1 \leq \|x\|_2 \leq C\|x\|_1 \qquad \text{für alle } x \in X. \tag{3.1}$$

Bevor wir Beispiele betrachten, sammeln wir zunächst einige fundamentale Eigenschaften. Jede Norm auf X induziert vermöge

$$d(x, y) := \|x - y\| \qquad \text{für alle } x, y \in X$$

eine Metrik; zu jedem normierten Vektorraum $(X, \|\cdot\|)$ gehört also stets ein kanonischer metrischer Raum (X, d), zwischen denen wir in Folge nicht unterscheiden werden. Wir können also von offenen Mengen, konvergenten Folgen, und stetigen Funktionen in normierten Vektorräumen sprechen.

Die durch die Norm vermittels ihrer kanonischen Metrik induzierte Topologie ist besonders gut mit der algebraischen Struktur des Vektorraums verträglich. Wir erinnern uns: Zwei Metriken sind genau dann äquivalent, wenn sie die selben konvergenten Folgen besitzen; ist die Metrik durch eine Norm induziert, gilt $x_n \to x$ genau dann, wenn $\|x_n - x\| \to 0$ gilt.

Satz 3.2. *Seien $\|\cdot\|_1$ und $\|\cdot\|_2$ Normen auf dem Vektorraum X und d_1 bzw. d_2 die dadurch induzierten Metriken. Dann sind $\|\cdot\|_1$ und $\|\cdot\|_2$ äquivalent genau dann, wenn d_1 und d_2 äquivalent sind.*

Beweis. Die erste Richtung folgt direkt aus der Definition der Äquivalenz von Normen und der Konvergenz von Folgen in normierten Vektorräumen.

Sei nun angenommen, dass $\|\cdot\|_1$ und $\|\cdot\|_2$ nicht äquivalent sind. Mindestens eine der Ungleichungen in (3.1) kann also nicht gelten; nehmen wir an, es existiert kein $C > 0$ mit $\|x\|_2 \leq C\|x\|_1$ für alle $x \in X$. Für jedes $n \in \mathbb{N}$ existiert dann ein $x_n \in X$ mit $\|x_n\|_2 > n\|x_n\|_1$. Setzen wir $y_n := (n\|x_n\|_1)^{-1}x_n$, so ist $\|y_n\|_1 = \frac{1}{n} \to 0$, aber für alle $n \in \mathbb{N}$ gilt $\|y_n\|_2 > 1$. Also konvergiert die Folge $\{y_n\}_{n\in\mathbb{N}}$ bezüglich d_1, aber nicht bezüglich d_2, gegen 0, und somit können d_1 und d_2 auch nicht äquivalent sein. □

Äquivalente Normen definieren also die selben konvergenten Folgen. Aus (3.1) folgt aber auch, dass sie die selben Cauchy-Folgen definieren. Im Gegensatz zu metrischen Räumen vererbt sich daher die Vollständigkeit zwischen äquivalenten normierten Vektorräumen. Die Vollständigkeit ist also hier eine stärkere Eigenschaft und verdient daher einen eigenen Namen.

Definition 3.3. Ein vollständiger normierter Vektorraum heißt *Banachraum*.

Folgerung 3.4. *Sind $\|\cdot\|_1$ und $\|\cdot\|_2$ äquivalente Normen auf X, dann ist $(X, \|\cdot\|_1)$ ein Banachraum genau dann, wenn $(X, \|\cdot\|_2)$ ein Banachraum ist.*

Eine weitere Möglichkeit zu zeigen, dass ein normierter Vektorraum vollständig ist, liefert das folgende nützliche Lemma. Zur Erinnerung: Ein Unterraum ist eine Teilmenge, die abgeschlossen bezüglich der Vektorraumoperationen ist.

Lemma 3.5. *Sei $(X, \|\cdot\|_X)$ ein Banachraum und $U \subset X$ ein Unterraum. Dann ist $(U, \|\cdot\|_X)$ ein Banachraum genau dann, wenn U abgeschlossen ist.*

Beweis. Man vergewissert sich zunächst leicht, dass $(U, \|\cdot\|_X)$ ein normierter Vektorraum ist. Sei nun U abgeschlossen und $\{x_n\}_{n\in\mathbb{N}} \subset U$ eine Cauchy-Folge. Da X vollständig ist, konvergiert $x_n \to x \in X$, und aus der Abgeschlossenheit von U folgt $x \in U$.

Sei andererseits U vollständig und $\{x_n\}_{n\in\mathbb{N}} \subset U$ mit $x_n \to x \in X$. Dann ist $\{x_n\}_{n\in\mathbb{N}}$ insbesondere eine Cauchy-Folge (in X und damit ebenso in U) und besitzt wegen der Vollständigkeit von U einen Grenzwert $\tilde{x} \in U$. Da Grenzwerte eindeutig sind, muss $x = \tilde{x} \in U$ gelten. Also ist U abgeschlossen. □

Weiterhin sind die Vektorraumoperationen sowie die Norm stetig.

Satz 3.6. *Sei X ein normierter Vektorraum und $\{x_n\}_{n\in\mathbb{N}}, \{y_n\}_{n\in\mathbb{N}} \subset X$, und $\{\lambda_n\}_{n\in\mathbb{N}} \subset \mathbb{K}$ konvergente Folgen mit $x_n \to x$, $y_n \to y$ und $\lambda_n \to \lambda$. Dann gelten:*

(i) $x_n + y_n \to x + y$;

(ii) $\lambda_n x_n \to \lambda x$;

(iii) $\|x_n\| \to \|x\|$.

Beweis. Die Eigenschaften (i) und (ii) erhält man wie in $\mathbb{R}^n$ direkt aus der Dreiecksungleichung. Für (iii) verwenden wir die umgekehrte Dreiecksungleichung in der Form

$$|\|x_n\| - \|x\|| = |\|x_n - x + x\| - \|x\|| \leq \|x_n - x\| \to 0. \qquad \square$$

In normierten Vektorräumen gilt weiterhin $B_r(x) = \{y \in X : \|x - y\| \leq r\}$ und damit

$$B_r(x) = x + B_r(0) := \{y \in X : y = x + z \text{ mit } z \in B_r(0)\}$$

sowie

$$B_r(0) = rB_1(0) := \{y \in X : y = rz \text{ mit } z \in B_1(0)\},$$

und analog für $U_r(x)$. Es genügt also in einem normierten Vektorraum, die *Einheitskugel* $B_X := B_1(0)$ zu kennen.

Analog zur Konvergenz von Folgen definieren wir auch Reihen in normierten Räumen über die Norm. Sei $\{x_n\}_{n\in\mathbb{N}} \subset X$. Eine Reihe $\sum_{n=1}^\infty x_n$ in X *konvergiert*, wenn die Folge ihrer Partialsummen $S_N := \sum_{n=1}^N x_n$ konvergiert, d. h. wenn ein Element $x \in X$ existiert mit

$$\lim_{N\to\infty} \left\|x - \sum_{n=1}^N x_n\right\| = 0.$$

Eine Reihe $\sum_{n=1}^\infty x_n$ in X *konvergiert absolut*, wenn gilt

$$\sum_{n=1}^\infty \|x_n\| < \infty.$$

Durch Betrachtung der Partialsummenfolgen der Normfolge zeigt man wie im $\mathbb{R}^n$ das folgende Resultat.

Lemma 3.7. *Sei X ein Banachraum. Ist die Reihe $\sum_{n=1}^{\infty} x_n$ absolut konvergent, so ist sie auch konvergent, und es gilt*

$$\left\| \sum_{n=1}^{\infty} x_n \right\| \leq \sum_{n=1}^{\infty} \|x_n\|.$$

Wir betrachten nun die kanonischen Beispiele für normierte Vektorräume.

ENDLICHDIMENSIONALE RÄUME

Zur Erinnerung: Eine Teilmenge V eines Vektorraums X heißt *Basis*, wenn sich jedes $x \in X$ *eindeutig* als Linearkombination $x = \sum_{v\in V} \alpha_v v$ mit $\alpha_v \in \mathbb{K}$ für alle $v \in V$ (den *Koeffizienten*) darstellen lässt. Ist $V = \{v_1, \dots, v_n\}$ endlich, so heißt die Zahl n die *Dimension* von X. Existiert keine endliche Basis, so ist X unendlichdimensional.

Aus der Analysis ist bekannt, dass $(\mathbb{K}, |\cdot|)$ vollständig und damit ein Banachraum ist. Ebenso ist $\mathbb{K}^n$ ein Banachraum, versehen mit einer der Normen

$$\|x\|_1 := \sum_{i=1}^{n} |x_i|, \qquad \|x\|_2 := \left(\sum_{i=1}^{n} |x_i|^2\right)^{\frac{1}{2}}, \qquad \|x\|_\infty := \max_{i=1,\dots,n} |x_i|;$$

dies folgt direkt aus der Vollständigkeit von $(\mathbb{K}, |\cdot|)$ und der Endlichkeit der Summen bzw. Maximumsbildung. In jedem Fall sind Folgen genau dann konvergent, wenn sie komponentenweise konvergieren; die Normen sind also äquivalent. Tatsächlich gilt dies für alle Normen auf endlichdimensionalen Räumen.

Satz 3.8. *Ist X ein endlichdimensionaler Vektorraum, so sind alle Normen auf X äquivalent.*

Beweis. Ist X endlichdimensional, so existiert eine Basis $\{v_1, \dots, v_n\}$. Wir werden zeigen, dass jede Norm $\|\cdot\|$ äquivalent ist zur *euklidischen Norm*

$$\|x\|_2 = \left\|\sum_{i=1}^{n} \alpha_i v_i\right\|_2 := \left(\sum_{i=1}^{n} |\alpha_i|^2\right)^{\frac{1}{2}}.$$

Eine Folge konvergiert genau dann in $(X, \|\cdot\|_2)$, wenn die zugehörigen Koeffizientenfolgen konvergieren. Insbesondere ist $(X, \|\cdot\|_2)$ vollständig, da $(\mathbb{K}^n, \|\cdot\|_2)$ vollständig ist.

Setze nun $M := \max\{\|v_1\|, \dots, \|v_n\|\} > 0$. Dann folgt aus der Dreiecks- und der Cauchy–Schwarz-Ungleichung

$$\|x\| = \left\|\sum_{i=1}^{n} \alpha_i v_i\right\| \leq \sum_{i=1}^{n} |\alpha_i| \|v_i\| \leq \left(\sum_{i=1}^{n} |\alpha_i|^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^{n} \|v_i\|^2\right)^{\frac{1}{2}} \leq M\sqrt{n}\|x\|_2$$

und damit die zweite Ungleichung in (3.1) mit $C := M\sqrt{n} > 0$.

Für die erste Ungleichung betrachten wir $S := \{x \in X : \|x\|_2 = 1\}$. Offensichtlich ist S bezüglich $\|\cdot\|_2$ beschränkt. Außerdem ist S abgeschlossen, denn S ist das Urbild der abgeschlossenen Menge $\{1\}$ unter der stetigen Funktion $\|\cdot\|_2$ (siehe Satz 3.6 (iii)). Da wir X durch $\|\cdot\|_2$ mit der euklidischen Metrik versehen haben, können wir den Satz 2.5 von Heine–Borel anwenden; also ist S kompakt. Nach dem Satz von Weierstraß (Folgerung 2.10) nimmt die bezüglich $\|\cdot\|_2$ stetige Funktion $\|\cdot\|$ (dies folgt aus der zuerst bewiesenen Ungleichung) auf S ihr Minimum an. Es existiert also ein $\bar{x} \in S$ mit

$$c := \|\bar{x}\| \leq \|x\| \qquad \text{für alle } x \in S.$$

Da $\|\cdot\|_2$ eine Norm ist und $0 \notin S$ gilt, muss $\bar{x} \neq 0$ sein. Sei nun $x \in X \setminus \{0\}$ beliebig. Dann ist $\frac{x}{\|x\|_2} \in S$ und damit

$$c \leq \left\| \frac{x}{\|x\|_2} \right\| = \|x\|_2^{-1} \|x\|,$$

woraus die erste Ungleichung folgt. □

Da Vollständigkeit beim Übergang zu äquivalenten Normen erhalten bleibt und $(X, \|\cdot\|_2)$ vollständig ist, erhalten wir

Folgerung 3.9. *Alle endlichdimensionalen normierten Vektorräume sind vollständig.*

Auch bezüglich Kompaktheit nehmen endlichdimensionale Räume eine Sonderstellung ein. Wir benötigen das folgende Lemma.

Lemma 3.10 (Riesz). *Sei X ein normierter Vektorraum und $U \subsetneq X$ ein echter abgeschlossener Unterraum. Dann existiert für alle $\delta \in (0,1)$ ein $x_\delta \in X$ mit $\|x_\delta\| = 1$ und*

$$\|x_\delta - u\| \geq \delta \qquad \text{für alle } u \in U.$$

Beweis. Sei $x \in X \setminus U$ beliebig. Da U abgeschlossen ist, gilt

$$d := \inf \{\|x - u\| : u \in U\} > 0,$$

denn andernfalls gäbe es eine Folge $\{u_n\}_{n\in\mathbb{N}} \subset U$ mit $u_n \to x$ und $x \in \operatorname{cl} U = U$. Wegen $d < d/\delta$ existiert also nach Definition des Infimums ein $u_\delta \in U$ mit

$$0 < d \leq \|x - u_\delta\| < d/\delta.$$

Setze $x_\delta := \frac{x - u_\delta}{\|x - u_\delta\|}$, so dass $\|x_\delta\| = 1$.

Sei nun $u \in U$ beliebig. Da U ein Unterraum ist, ist auch $u_\delta + (\|x - u_\delta\|)u \in U$, und daraus folgt durch Einsetzen von x_δ

$$\|x_\delta - u\| = \|x - u_\delta\|^{-1} \|x - u_\delta - (\|x - u_\delta\|)u\| \geq \|x - u_\delta\|^{-1} d > \delta. \qquad \square$$

Wir definieren weiter den *Spann* (oder die *lineare Hülle*) einer (nicht notwendigerweise endlichen) Teilmenge A von X als

$$\operatorname{span} A = \left\{ \sum_{k=1}^{N} \lambda_k a_k : N \in \mathbb{N}, \lambda_k \in \mathbb{K}, a_k \in A \right\},$$

d. h. die Menge aller *endlichen* Linearkombinationen von Elementen aus A.

Satz 3.11. *Sei X ein normierter Raum. Dann ist die Einheitskugel B_X genau dann kompakt, wenn X endlichdimensional ist.*

Beweis. Ist X endlichdimensional, so folgt die Kompaktheit von B_X aus dem Satz 2.5 von Heine–Borel, denn sämtliche topologischen Eigenschaften wie Abgeschlossenheit, Beschränktheit und Kompaktheit bleiben nach Satz 3.2 beim Übergang zu der äquivalenten Norm im Beweis von Satz 3.8 erhalten.

Sei umgekehrt B_X kompakt. Dann existieren endlich viele offene Kugeln mit Radius $\frac{1}{2}$ mit $B_X \subset \bigcup_{i=1}^{n} U_{\frac{1}{2}}(x_i)$ mit $x_i \in B_X$. Wir zeigen nun, dass $\{x_1, \dots, x_n\}$ ein Erzeugendensystem von X darstellt und damit X endlichdimensional ist. Angenommen, dies wäre nicht der Fall. Dann ist $\operatorname{span}\{x_1, \dots, x_n\}$ ein echter und (wegen der Endlichkeit der Menge) abgeschlossener Unterraum von X, und nach dem Rieszschen Lemma 3.10 existiert ein $x_{\frac{1}{2}} \in B_X$ (wegen $\|x_{\frac{1}{2}}\| = 1$) mit $\|x_{\frac{1}{2}} - x_i\| > \frac{1}{2}$ für alle $i = 1, \dots, n$, im Widerspruch zur Wahl der x_i. □

Abgeschlossene beschränkte Mengen sind in unendlichdimensionalen normierten Räumen also *nicht* kompakt; das Fehlen dieser nützlichen Eigenschaft hat die Entwicklung eigenständiger funktionalanalytischer Werkzeuge entscheidend geprägt.

FOLGENRÄUME

Wir betrachten nun die einfachsten Beispiele für unendlichdimensionale normierte Räume. Wir bezeichnen die Menge aller Folgen in $\mathbb{K}$ mit

$$\mathbb{K}^{\mathbb{N}} := \left\{ \{x_k\}_{k\in\mathbb{N}} : x_k \in \mathbb{K} \text{ für alle } k \in \mathbb{N} \right\}$$

und definieren die Teilmengen

$$\begin{aligned}
\ell^\infty(\mathbb{K}) &:= \left\{ x \in \mathbb{K}^{\mathbb{N}} : x = \{x_k\}_{k\in\mathbb{N}} \text{ ist beschränkt} \right\},\\
c(\mathbb{K}) &:= \left\{ x \in \mathbb{K}^{\mathbb{N}} : x = \{x_k\}_{k\in\mathbb{N}} \text{ ist konvergent} \right\},\\
c_0(\mathbb{K}) &:= \left\{ x \in \mathbb{K}^{\mathbb{N}} : x = \{x_k\}_{k\in\mathbb{N}} \text{ ist Nullfolge} \right\},\\
c_e(\mathbb{K}) &:= \left\{ x \in \mathbb{K}^{\mathbb{N}} : x = \{x_k\}_{k\in\mathbb{N}} \text{ ist endliche Folge} \right\},
\end{aligned}$$

d. h. $\{x_k\}_{k\in\mathbb{N}} \in c_e(\mathbb{K})$ genau dann, wenn ein $N \in \mathbb{N}$ existiert mit $x_k = 0$ für alle $k \geq N$. Man vergewissert sich leicht, dass diese Mengen bezüglich der komponentenweisen Addition und Skalarmultiplikation abgeschlossen sind und Vektorräume bilden. Wir versehen diese Räume nun mit der *Supremumsnorm*

$$\|x\|_\infty := \sup_{k\in\mathbb{N}} |x_k| \qquad \text{für } x = \{x_k\}_{k\in\mathbb{N}}.$$

Satz 3.12. *Der Raum $(\ell^\infty(\mathbb{K}), \|\cdot\|_\infty)$ ist ein Banachraum.*

Beweis. Wir vergewissern uns zuerst, dass dadurch tatsächlich ein normierter Raum definiert ist. Beachte, dass $x \in \mathbb{K}^\mathbb{N}$ nach Definition genau dann beschränkt ist, wenn $\|x\|_\infty$ endlich ist. Nichtdegeneriertheit und Homogenität sind offensichtlich. Seien nun $x, y \in \ell^\infty(\mathbb{K})$ und $k \in \mathbb{N}$ beliebig. Dann gilt

$$|x_k + y_k| \leq |x_k| + |y_k| \leq \sup_{k\in\mathbb{N}} |x_k| + \sup_{k\in\mathbb{N}} |y_k| = \|x\|_\infty + \|y\|_\infty,$$

und Übergang zum Supremum über alle $k \in \mathbb{N}$ ergibt die Dreiecksungleichung.

Für die Vollständigkeit müssen wir Folgen von Folgen betrachten; wir ändern dafür wieder die Notation und schreiben $x = \{x(k)\}_{k\in\mathbb{N}} \in \mathbb{K}^\mathbb{N}$. Sei nun $\{x_n\}_{n\in\mathbb{N}}$ eine Cauchy-Folge in $(\ell^\infty(\mathbb{K}), \|\cdot\|_\infty)$. Dann ist wegen $|x_n(k)| \leq \|x_n\|_\infty$ für alle $k \in \mathbb{N}$ auch $\{x_n(k)\}_{n\in\mathbb{N}}$ eine Cauchy-Folge, die wegen der Vollständigkeit von $\mathbb{K}$ einen Grenzwert $x(k) \in \mathbb{K}$ besitzt. Dadurch wird eine Folge $x := \{x(k)\}_{k\in\mathbb{N}}$ definiert, für die wir nun nachweisen müssen, dass einerseits $x \in \ell^\infty(\mathbb{K})$ und andererseits $\|x_n - x\|_\infty \to 0$ gilt. Da $\{x_n\}_{n\in\mathbb{N}}$ eine Cauchy-Folge ist, existiert für $\varepsilon > 0$ beliebig ein $N \in \mathbb{N}$ mit

$$|x_n(k) - x_m(k)| \leq \|x_n - x_m\|_\infty \leq \varepsilon \qquad \text{für alle } n, m \geq N,\ k \in \mathbb{N}.$$

Sei nun $k \in \mathbb{N}$ beliebig. Wegen $x_n(k) \to x(k)$ existiert weiterhin ein $M(k)$ mit

$$|x_{M(k)}(k) - x(k)| \leq \varepsilon,$$

wobei wir ohne Einschränkung $M(k) \geq N$ annehmen dürfen. Es gilt also für alle $n \geq N$

$$|x_n(k) - x(k)| \leq |x_n(k) - x_{M(k)}(k)| + |x_{M(k)}(k) - x(k)| \leq 2\varepsilon.$$

Daraus folgt zum einen für alle $k \in \mathbb{N}$

$$|x(k)| \leq |x_N(k)| + |x(k) - x_N(k)| \leq \|x_N\|_\infty + 2\varepsilon < \infty$$

und damit $x \in \ell^\infty(\mathbb{K})$ durch Supremum über alle $k \in \mathbb{N}$, zum anderen analog

$$\|x_n - x\|_\infty \leq 2\varepsilon \qquad \text{für alle } n \geq N$$

und damit $x_n \to x$ in der Supremumsnorm. □

Für die anderen Vektorräume verwenden wir Lemma 3.5.

Satz 3.13. *Versehen mit der Supremumsnorm sind $c(\mathbb{K})$ und $c_0(\mathbb{K})$, aber nicht $c_e(\mathbb{K})$, Banachräume.*

Beweis. Man sieht leicht, dass alle drei Räume Untervektorräume von $\ell^\infty(\mathbb{K})$ und deshalb zusammen mit der Supremumsnorm normierte Vektorräume sind. Es bleibt also nur zu zeigen, dass $c(\mathbb{K})$ und $c_0(\mathbb{K})$, aber nicht $c_e(\mathbb{K})$ abgeschlossen sind.

Sei zunächst $\{x_n\}_{n\in\mathbb{N}} \subset c(\mathbb{K})$ eine in $\ell^\infty(\mathbb{K})$ konvergente Folge mit Grenzwert $x \in \ell^\infty(\mathbb{K})$. Wir zeigen, dass $x = \{x(k)\}_{k\in\mathbb{N}}$ selber Cauchy-Folge (in $\mathbb{K}$) ist. Sei $\varepsilon > 0$ gegeben. Dann existiert wegen der Konvergenz der Folge $\{x_n\}_{n\in\mathbb{N}}$ ein $N \in \mathbb{N}$ mit $\|x_N - x\|_\infty \leq \varepsilon$. Da die Folge $x_N = \{x_N(k)\}_{k\in\mathbb{N}} \in c(\mathbb{K})$ konvergent und damit Cauchy-Folge ist, existiert weiterhin ein $M \in \mathbb{N}$ mit $|x_N(k) - x_N(l)| \leq \varepsilon$ für alle $k, l \geq M$. Damit gilt für alle $k, l \geq M$

$$\begin{aligned}|x(k) - x(l)| &\leq |x(k) - x_N(k)| + |x_N(k) - x_N(l)| + |x_N(l) - x(l)| \\ &\leq 2\|x - x_N\|_\infty + \varepsilon \leq 3\varepsilon,\end{aligned}$$

d. h. x ist Cauchy-Folge und damit $x \in c(\mathbb{K})$.

Sei nun $\{x_n\}_{n\in\mathbb{N}} \subset c_0(\mathbb{K})$ eine in $\ell^\infty(\mathbb{K})$ konvergente Folge mit Grenzwert $x \in \ell^\infty(\mathbb{K})$. Wir wissen bereits $x \in c(\mathbb{K})$ und müssen lediglich noch $\lim_{k\to\infty} x(k) = 0$ zeigen. Gehen wir wie eben vor, erhalten wir für $\varepsilon > 0$ ein $N \in \mathbb{N}$ und ein $M \in \mathbb{N}$ mit

$$|x(M)| \leq |x(M) - x_N(M)| + |x_N(M)| \leq \|x - x_N\|_\infty + |x_N(M)| \leq \varepsilon + \varepsilon,$$

da $x_N \in c_0(\mathbb{K})$. Also ist auch x eine Nullfolge.

Für $c_e(\mathbb{K})$ betrachte für $n, k \in \mathbb{N}$

$$x_n(k) := \begin{cases} \frac{1}{k} & \text{falls } k \leq n, \\ 0 & \text{sonst,} \end{cases} \qquad x(k) := \frac{1}{k}.$$

Dann ist $x_n \in c_e(\mathbb{K})$ für alle $n \in \mathbb{N}$ und $\|x_n - x\|_\infty = \frac{1}{n+1} \to 0$, aber $x \notin c_e(\mathbb{K})$. □

Wie im letzten Schritt kann man zeigen, dass $c_e(\mathbb{K})$ dicht in $c_0(\mathbb{K})$ liegt (betrachte für $x \in c_0(\mathbb{K})$ die Folge $\{x_n\}_{n\in\mathbb{N}} \subset c_e(\mathbb{K})$ mit $x_n(k) = x(k)$ für $k \leq n$ und $x_n(k) = 0$ sonst).

Eine weitere Klasse von Banachräumen erhält man durch die *p-Normen.* Wir definieren für $x = \{x_k\}_{k\in\mathbb{N}} \in \mathbb{K}^\mathbb{N}$

$$\|x\|_p := \left(\sum_{k\in\mathbb{N}} |x_k|^p\right)^{\frac{1}{p}}, \qquad 1 \leq p < \infty.$$

und setzen

$$\ell^p(\mathbb{K}) := \left\{x \in \mathbb{K}^\mathbb{N} : \|x\|_p < \infty\right\}.$$

Satz 3.14. *Für $1 \le p < \infty$ ist $\ell^p(\mathbb{K})$ ein Banachraum.*

Beweis. Wieder überprüft man leicht die Nichtdegeneriertheit und Homogenität der Abbildung $\|\cdot\|_p$. Für die Dreiecksungleichung verwenden wir die *Minkowski-Ungleichung* für endliche Summen

$$\left(\sum_{k=1}^{N} |x_k + y_k|^p\right)^{\frac{1}{p}} \le \left(\sum_{k=1}^{N} |x_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{N} |y_k|^p\right)^{\frac{1}{p}}.$$

Wir erhalten

$$\sum_{k=1}^{N} |x_k + y_k|^p \le (\|x\|_p + \|y\|_p)^p \qquad \text{für alle } N \in \mathbb{N},$$

woraus durch Grenzübergang $N \to \infty$ und Ziehen der p-ten Wurzel folgt

$$\|x + y\|_p \le \|x\|_p + \|y\|_p.$$

Also ist $x + y \in \ell^p(\mathbb{K})$ und $\ell^p(\mathbb{K})$ ein normierter Raum.

Sei nun $\{x_n\}_{n\in\mathbb{N}} \subset \ell^p(\mathbb{K})$ eine Cauchy-Folge. Dann gilt

$$|x_n(k) - x_m(k)|^p \le \sum_{j=1}^{\infty} |x_n(j) - x_m(j)|^p = \|x_n - x_m\|_p^p \qquad \text{für alle } k, n, m \in \mathbb{N}.$$

Also ist $\{x_n(k)\}_{n\in\mathbb{N}}$ Cauchy-Folge für alle $k \in \mathbb{N}$ und konvergiert daher gegen ein $x(k) \in \mathbb{K}$. Für $\varepsilon > 0$ finden wir daher ein $M \in \mathbb{N}$ mit

$$\|x_n - x_m\|_p \le \varepsilon \qquad \text{für alle } m, n \ge M,$$

und für jedes $N \in \mathbb{N}$ ein $m = m(N) \ge N$, so dass gilt

$$\left(\sum_{k=1}^{N} |x_m(k) - x(k)|^p\right)^{\frac{1}{p}} \le \varepsilon$$

(dies ist möglich, da $x_n(k) \to x(k)$ für alle $k \le N$). Für alle $n \ge M$ gilt dann (wieder unter Verwendung der Minkowski-Ungleichung)

$$\begin{aligned}
\left(\sum_{k=1}^{N} |x_n(k) - x(k)|^p\right)^{\frac{1}{p}} &\le \left(\sum_{k=1}^{N} |x_n(k) - x_m(k)|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{N} |x_m(k) - x(k)|^p\right)^{\frac{1}{p}} \\
&\le \|x_n - x_m\|_p + \left(\sum_{k=1}^{N} |x_m(k) - x(k)|^p\right)^{\frac{1}{p}} \\
&\le \varepsilon + \varepsilon.
\end{aligned}$$

Grenzübergang $N \to \infty$ liefert $\|x_n - x\|_p \le 2\varepsilon$ für alle $n \ge M$ und damit $x_n \to x$ sowie $x_n - x \in \ell^p(\mathbb{K})$, woraus $x = (x - x_n) + x_n \in \ell^p(\mathbb{K})$ folgt. □

Ähnlich wie im Beweis von Satz 3.13 zeigt man, dass $c_e(\mathbb{K})$ dicht in $\ell^p(\mathbb{K})$ für $1 \leq p < \infty$ liegt.

Satz 3.15. *Die Räume $c_0(\mathbb{K})$ und $\ell^p(\mathbb{K})$ für $1 \leq p < \infty$ sind separabel. Der Raum $\ell^\infty(\mathbb{K})$ ist nicht separabel.*

Beweis. Für $c_0(\mathbb{K})$ und $\ell^p(\mathbb{K})$ betrachte für $\mathbb{K} = \mathbb{R}$ den Raum $c_e(\mathbb{Q})$ der rationalen endlichen Folgen bzw. $c_e(\mathbb{Q} + i\mathbb{Q})$ für $\mathbb{K} = \mathbb{C}$. Da die rationalen Zahlen dicht in $\mathbb{R}$ liegen, zeigt ein Diagonalfolgenargument, dass diese Mengen dicht in $c_0(\mathbb{K})$ und $\ell^p(\mathbb{K})$ liegen. Weiterhin sind sie abzählbar, woraus die Separabilität folgt.

Für $\ell^\infty(\mathbb{K})$ betrachte eine beliebige Teilmenge $M \subset \mathbb{N}$ und definiere $x_M \in \ell^\infty(\mathbb{K})$ durch

$$x_M(k) := \begin{cases} 1 & \text{falls } k \in M, \\ 0 & \text{sonst.} \end{cases}$$

Für $M, N \subset \mathbb{N}$ mit $M \neq N$ gilt dann $\|x_M - x_N\|_\infty = 1$. Ist nun $A \subset \ell^\infty(\mathbb{K})$ eine beliebige abzählbare Teilmenge, so kann für jedes $x \in A$ die offene Kugel $U_{\frac{1}{2}}(x)$ höchstens ein solches x_M enthalten. Da aber die Menge aller Teilmengen von $\mathbb{N}$ und damit die Anzahl solcher x_M überabzählbar ist, kann A nicht dicht liegen. □

Funktionenräume

Auch $\mathbb{K}$-wertige Funktionen bilden einen $\mathbb{K}$-Vektorraum, wenn man Addition und Skalarmultiplikation punktweise definiert. Ist (X, d) ein metrischer Raum, so definieren wir den Raum der beschränkten Funktionen auf X,

$$B(X) := \{f : X \to \mathbb{K} : f \text{ beschränkt}\}$$

sowie die Supremumsnorm

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

Man zeigt wörtlich wie im Beweis von Satz 3.12 (man ersetze lediglich überall $k \in \mathbb{N}$ durch $x \in X$), dass $(B(X), \|\cdot\|_\infty)$ ein Banachraum ist. Analog beweist man, dass $C_b(X)$ ein abgeschlossener Unterraum von $B(X)$ und damit $(C_b(X), \|\cdot\|_\infty)$ ebenfalls ein Banachraum ist; dies folgt auch daraus, dass $\|\cdot\|_\infty$ genau die Metrik in (1.2) und damit die Topologie der gleichmäßigen Konvergenz induziert, und dass gleichmäßig konvergente Folgen stetiger Funktionen einen stetigen Grenzwert haben. Der $c_0(\mathbb{K})$ entsprechende Funktionenraum ist der Raum der „zum Rand hin“ verschwindenden Funktionen

$$C_0(X) := \{f \in C(X) : \text{Für alle } \varepsilon > 0 \text{ ist } \{x \in X : |f(x)| \geq \varepsilon\} \text{ kompakt}\},$$

der wiederum ein abgeschlossener Unterraum in $C_b(X)$ und damit zusammen mit der Supremumsnorm ein Banachraum ist.

Nach dem Weierstraßschen Approximationssatz[1] lässt sich jede stetige Funktion auf kompakten Mengen beliebig gut durch Polynome annähern. Ein Diagonalfolgenargument wie in Satz 3.15 zeigt dann, dass für kompaktes $X \subset \mathbb{R}^n$ die Polynome mit rationalen Koeffizienten dicht in $C_b(X)$ liegen und damit $C_b(X)$ separabel ist.

Man kann auch Funktionenräume analog zu $\ell^p(\mathbb{K})$, $1 \leq p \leq \infty$, definieren. Die Konstruktion ist aufwendig und erfordert einige maßtheoretische Vorarbeit; daher behandeln wir diese Räume nur kursorisch (insbesondere, da die rein funktionalanalytischen Argumente im wesentlichen die selben sind wie für $\ell^p(\mathbb{K})$) und verweisen für Details auf [Dobrowolski 2010, Kapitel 4] oder [Brokate & Kersting 2019, Kapitel VI].

Sei $\Omega \subset \mathbb{R}^n$ eine Lebesgue-messbare[2] Teilmenge und definiere für eine Lebesgue-messbare Funktion $f : \Omega \to \mathbb{R}$

$$
\begin{aligned}
\|f\|_p &:= \left(\int_\Omega |f(x)|^p \, dx\right)^{\frac{1}{p}}, \qquad 1 \leq p < \infty, \\
\|f\|_\infty &:= \operatorname*{ess\,sup}_{x\in\Omega} |f(x)|,
\end{aligned}
$$

wobei das *essentielle Supremum* von $|f|$ definiert ist als

$$
\operatorname*{ess\,sup}_{x\in\Omega} |f(x)| := \inf\left\{M > 0 : \{x \in \Omega : |f(x)| > M\} \text{ hat Lebesgue-Maß } 0\right\}.
$$

Im folgenden fassen wir Funktionen, die sich nur auf einer Menge vom Lebesgue-Maß 0 unterscheiden, zu einer Äquivalenzklasse zusammen, die wir der Übersichtlichkeit halber wieder mit f bezeichnen. Dann ist für alle $1 \leq p \leq \infty$ der Raum

$$
L^p(\Omega) := \left\{f : \Omega \to \mathbb{R} : \|f\|_p < \infty\right\}
$$

zusammen mit der entsprechenden Norm ein Banachraum.[3] Man kann zeigen, dass $C_b(\Omega)$ dicht in $L^p(\Omega)$ für $1 \leq p < \infty$ liegt, woraus die Separabilität dieser Räume folgt. Dagegen ist $L^\infty(\Omega)$ nicht separabel, was man mit ähnlichen Argumenten wie für $\ell^\infty(\Omega)$ zeigen kann.

[1] siehe z. B. [Werner 2018, Satz I.2.11]

[2] Siehe z. B. [Brokate & Kersting 2019, Kapitel II und III] für diese und die folgenden Begriffe aus der Maßtheorie.

[3] Diese Konstruktion ist auch für allgemeine Maßräume möglich, was wichtig für die Wahrscheinlichkeitstheorie ist; siehe z. B. [Alt 2012, § 1.14–1.19] oder [Brokate & Kersting 2019, Kapitel VI].

4 LINEARE OPERATOREN

Wir betrachten nun Abbildungen zwischen normierten Räumen, und nutzen auch hier das Zusammenspiel algebraischer und topologischer Eigenschaften aus. Für normierte Räume $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ interessieren wir uns daher für Abbildungen $T : X \to Y$, die

(i) *linear*, d. h. $T(\lambda x_1 + x_2) = \lambda T(x_1) + T(x_2)$ für $x_1, x_2 \in X$, $\lambda \in \mathbb{K}$, und

(ii) stetig im Sinne von Definition 1.7 sind, d. h. für die $x_n \to x$ impliziert, dass $Tx_n \to Tx$.

Eine solche Abbildung nennt man auch *stetigen (linearen) Operator*; um die Linearität zu verdeutlichen, schreibt man auch oft $Tx := T(x)$. Für $T : X \to Y$ definieren wir analog zur linearen Algebra

(i) den *Kern* $\ker T := T^{-1}(\{0\}) = \{x \in X : Tx = 0\} \subset X$;

(ii) das *Bild* $\operatorname{ran} T := T(X) = \{Tx : x \in X\} \subset Y$;

(iii) den *Graph* $\operatorname{graph} T := \{(x, Tx) : x \in X\} \subset X \times Y$.

Aus der Linearität von T folgt sofort, dass dies Unterräume sind. Ist T stetig, ist $\ker T$ als Urbild der abgeschlossenen Menge $\{0\}$ sogar abgeschlossen. Dies gilt für $\operatorname{ran} T$ (und $\operatorname{graph} T$) dagegen nicht unbedingt, was eine wesentliche Schwierigkeit im Umgang mit Abbildungen zwischen unendlichdimensionalen Vektorräumen darstellt.

Die zusätzliche lineare Struktur erlaubt eine einfachere Charakterisierung der Stetigkeit.

Lemma 4.1. *Für eine lineare Abbildung $T : X \to Y$ zwischen normierten Räumen $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ sind äquivalent:*

(i) T ist stetig auf X;

(ii) T ist stetig in $0 \in X$;

(iii) T ist beschränkt, *d. h. es gibt eine Konstante $C \geq 0$ mit*

$$\|Tx\|_Y \leq C\|x\|_X \qquad \text{für alle } x \in X.$$

Beweis. *(i)* $\Longrightarrow$ *(ii)* ist klar.

(ii) $\Longrightarrow$ *(iii)*: Nach Definition 1.7 (ii) existiert für $\varepsilon = 1$ ein $\delta > 0$ mit $T(B_\delta(0)) \subset B_1(T(0)) = B_1(0)$. Aus der Definition der abgeschlossenen Kugeln folgt damit

$$\|Tx\|_Y \leq 1 \qquad \text{für alle } x \in X \text{ mit } \|x\|_X \leq \delta. \tag{4.1}$$

Für beliebiges $x \in X \setminus \{0\}$ ist nun $\delta \frac{x}{\|x\|_X} \in B_\delta(0)$, und daher kann (4.1) mit Hilfe der Linearität von T und der Homogenität der Norm umgeformt werden zu

$$\|Tx\|_Y \leq \frac{1}{\delta}\|x\|_X,$$

woraus die Behauptung mit $C := \delta^{-1}$ folgt.

(iii) $\Longrightarrow$ *(i)*: Sei $x \in X$ und $\varepsilon > 0$ gegeben. Wir zeigen, dass $\delta > 0$ mit $T(B_\delta(x)) \subset B_\varepsilon(Tx)$ existiert. Konkret wählen wir $\delta := \frac{\varepsilon}{C}$; dann gilt für alle $z \in B_\delta(x)$

$$\|Tz - Tx\|_Y = \|T(z-x)\|_Y \leq C\|z-x\|_X \leq C\delta = \varepsilon,$$

was zu zeigen war. □

Auf endlichdimensionalen Räumen sind lineare Operatoren automatisch stetig.

Lemma 4.2. *Sind X und Y normierte Räume mit X endlichdimensional und $T : X \to Y$ linear, so ist T stetig.*

Beweis. Da X endlichdimensional ist, existiert eine Basis $\{v_1, \dots, v_n\}$ von X. Für

$$x = \sum_{k=1}^n x_k v_k \in X, \qquad x_k \in \mathbb{K} \quad \text{für alle} \quad k = 1, \dots, n,$$

ist dann

$$\|Tx\|_Y \leq \sum_{k=1}^n |x_k| \|Tv_k\|_Y \leq \max_{k=1,\dots,n} \|Tv_k\|_Y \sum_{k=1}^n |x_k| =: M\|x\|_1$$

für $M := \max_{k=1,\dots,n} \|Tv_k\|_Y < \infty$. Da nach Satz 3.8 alle Normen auf dem endlichdimensionalen Raum X äquivalent sind, existiert ein $C > 0$ mit $\|x\|_1 \leq C\|x\|_X$ für alle $x \in X$, woraus mit Lemma 4.1 die Stetigkeit folgt. □

Die Menge aller stetigen Operatoren von X nach Y wird mit $L(X, Y)$ bezeichnet; dieser Raum wird zu einem Vektorraum, indem wir Operatoren punktweise addieren und skalieren, d. h. $T_1 + \alpha T_2 \in L(X, Y)$ für $T_1, T_2 \in L(X, Y)$ und $\alpha \in \mathbb{K}$ definieren wir durch $(T_1 + \alpha T_2)x := T_1 x + \alpha T_2 x$ für alle $x \in X$. Es liegt nahe, als Norm eines stetigen Operators die kleinste mögliche Konstante in Lemma 4.1 (iii) zu wählen; konkret definieren wir die *Operatornorm*

$$\|T\|_{L(X,Y)} := \sup_{x \in B_X} \|Tx\|_Y. \tag{4.2}$$

Dass diese Definition tatsächlich zu der Motivation passt, besagt das folgende Lemma, dass man durch einfache Abschätzung bzw. geeignete Wahl von $x \in X$ zeigt.

Lemma 4.3. *Für $T \in L(X,Y)$ gelten:*

(i) $\|T\|_{L(X,Y)} = \sup_{x\in X, \|x\|_X<1} \|Tx\|_Y = \sup_{x\in X,\|x\|_X=1} \|Tx\|_Y = \sup_{x\in X\setminus\{0\}} \frac{\|Tx\|_Y}{\|x\|_X}$;

(ii) $\|T\|_{L(X,Y)} = \inf\{C>0 : \|Tx\|_Y \leq C\|x\|_X$ *für alle* $x \in X\}$;

(iii) $\|Tx\|_Y \leq \|T\|_{L(X,Y)}\|x\|_X$ *für alle* $x \in X$.

Offensichtlich ist für jeden normierten Raum X die *Identität* $\mathrm{Id}_X : X \to X$, $x \mapsto x$, ein stetiger Operator mit Operatornorm 1. Weiterhin entspricht die Definition genau der aus der Numerik bekannten induzierten Matrixnorm; das wichtigste Beispiel ist die von der euklidischen Norm (d. h. $X = (\mathbb{R}^n, \|\cdot\|_2)$) induzierte *Spektralnorm*. Die folgenden Beispiele illustrieren die Definition in unendlichdimensionalen Räumen.

Beispiel 4.4.

(i) Sei $X = c(\mathbb{K})$ und $T : X \to \mathbb{K}$, $x = \{x_k\}_{k\in\mathbb{N}} \mapsto \lim_{k\to\infty} x_k$. Aus den Rechenregeln für Grenzwerte folgt sofort die Linearität von T. Für die Stetigkeit betrachten wir

$$|Tx| = \left|\lim_{k\to\infty} x_k\right| = \lim_{k\to\infty} |x_k| \leq \sup_{k\in\mathbb{N}} |x_k| = \|x\|_\infty.$$

Da für die konstante Folge $x = \{1\}_{k\in\mathbb{N}}$ gilt $|Tx| = 1 = \|x\|_\infty$, ist $\|T\|_{L(X,\mathbb{K})} = 1$.[1]

(ii) Analog zeigt man, dass für $X = C([0,1])$ die *Punktauswertung* $T : X \to \mathbb{R}$, $x \mapsto x(0)$, linear und stetig ist mit Operatornorm $\|T\|_{L(C,\mathbb{R})} = 1$.

(iii) Sei $X = C([a,b])$ und $T : X \to \mathbb{R}$, $x \mapsto \int_a^b x(t)\,dt$. Dann ist T wieder linear und stetig wegen

$$|Tx| = \left|\int_a^b x(t)\,dt\right| \leq (b-a)\|x\|_\infty$$

mit Gleichheit für konstante Funktionen. Also ist $\|T\|_{L(X,\mathbb{R})} = b-a$.

(iv) Sei $X = C^1([0,1]) := \{f : [0,1] \to \mathbb{R} : f \text{ stetig differenzierbar}\}$ und $Y = C([0,1])$, und betrachte den *Ableitungsoperator* $D : X \to Y$, $x \mapsto x'$, der bekanntermaßen linear ist. Versehen wir $C([0,1])$ und $C^1([0,1])$ mit der Supremumsnorm, so ist D nicht stetig, denn für $x_n(t) := t^n$ gilt $\|x_n\|_\infty = 1$ für alle $n \in \mathbb{N}$, aber $\|Dx_n\|_\infty = \|nt^{n-1}\|_\infty = n \to \infty$.

Versehen wir $C^1([0,1])$ dagegen mit der Norm $\|x\|_{C^1} := \|x\|_\infty + \|x'\|_\infty$, so ist D wegen $\|Dx\|_\infty = \|x'\|_\infty \leq \|x\|_{C^1}$ stetig mit Norm 1.

Es steht noch aus zu zeigen, dass (4.2) tatsächlich eine Norm definiert.

[1] Wegen $c_0(\mathbb{K}) = \ker T$ liefert die Stetigkeit zusammen mit Folgerung 1.9 übrigens einen wesentlich eleganteren Beweis für die Abgeschlossenheit von $c_0(\mathbb{K})$ in $c(\mathbb{K})$.

Satz 4.5. *Das Paar* $(L(X,Y), \|\cdot\|_{L(X,Y)})$ *ist ein normierter Raum. Ist* Y *vollständig, dann ist* $L(X,Y)$ *ein Banachraum.*

Beweis. Aus Lemma 4.3 (ii) folgt, dass für einen stetigen Operator $T : X \to Y$ gilt $\|T\|_{L(X,Y)} < \infty$. Die Homogenität und Nichtdegeneriertheit folgt aus den entsprechenden Eigenschaften der Norm in Y; letzteres unter Verwendung der Tatsache, dass $T = 0$ genau dann gilt, wenn $Tx = 0$ für alle $x \in X$ gilt. Für die Dreiecksungleichung sei $x \in B_X$ beliebig. Dann gilt nach Lemma 4.3 (iii) für alle $S, T \in L(X,Y)$

$$\|(S+T)x\|_Y = \|Sx + Tx\|_Y \leq \|Sx\|_Y + \|Tx\|_Y \leq \|S\|_{L(X,Y)} + \|T\|_{L(X,Y)},$$

und Supremum über alle $x \in B_X$ liefert die Aussage.

Für die Vollständigkeit sei $\{T_n\}_{n\in\mathbb{N}}$ eine Cauchy-Folge in $L(X,Y)$. Für alle $x \in X$ ist wegen Lemma 4.3 (iii) dann $\{T_n x\}_{n\in\mathbb{N}}$ eine Cauchy-Folge in Y, die nach Annahme gegen ein Element $y_x \in Y$ konvergiert. Dies definiert eine Abbildung $T : X \to Y$, $x \mapsto y_x$. Dass T linear ist, folgt aus

$$T(\lambda x_1 + x_2) = \lim_{n\to\infty} T_n(\lambda x_1 + x_2) = \lim_{n\to\infty} \lambda T_n x_1 + \lim_{n\to\infty} T_n x_2 = \lambda T(x_1) + T(x_2),$$

wobei wir neben der Linearität von T_n die Stetigkeit der Addition und Skalarmultiplikation (Satz 3.6 (i,ii)) verwendet haben.

Wir zeigen jetzt $\|T\|_{L(X,Y)} < \infty$ (und damit $T \in L(X,Y)$) und $\|T_n - T\|_{L(X,Y)} \to 0$. Da $\{T_n\}_{n\in\mathbb{N}}$ eine Cauchy-Folge ist, finden wir für $\varepsilon > 0$ ein $N \in \mathbb{N}$ mit

$$\|T_n - T_m\|_{L(X,Y)} \leq \varepsilon \qquad \text{für alle } n, m \geq N.$$

Sei nun $x \in B_X$ beliebig. Wegen $T_n x \to Tx$ finden wir ein $M = M(\varepsilon, x) \geq N$ mit

$$\|T_M x - Tx\|_Y \leq \varepsilon.$$

Dann gilt

$$\|T_n x - Tx\|_Y \leq \|T_n x - T_M x\|_Y + \|T_M x - Tx\|_Y \leq \|T_n - T_m\|_{L(X,Y)} + \varepsilon \leq 2\varepsilon.$$

Supremum über alle $x \in B_X$ ergibt nun $\|T_n - T\|_{L(X,Y)} \leq 2\varepsilon$. Daraus folgt einerseits mit Hilfe der Dreiecksungleichung $\|T\|_{L(X,Y)} \leq \|T_n\|_{L(X,Y)} + 2\varepsilon < \infty$, andererseits wegen der Beliebigkeit von ε die Konvergenz $\|T_n - T\|_{L(X,Y)} \to 0$. □

Wir zeigen noch zwei nützliche Eigenschaften. Zunächst folgt aus Lemma 4.3 (iii) sofort

Folgerung 4.6. *Seien* X, Y, Z *normierte Räume,* $T \in L(X,Y)$ *und* $S \in L(Y,Z)$. *Dann ist* $S \circ T \in L(X,Z)$ *mit* $\|S \circ T\|_{L(X,Z)} \leq \|S\|_{L(Y,Z)} \|T\|_{L(X,Y)}$.

Der folgende Satz ist hilfreich bei der Konstruktion von stetigen Operatoren mit gewünschten Eigenschaften.

Satz 4.7. *Ist $U \subset X$ ein dichter Unterraum, Y ein Banachraum und $T \in L(U, Y)$, so existiert eine eindeutige* stetige Fortsetzung $S \in L(X, Y)$ *mit* $S|_U = T$ *und* $\|S\|_{L(X,Y)} = \|T\|_{L(U,Y)}$.

Beweis. Sei $x \in X$. Dann existiert nach Annahme eine Folge $\{x_n\}_{n\in\mathbb{N}} \subset U$ mit $x_n \to x$. Insbesondere ist $\{x_n\}_{n\in\mathbb{N}}$ eine Cauchy-Folge. Wegen

$$\|Tx_n - Tx_m\|_Y \leq \|T\|_{L(U,Y)}\|x_n - x_m\|_X \qquad n, m \in \mathbb{N} \tag{4.3}$$

ist auch $\{Tx_n\}_{n\in\mathbb{N}}$ Cauchy-Folge und konvergiert im Banachraum Y gegen ein $y_x \in Y$. Dass dieses Element eindeutig ist (also nicht von der Wahl der Cauchy-Folge abhängt), sieht man wie folgt: Konvergieren zwei Cauchy-Folgen $\{x_n\}_{n\in\mathbb{N}}$ und $\{\tilde{x}_n\}_{n\in\mathbb{N}}$ gegen denselben Grenzwert x, so ist die Differenzenfolge $\{x_n - \tilde{x}_n\}_{n\in\mathbb{N}}$ eine Nullfolge, und aus der Abschätzung (4.3) folgt, dass auch $Tx_n - T\tilde{x}_n \to 0$ konvergiert. Wie oben zeigt man nun, dass $S : X \to Y$, $x \mapsto y_x$, einen linearen Operator definiert. Insbesondere gilt $y_x = Tx$ für $x \in U$, da wir dann die konstante Folge $\{x\}_{n\in\mathbb{N}}$ wählen können; dies zeigt $S|_U = T$.

Mit der selben Wahl von $x_n \to x$ erhalten wir aus der Stetigkeit von T und der der Norm (Satz 3.6 (iii)) zusammen mit Lemma 4.3 (iii), dass

$$\begin{aligned}\|Sx\|_Y &= \lim_{n\to\infty} \|Sx_n\|_Y = \lim_{n\to\infty} \|Tx_n\|_Y \\ &\leq \|T\|_{L(U,Y)} \left(\lim_{n\to\infty} \|x_n\|_X\right) = \|T\|_{L(U,Y)}\|x\|_X \qquad \text{für alle } x \in X.\end{aligned}$$

Supremum über alle $x \in B_X$ ergibt dann die Abschätzung $\|S\|_{L(X,Y)} \leq \|T\|_{L(U,Y)}$ und damit bereits die Stetigkeit von S. Die umgekehrte Abschätzung erhalten wir aus

$$\|T\|_{L(U,Y)} = \sup_{x\in B_U} \|Tx\|_Y = \sup_{x\in B_U} \|Sx\|_Y \leq \sup_{x\in B_X} \|Sx\|_Y = \|S\|_{L(X,Y)},$$

da wegen $B_U = \{x \in U : \|x\|_X \leq 1\} \subset B_X$ das Supremum über eine größere Menge gebildet wird. □

Der Beweis ist ein klassisches Beispiel für ein in der Funktionalanalysis häufig angewendetes *Dichtheitsargument*: Um eine Eigenschaft für alle Elemente aus X nachzuweisen, zeigt man, dass sie in einer dichten Teilmenge gilt und durch Grenzübergang erhalten bleibt.

Über stetige Operatoren zwischen zwei normierten Räumen lassen sich viele Eigenschaften von einem Raum auf den anderen übertragen.

Wir erinnern: Ein Operator $T : X \to Y$ heißt *injektiv*, falls $\ker T = \{0\}$, und *surjektiv*, falls $\operatorname{ran} T = Y$. Ein injektiver und surjektiver Operator heißt *bijektiv*; in diesem Fall können wir einen linearen Operator $T^{-1} : Y \to X$ vermittels der eindeutigen Zuordnung

$y \mapsto x \in T^{-1}(\{y\})$ definieren, der $T^{-1}T = \mathrm{Id}_X$ und $TT^{-1} = \mathrm{Id}_Y$ erfüllt. Wir nennen dann T^{-1} *Inverse* und T *invertierbar*. Ist $T^{-1} \in L(Y, X)$ (was nicht automatisch der Fall ist!), so heißt T *stetig invertierbar*. Ein stetig invertierbarer Operator heißt auch *Isomorphismus*. Gilt $\|Tx\|_Y = \|x\|_X$ für alle $x \in X$, so heißt T *Isometrie*.

Wir nennen nun X und Y *isomorph* und schreiben $X \simeq Y$, wenn es einen Isomorphismus $T : X \to Y$ gibt. Analog nennen wir X und Y *isometrisch isomorph* und schreiben $X \cong Y$, wenn ein $T \in L(X, Y)$ existiert, der sowohl Isomorphismus als auch Isometrie ist. Isomorphe Räume sind in gewisser Weise nur unterschiedliche Darstellungen des selben Raumes; für isometrisch isomorphe Räume sind die Darstellungen „uniform". Analog zu Folgerung 3.4 gilt, dass falls X Banachraum und $X \simeq Y$ ist, auch Y ein Banachraum ist.

Sind zwei Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ auf X äquivalent, so ist die Identität $\mathrm{Id}_X : (X, \|\cdot\|_1) \to (X, \|\cdot\|_2)$ ein Isomorphismus. Die Isomorphie von verschiedenen Räumen stellt also eine Verallgemeinerung der Äquivalenz verschiedener Normen auf dem selben Raum dar. Analog zu Satz 3.8 gilt, dass endlichdimensionale Räume gleicher Dimension stets (nicht unbedingt isometrisch) isomorph sind.

Satz 4.8. *Sind X, Y endlichdimensionale normierte Vektorräume mit $\dim(X) = \dim(Y)$, so gilt $X \simeq Y$.*

Beweis. Wir zeigen, dass jeder n-dimensionale normierte Vektorraum $(X, \|\cdot\|)$ isomorph zu $(\mathbb{K}^n, \|\cdot\|_2)$ ist. Sei $\{v_1, \dots, v_n\}$ eine Basis von X. Dann ist

$$T : X \to \mathbb{K}^n, \qquad x = \sum_{k=1}^{n} x_k v_k \mapsto (x_1, \dots, x_n),$$

linear und nach Lemma 4.2 stetig, und ebenso die Inverse

$$T^{-1} : \mathbb{K}^n \to X, \qquad (x_1, \dots, x_n) \mapsto \sum_{k=1}^{n} x_k v_k =: x.$$

Da Kompositionen stetiger Abbildungen wieder stetig sind, folgt aus $X \simeq (\mathbb{K}^n, \|\cdot\|) \simeq Y$ die gewünschte Aussage $X \simeq Y$. □

Ein weniger offensichtliches Beispiel ist $c(\mathbb{K}) \simeq c_0(\mathbb{K})$, vermöge der Abbildung

$$T : c(\mathbb{K}) \to c_0(\mathbb{K}), \quad (x_1, x_2, x_3, \dots) \mapsto \left(\lim_{k\to\infty} x_k, x_1 - \lim_{k\to\infty} x_k, x_2 - \lim_{k\to\infty} x_k, \dots \right).$$

Eine unterschiedliche Verallgemeinerung der Normäquivalenz ist möglich für normierte Räume X, Y mit $X \subset Y$. In diesem Fall heißt X *stetig eingebettet* in Y, geschrieben $X \hookrightarrow Y$, falls die Identität $\mathrm{Id} : X \to Y$ stetig ist, d. h. ein $C > 0$ existiert mit

$$\|x\|_Y \leq C\|x\|_X \qquad \text{für alle } x \in X.$$

Zum Beispiel gilt $\ell^p(\mathbb{K}) \hookrightarrow \ell^q(\mathbb{K})$ für $1 \leq p \leq q \leq \infty$ und, falls $\Omega \subset \mathbb{R}^n$ beschränkt ist, $L^p(\Omega) \hookrightarrow L^q(\Omega)$ für $1 \leq q \leq p \leq \infty$ (beachte die unterschiedliche Reihenfolge).

5 DAS PRINZIP DER GLEICHMÄSSIGEN BESCHRÄNKTHEIT

Wir kommen nun zu einem Herzstück der Funktionalanalysis: aus der Vollständigkeit von Banachräumen folgt, dass bestimmte punktweise Eigenschaften von linearen Operatoren *gleichmäßig* gelten. Als Konsequenz erhalten wir in diesem Kapitel einige der Hauptsätze über lineare Operatoren; weitere wichtige Folgerungen werden in Teil III auftauchen.

Da die Vollständigkeit eine metrische Eigenschaft ist, beruhen alle diese Sätze auf einer abstrakten Eigenschaft vollständiger metrischer Räume, die als *Satz von Baire* bekannt ist.[1] Es gibt mehrere äquivalente Varianten; wir benötigen hier die folgende.

Satz 5.1 (Baire). *Sei X ein vollständiger metrischer Raum und $\{A_n\}_{n\in\mathbb{N}}$ eine Folge von abgeschlossenen Teilmengen $A_n \subset X$. Enthält $A := \bigcup_{n\in\mathbb{N}} A_n$ einen inneren Punkt, so existiert ein $j \in \mathbb{N}$, so dass A_j einen inneren Punkt enthält.*

Beweis. Wir führen einen Widerspruchsbeweis. Angenommen, A enthält einen inneren Punkt, aber keines der A_n. Ersteres bedeutet, dass A eine offene Kugel $U_0 := U_{\varepsilon_0}(x_0)$ enthält; aus letzterem folgt $(X \setminus A_n) \cap U_\varepsilon(x) \neq \emptyset$ für alle $n \in \mathbb{N}$, $\varepsilon > 0$ und $x \in X$ (sonst würde ein A_n eine offene Kugel und damit innere Punkte enthalten).

Wir definieren nun induktiv eine Folge $\{B_{\varepsilon_n}(x_n)\}_{n\in\mathbb{N}}$ von geschachtelten abgeschlossenen Kugeln mit $\varepsilon_n < \frac{1}{n}$ wie folgt: Für $n = 1$ wählen wir $\varepsilon_1 < \min\{1, \varepsilon_0/2\}$ und $x_1 := x_0$. Haben wir x_n und ε_n gefunden, so wählen wir $\varepsilon_{n+1} < \frac{1}{n+1}$ und $x_{n+1} \in X$ so, dass gilt

$$B_{\varepsilon_{n+1}}(x_{n+1}) \subset (X \setminus A_{n+1}) \cap U_{\varepsilon_n}(x_n).$$

Dies ist möglich, da die Menge auf der rechten Seite nach Voraussetzung an A_n offen und nach Annahme nichtleer ist. Nach Konstruktion ist dann

$$U_{\varepsilon_{n+1}}(x_{n+1}) \subset B_{\varepsilon_{n+1}}(x_{n+1}) \subset U_{\varepsilon_n}(x_n) \subset B_{\varepsilon_n}(x_n).$$

Die Folge $\{x_n\}_{n\in\mathbb{N}}$ erfüllt also $x_m \in U_{\varepsilon_n}(x_n)$ für alle $m \geq n$ und ist daher wegen $\varepsilon_n \to 0$ eine Cauchy-Folge, die aufgrund der Vollständigkeit von X einen Grenzwert $x \in X$ besitzt.

[1] Er wird in der Literatur auch als *Bairescher Kategoriensatz* bezeichnet; dieser Name basiert auf einer historischen Terminologie, die hier aber nicht relevant ist.

Wegen $x_n \in B_{\varepsilon_k}(x_k)$ für alle $k \le n$ gilt auch $x \in B_{\varepsilon_k}(x_k)$ für alle $k \in \mathbb{N}$ (gerade dafür haben wir abgeschlossene Kugeln für die Konstruktion gewählt). Damit gilt einerseits

$$x \in \bigcap_{k\in\mathbb{N}} B_{\varepsilon_k}(x_k) \subset \bigcap_{k\in\mathbb{N}}(X \setminus A_k) = X \setminus \left(\bigcup_{k\in\mathbb{N}} A_k\right) = X \setminus A$$

und damit $x \notin A$, andererseits $x \in B_{\varepsilon_1}(x_1) \subset U_0 \subset A$, womit wir den gewünschten Widerspruch erhalten. □

Der Satz von Baire garantiert nun eine besondere Verträglichkeit von algebraischer und topologischer Struktur in Banachräumen. Dazu definieren wir für eine Teilmenge A eines Vektorraums X das *algebraische Innere*

$$\operatorname{core} A := \{x \in A : \text{für alle } h \in X \text{ existiert } \delta_h > 0 \text{ mit } x + th \in A \text{ für alle } t \in [0, \delta_h]\}\,.$$

Für alle $x \in \operatorname{core} A$ kann man also in alle Richtungen zumindest eine kleine Strecke gehen, ohne A zu verlassen. Wir nennen weiterhin A *konvex*, falls für alle $x, y \in A$ und $t \in [0, 1]$ auch $tx + (1 - t)y \in A$ gilt. Eine konvexe Menge enthält also sämtliche Strecken zwischen zwei Punkten.

Lemma 5.2 (core–int). *Sei X ein Banachraum und $A \subset X$ abgeschlossen und konvex. Dann gilt* $\operatorname{core} A = \operatorname{int} A$.

Beweis. Die Inklusion $\operatorname{int} A \subset \operatorname{core} A$ folgt sofort aus der Definition von inneren Punkten in normierten Räumen: Ist $x \in A$ innerer Punkt, so existiert ein $U_\varepsilon(x) \subset A$, und für alle $h \in X$ und $t < \varepsilon\|h\|_X^{-1} =: \delta_h$ ist dann $x + th \in U_\varepsilon(x) \subset A$.

Sei umgekehrt $x \in \operatorname{core} A$ gegeben; wir nehmen der Einfachheit halber $x = 0$ an. (Der allgemeine Fall folgt daraus durch die Translationsinvarianz der Definitionen von core und int.) Nach Voraussetzung existiert dann für jedes $h \in X$ ein $n \in \mathbb{N}$ groß genug mit $n^{-1}h \in A$, d. h. es gilt $h \in nA$ für $n > 0$ klein genug. Also ist $\bigcup_{n\in\mathbb{N}}(nA) = X$ offen, und nach dem Satz 5.1 von Baire ist daher $\operatorname{int}(nA) \neq \emptyset$ für ein $n \in \mathbb{N}$, was nur möglich ist, falls $\operatorname{int} A \neq \emptyset$ ist.

Bleibt zu zeigen, dass $0 \in \operatorname{int} A$ liegt. Sei dazu $x \in A$ ein innerer Punkt, d. h. $U_\varepsilon(x) \subset A$ für ein $\varepsilon > 0$. Wegen $0 \in \operatorname{core} A$ existiert für $h = -x$ ein $\delta > 0$ mit $-\delta x \in A$. Aus der Konvexität von A folgt dann, dass mit $y \in U_\varepsilon(x) \subset A$ und $-\delta x \in A$ für $t = \frac{1}{1+\delta} < 1$ auch

$$z := t(-\delta x) + (1 - t)y = \frac{\delta}{1 + \delta}(y - x) \in A$$

liegt. Wegen $\|y - x\|_X \le \varepsilon$ folgt daraus $z \in U_r(0)$ für $r = \frac{\delta\varepsilon}{1+\delta}$, und alle Elemente in $U_r(0)$ lassen sich auf diese Weise erzeugen. Damit gilt $U_r(0) \subset A$, woraus $0 \in \operatorname{int} A$ folgt. □

Dieses Lemma ist unscheinbar aber dennoch von zentraler Bedeutung, denn es stellt die eingangs versprochene Verknüpfung zwischen algebraischen (core) und topologischen (int) Eigenschaften her. Es garantiert im wesentlichen, dass eine Eigenschaft, die für alle $x \in A$ gilt und bei Grenzwertbildung (und Konvexkombination) erhalten bleibt, *gleichmäßig* für alle $x \in A$ gelten muss. (Natürlich gibt es nichts geschenkt – die Abgeschlossenheit von A ist in der Praxis eine nichttriviale Forderung, die oft genug nicht gilt.)

In der Funktionalanalysis wird dieses *Prinzip der gleichmäßigen Beschränktheit* häufig in der folgenden Form angewendet.

Satz 5.3 (Banach–Steinhaus). *Sei X ein Banachraum und Y ein normierter Vektorraum. Gilt für eine Teilmenge $\mathcal{T} \subset L(X,Y)$*

$$\sup_{T\in\mathcal{T}} \|Tx\|_Y < \infty \qquad \text{für alle } x \in X,$$

so gilt

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

Beweis. Wir wenden das core–int-Lemma 5.2 an auf die Menge

$$A := \left\{x \in X : \sup_{T\in\mathcal{T}} \|Tx\|_Y \leq 1\right\}.$$

Wegen $A = \bigcap_{T\in\mathcal{T}} T^{-1}(B_Y)$ und der Stetigkeit von $T \in L(X,Y)$ ist A abgeschlossen. Aus der Linearität von T^{-1} und den Normaxiomen folgt, dass $T^{-1}(B_Y)$ konvex ist; damit ist auch A als Schnitt konvexer Mengen konvex. Schließlich ist nach Voraussetzung $\sup_{T\in\mathcal{T}} \|Tx\|_Y < \infty$ für alle $x \in X$; somit existiert wegen der Linearität von T und der Homogenität der Norm für alle $h \in X$ ein $\delta > 0$ mit $\|T(\delta h)\|_Y \leq 1$ für alle $T \in \mathcal{T}$, woraus $\delta h \in A$ und damit $0 \in \operatorname{core} A$ folgt. Nach dem core–int-Lemma 5.2 ist daher $0 \in \operatorname{int} A$, d. h. es existiert ein $\varepsilon > 0$ mit $U_\varepsilon(0) \subset A$. Dies bedeutet aber, dass für alle $T \in \mathcal{T}$ aus $\|x\|_X < \varepsilon$ stets $\|Tx\|_Y \leq 1$ folgt. Wegen Lemma 4.3 (i) gilt daher für alle $T \in \mathcal{T}$

$$\|T\|_{L(X,Y)} = \sup_{x\in X, \|x\|_X<1} \|Tx\|_Y = \sup_{x\in X, \|x\|_X<1} \frac{1}{\varepsilon}\|T(\varepsilon x)\|_Y \leq \frac{1}{\varepsilon}$$

und damit die behauptete gleichmäßige Beschränktheit. □

Daraus folgt ein sehr nützliches Resultat: Bereits der *punktweise* Grenzwert einer Folge stetiger Operatoren ist stetig.

Folgerung 5.4. *Seien X ein Banachraum, Y ein normierter Raum, und $\{T_n\}_{n\in\mathbb{N}} \subset L(X,Y)$. Existiert für alle $x \in X$ der Grenzwert $Tx := \lim_{n\to\infty} T_n x \in Y$, so wird durch $x \mapsto Tx$ ein $T \in L(X,Y)$ definiert.*

Beweis. Die Linearität von T zeigt man wie im Beweis von Satz 4.5; bleibt die Stetigkeit. Aus der Konvergenz $T_n x \to Tx \in Y$ für alle $x \in X$ folgt direkt, dass die Folge $\{T_n x\}_{n\in\mathbb{N}}$ in Y beschränkt ist, d. h. $\sup_{n\in\mathbb{N}} \|T_n x\|_Y < \infty$. Der Satz 5.3 von Banach–Steinhaus liefert dann $\sup_{n\in\mathbb{N}} \|T_n\|_{L(X,Y)} < \infty$, und aus der Stetigkeit der Norm (Satz 3.6 (iii)) folgt

$$\|Tx\|_Y = \lim_{n\to\infty} \|T_n x\|_Y \leq \sup_{n\in\mathbb{N}} \|T_n x\|_Y \leq \sup_{n\in\mathbb{N}} \|T_n\|_{L(X,Y)} \|x\|_X \qquad \text{für alle } x \in X.$$

Also ist T beschränkt und daher stetig. □

Aus dem Prinzip der gleichmäßigen Beschränktheit folgen auch drei Hauptsätze über stetige Operatoren: die Sätze von der offenen Abbildung, der stetigen Inversen und dem abgeschlossenen Graphen. Dabei sind diese Sätze äquivalent: Man kann aus jedem jeweils die anderen direkt herleiten. Welchen man also aus dem Prinzip der gleichmäßigen Beschränktheit ableitet, ist Geschmackssache; üblicherweise beginnt man mit dem Satz von der offenen Abbildung.

Dafür nennen wir eine Abbildung $f : X \to Y$ zwischen zwei metrischen Räumen X und Y *offen*, wenn für alle offene Mengen $U \subset X$ auch $f(U) \subset Y$ offen ist. Der Satz von der offenen Abbildung sagt nun, dass ein stetiger linearer Operator zwischen Banachräumen genau dann offen ist, wenn er surjektiv ist. Wir bezeichnen hier mit U_X und U_Y die *offenen* Einheitskugeln in X bzw. Y.

Satz 5.5 (von der offenen Abbildung). *Seien X, Y Banachräume und $T \in L(X,Y)$. Dann sind äquivalent:*

(i) T ist offen;

(ii) Es gibt ein $\delta > 0$ mit $\delta U_Y \subset T(U_X)$;

(iii) T ist surjektiv.

Beweis. Wir zeigen, dass sowohl (i) und (ii) als auch (ii) und (iii) äquivalent sind.

(i) ⇒ (ii) folgt direkt aus der Definition von offenen Mengen und Abbildungen: Da T und U_X offen sind, ist auch $T(U_X)$ offen; also existiert insbesondere eine offene Kugel um $0 = T0 \in T(U_X)$.

(ii) ⇒ (i): Sei $U \subset X$ offen und $y \in T(U)$ beliebig. Wir zeigen, dass y ein innerer Punkt ist. Wähle dazu $x \in U$ mit $Tx = y$ sowie $\varepsilon > 0$ mit $U_\varepsilon(x) \subset U$. Mit $\delta > 0$ aus (ii) und der Linearität von T folgt dann

$$U_{\delta\varepsilon}(y) = U_{\delta\varepsilon}(Tx) = Tx + \varepsilon\delta U_Y \subset Tx + \varepsilon T(U_X) = T(U_\varepsilon(x)) \subset T(U),$$

d. h. $y \in \operatorname{int} T(U)$, und damit ist $T(U)$ offen.

(ii) ⇒ (iii) folgt wiederum aus der Linearität von T: Für $y \in Y$ beliebig ist $\tilde{y} := \frac{\delta}{2}\|y\|_Y^{-1} y \in \delta U_Y$. Nach (ii) existiert also ein $\tilde{x} \in U_X$ mit $T\tilde{x} = \tilde{y}$, d. h. für $x := \frac{2}{\delta}\|y\|_Y \tilde{x}$ gilt $Tx = y$.

(iii) ⇒ (ii): Für diese Richtung benötigen wir das Prinzip der gleichmäßigen Beschränktheit. Wir setzen dafür $A := \operatorname{cl} T(U_X) \subset Y$. Dann ist A abgeschlossen und konvex (denn mit U_X ist wegen der Linearität auch $T(U_X)$ konvex, und der Abschluss einer konvexen Menge ist auch konvex). Weiterhin existiert nach Voraussetzung für alle $h \in Y$ ein $x \in X$ mit $Tx = h$; für $\delta := \frac{1}{2}\|x\|_X^{-1} > 0$ ist dann $\delta x \in U_X$ und damit $\delta h = T(\delta x) \in A$. Da Y vollständig ist, können wir das core–int-Lemma 5.2 anwenden und erhalten $0 \in \operatorname{core} A = \operatorname{int} A$. Die Menge A enthält also eine offene Kugel um 0 mit Radius δ für ein $\delta > 0$.

Wir nutzen nun die Vollständigkeit von X, um $\delta U_Y \subset T(U_X)$ zu zeigen (gegebenenfalls für ein kleineres δ). Da wir gerade $\delta U_Y \subset A = \operatorname{cl} T(U_X)$ gezeigt haben, existiert für jedes $y \in Y$ mit $\|y\|_Y < \delta$ eine Folge $\{x_n\}_{n\in\mathbb{N}} \in U_X$ mit $Tx_n \to y$. Allerdings können wir daraus noch nicht schließen, dass $\{x_n\}_{n\in\mathbb{N}}$ selber konvergiert. Um das gesuchte Urbild zu erhalten, konstruieren wir daher eine neue Folge $\{\tilde{x}_n\}_{n\in\mathbb{N}} \in U_X$ mit $\tilde{x}_n \to x \in U_X$ und $Tx = y$. Dafür gehen wir wieder induktiv vor. Sei zunächst $y \in \delta U_Y$ gegeben und setze $y_0 := y \in \delta U_Y$. Sei nun $y_n \in \delta U_Y \subset \operatorname{cl} T(U_X)$ gegeben. Wir finden dann für $\varepsilon := \frac{\delta}{2} > 0$ ein $x_n \in U_X$ mit $\|y_n - Tx_n\|_Y < \varepsilon$, so dass

$$y_{n+1} := 2(y_n - Tx_n) \in \delta U_Y \subset A.$$

Es folgt

$$2^{-(n+1)} y_{n+1} = 2^{-n} y_n - T(2^{-n} x_n) \qquad \text{für alle } n \in \mathbb{N}.$$

Definieren wir $\tilde{x}_m := \sum_{n=0}^m 2^{-n} x_n$ für $m \in \mathbb{N}$ beliebig, so erhalten wir durch Auflösen der letzten Gleichung und Anwenden der Teleskopsumme

$$T\tilde{x}_m = y_0 - 2^{-(m+1)} y_{m+1} \to y_0 = y \qquad \text{für } m \to \infty, \tag{5.1}$$

denn $\{y_n\}_{n\in\mathbb{N}} \subset \delta U_Y$ ist beschränkt. Nun gilt wegen $\|x_n\|_X < 1$

$$\sum_{n=0}^m 2^{-n}\|x_n\|_X < 2 \qquad \text{für alle } m \in \mathbb{N}.$$

Durch Supremum über alle $m \in \mathbb{N}$ folgt, dass die zugehörige Reihe $\sum_{n=0}^\infty 2^{-n} x_n$ absolut konvergiert und damit aufgrund der Vollständigkeit von X und Lemma 3.7 gegen ein $x \in X$ konvergiert. Wegen der Abgeschlossenheit der Einheitskugel gilt sogar $x \in 2B_X$, und aus (5.1) folgt $Tx = y \in \delta U_Y$. Durch Skalierung $\delta \mapsto \delta/4$ folgt daraus die Existenz von $x \in \frac{1}{2}B_X \subset U_X$ mit $Tx = y$ für beliebige $y \in \delta U_Y$, d. h. $\delta U_Y \subset T(U_X)$. □

Beachten Sie, dass nur im letzten Schritt Stetigkeit von T und Vollständigkeit von X und Y benötigt wurde; jede offene lineare Abbildung zwischen normierten Räumen ist also surjektiv. Der letzte Schritt liefert dagegen eine sehr starke Aussage: Existiert für $y \in Y$ (klein genug) ein Urbild, so hat y sogar ein *beschränktes* Urbild! (Da T nicht als injektiv angenommen wurde, müssen das nicht die selben Urbilder sein.) Für lineare Operatoren ist „klein genug“ natürlich keine Einschränkung, da wir immer skalieren können. Wir erhalten daraus das folgende zentrale Resultat.

Satz 5.6 (von der stetigen Inversen). *Seien X, Y Banachräume und $T \in L(X, Y)$ bijektiv. Dann ist $T^{-1} \in L(Y, X)$.*

Beweis. Ist T bijektiv, dann ist T insbesondere surjektiv und damit nach dem Satz 5.5 von der offenen Abbildung offen. Also gilt für jede offene Menge $U \subset X$, dass $T(U) = (T^{-1})^{-1}(U)$ offen ist, d. h. die Urbilder von offenen Mengen unter T^{-1} sind offen. Damit ist T^{-1} nach Satz 1.8 stetig. □

Ist T nicht surjektiv, so möchte man zumindest die stetige Invertierbarkeit auf dem Bild erhalten. Das folgende nützliche Resultat zeigt, wann dies möglich ist.

Folgerung 5.7. *Seien X, Y Banachräume und $T \in L(X, Y)$ injektiv. Dann ist $T^{-1} : \operatorname{ran} T \to X$ stetig genau dann, wenn $\operatorname{ran} T$ abgeschlossen ist.*

Beweis. Ist $\operatorname{ran} T \subset Y$ abgeschlossen, so ist $\operatorname{ran} T$ nach Lemma 3.5 ein Banachraum und damit hat nach Satz 5.6 die Einschränkung $T : X \to \operatorname{ran} X$ eine stetige Inverse. Ist umgekehrt T^{-1} stetig, so ist T ein Isomorphismus von X nach $\operatorname{ran} T$. Da X ein Banachraum ist, muss daher auch $\operatorname{ran} T$ ein Banachraum und damit nach Lemma 3.5 abgeschlossen sein. □

Wieder ist die Forderung der Abgeschlossenheit nichttrivial; tatsächlich stellt die Tatsache, dass $\operatorname{ran} T$ nicht abgeschlossen sein muss, eine der wesentlichen Schwierigkeiten der Analysis in unendlichdimensionalen Räumen dar. Zum Beispiel ist in diesem Fall auch für $y \in \operatorname{ran} T$ die Gleichung $Tx = y$ nicht stabil lösbar; man spricht dann von einem *schlecht gestellten* Problem. Solche Probleme treten in der medizinischen Bildgebung (z. B. in der Computertomographie) und der Parameteridentifizierung auf; man spricht auch von *inversen Problemen*. Für ihre Lösung werden spezielle *Regularisierungsverfahren* benötigt; siehe etwa [Engl, Hanke & Neubauer 1996].

Wir zeigen zuletzt ein zu Lemma 3.5 analoges Resultat für Operatoren. Dafür betrachten wir für $T : X \to Y$ den Graphen $\operatorname{graph} T = \{(x, Tx) : x \in X\} \subset X \times Y$, versehen mit der Produktnorm $\|(x, y)\|_{X\times Y} := \|x\|_X + \|y\|_Y$.

Satz 5.8 (vom abgeschlossenen Graphen). *Seien X, Y Banachräume und $T : X \to Y$ linear. Dann ist T stetig genau dann, wenn $\operatorname{graph} T$ abgeschlossen ist.*

Beweis. Ist $\{(x_n, y_n)\}_{n\in\mathbb{N}} \subset \operatorname{graph} T$ mit $(x_n, y_n) \to (x, y)$ in $X \times Y$, so gilt insbesondere $x_n \to x$ und $y_n \to y$, und aus der Stetigkeit von T folgt $Tx_n \to Tx$. Ebenso gilt $Tx_n = y_n \to y$, und die Eindeutigkeit des Grenzwertes impliziert dann $Tx = y$, d. h. $(x, y) \in \operatorname{graph} T$.

Man verifiziert leicht, dass für einen linearen Operator $\operatorname{graph} T$ ein Unterraum von $X \times Y$ ist. Ist $\operatorname{graph} T$ abgeschlossen, dann ist $\operatorname{graph} T$ daher nach Lemma 3.5 ein Banachraum. Die *kanonischen Projektionen*

$$P_X : X \times Y \to X, \quad (x, y) \mapsto x, \qquad P_Y : X \times Y \to Y, \quad (x, y) \mapsto y,$$

sind daher linear und stetig (durch die Wahl der Produktnorm). Weiterhin ist die Einschränkung von P_X auf $\operatorname{graph} T$ bijektiv, hat also nach Satz 5.6 eine stetige Inverse $Q := (P_X|_{\operatorname{graph} T})^{-1} : X \to \operatorname{graph} T$ mit $Qx = (x, Tx)$. Nun gilt für alle $x \in X$

$$Tx = P_Y(x, Tx) = P_Y Qx.$$

Also ist $T = P_Y \circ Q$ nach Folgerung 4.6 stetig. □

Um zu zeigen, dass eine lineare Abbildung zwischen Banachräumen stetig ist, reicht es also zu zeigen, dass aus $x_n \to x$ und $Tx_n \to y$ bereits $Tx = y$ folgt. Man darf also bereits verwenden, dass Tx_n konvergiert, was beim Nachweis der Definition der Stetigkeit mit zu zeigen ist. Für allgemeine (nichtlineare) Abbildungen ist dies eine schwächere Eigenschaft, die aber in vielen Fällen einen ausreichenden Ersatz für die Stetigkeit darstellt. Solche Abbildungen nennt man *abgeschlossen*.

6 QUOTIENTENRÄUME

Nach dem Hauptsatz der Differential- und Integralrechung besitzt jede stetige Funktion eine Stammfunktion, die bis auf eine Konstante eindeutig ist. Analog ist die Ableitung als linearer Operator „bis auf eine Konstante“ injektiv. Ist man nun an dieser Konstante nicht wirklich interessiert oder möchte Aussagen gesammelt für alle Konstanten treffen, dann ist die in diesem Kapitel betrachtete Formalisierung der „Betrachtung bis auf eine Konstante“ nützlich. (Wir werden sie insbesondere für den Beweis eines weiteren Hauptsatzes, Satz 9.10 vom abgeschlossenen Bild, benötigen.)

Wir beginnen mit einer allgemeinen Konstruktion[1] und definieren dafür eine *Halbnorm* auf einem Vektorraum X als eine Abbildung $|\cdot| : X \to \mathbb{R}^+$, die die Eigenschaften (ii) und (iii) in Definition 3.1 erfüllt (d. h. $|x| = 0$ für $x \neq 0$ ist erlaubt). Man rechnet leicht nach, dass durch $x \sim y$ falls $|x - y| = 0$ eine Äquivalenzrelation definiert wird. Die zugehörigen Äquivalenzklassen

$$[x] := \{y \in X : x \sim y\}$$

bilden den *Quotientenraum*

$$X/\sim := \{[x] : x \in X\},$$

der mit den Operationen

$$[x] + [y] := [x + y], \qquad \lambda[x] := [\lambda x], \qquad x, y \in X, \ \lambda \in \mathbb{K},$$

zum Vektorraum wird. Diesen Vektorraum versehen wir nun mit der *Quotientennorm*

$$\|[x]\|_\sim := |x|, \qquad [x] \in X/\sim.$$

Satz 6.1. *Durch die Quotientennorm wird $(X/\sim, \|\cdot\|_\sim)$ zu einem normierten Raum. Konvergiert in X jede Cauchy-Folge bezüglich $|\cdot|$, so ist $(X/\sim, \|\cdot\|_\sim)$ ein Banachraum.*

Beweis. Wir zeigen zuerst, dass die Quotientennorm als Abbildung wohldefiniert (d. h. unabhängig von der Wahl des Repräsentanten) ist. Dafür betrachte $x, y \in X$ mit $[x] = [y]$, d. h. $|x - y| = 0$; aus der Dreiecksungleichung für die Halbnorm folgt dann

$$|x| \leq |x - y| + |y| = |y| \leq |y - x| + |x| = |x|,$$

[1] die auch in der Konstruktion der $L^p(\Omega)$-Räume zum Einsatz kommt

d. h. $\|[x]\|_\sim = |x| = |y| = \|[y]\|_\sim$.

Homogenität und Dreiecksungleichung für $\|\cdot\|_\sim$ folgen aus den entsprechenden Eigenschaften für die Halbnorm. Für die Nichtdegeneriertheit sei $\|[x]\|_\sim = 0$; dann ist nach Definition $|x-0| = |x| = 0$ und damit $[x] = [0]$. Also ist $(X/\sim, \|\cdot\|_\sim)$ ein normierter Vektorraum.

Wir zeigen nun die Vollständigkeit. Sei $\{[x]_n\}_{n\in\mathbb{N}} \subset X/\sim$ eine Cauchy-Folge, und wähle $x_n \in X$ mit $[x_n] = [x]_n$ für alle $n \in \mathbb{N}$. Nach Definition der Quotientennorm und den Rechenregeln für Äquivalenzklassen gilt dann

$$|x_n - x_m| = \|[x_n - x_m]\|_\sim = \|[x]_n - [x]_m\|_\sim \to 0 \qquad \text{für } m \geq n \to \infty.$$

Also ist auch $\{x_n\}_{n\in\mathbb{N}} \subset X$ eine Cauchy-Folge bezüglich $|\cdot|$ und konvergiert nach Voraussetzung gegen ein $x \in X$ in $(X, |\cdot|)$. Daraus folgt analog

$$\|[x]_n - [x]\|_\sim = \|[x_n - x]\|_\sim = |x_n - x| \to 0,$$

und damit $[x]_n \to [x]$ in $X/\sim$. □

Man kann diese Konstruktion auch anders betrachten. Aus den Halbnormeigenschaften folgt direkt, dass $U := \{x \in X : |x| = 0\}$ ein Unterraum von X ist; nach Definition gilt damit $x \sim y$ genau dann, wenn $x - y \in U$ ist. (Man kann sagen, dass der Quotientenraum durch „Faktorisieren" von X durch U entstanden ist.) Umgekehrt lässt sich dadurch beliebigen Unterräumen ein Quotientenraum zuordnen. Sei X ein normierter Raum, $U \subset X$ ein Unterraum, und definiere für $x \in X$ den *Abstand* von x zu U,

$$d_U(x) := \inf_{u\in U} \|x - u\|.$$

Wir betrachten nun den zu $|x| := d_U(x)$ gehörigen Quotientenraum

$$X/U := \{[x] : x \in X\}, \qquad [x] := \{y \in X : x - y \in U\},$$

versehen mit $\|[x]\|_U := d_U(x)$.

Satz 6.2. *(i) Ist X ein normierter Vektorraum und $U \subset X$ ein Unterraum, dann ist $\|\cdot\|_U$ eine Halbnorm auf X/U.*

(ii) Ist U abgeschlossen, dann ist auch $(X/U, \|\cdot\|_U)$ ein normierter Vektorraum.

(iii) Ist zusätzlich X ein Banachraum, dann ist auch $(X/U, \|\cdot\|_U)$ ein Banachraum.

Beweis. Zu (i): Zunächst ist $\|\cdot\|_U$ wohldefiniert, denn aus $[x] = [y]$ folgt $x - y \in U$ und damit $y = x - v$ für ein $v \in U$. Also ist

$$d_U(y) = d_U(x - v) = \inf_{u\in U} \|x - (v + u)\| = \inf_{\tilde{u}\in U} \|x - \tilde{u}\| = d_U(x),$$

da mit u auch $\tilde{u} := v + u$ den ganzen Unterraum durchläuft.

Mit dem selben Argument für $\tilde{u} := \lambda u$ folgt auch die Homogenität. Für die Dreiecksungleichung verwenden wir, dass nach Definition des Infimums für festes $x \in X$ und jedes $\varepsilon > 0$ ein $u_\varepsilon \in U$ existiert mit

$$\|x - u_\varepsilon\| < \inf_{u \in U} \|x - u\| + \varepsilon = d_U(x) + \varepsilon.$$

Seien nun $x, y \in X$ und $\varepsilon > 0$ beliebig, und wähle für x und y entsprechend u_ε bzw. v_ε. Dann gilt wegen $u_\varepsilon + v_\varepsilon \in U$

$$\begin{aligned} d_U(x+y) = \inf_{u \in U} \|(x+y) - u\| &\leq \|(x+y) - (u_\varepsilon + v_\varepsilon)\| \leq \|x - u_\varepsilon\| + \|y - v_\varepsilon\| \\ &< d_U(x) + d_U(y) + 2\varepsilon. \end{aligned}$$

Da $\varepsilon > 0$ beliebig war, folgt

$$\|[x] + [y]\|_U = d_U(x+y) \leq d_U(x) + d_U(y) = \|[x]\|_U + \|[y]\|_U,$$

und damit ist $\|\cdot\|_U$ eine Halbnorm auf X/U.

Zu (ii): Sei nun U abgeschlossen und $[x] \in X/U$ mit $\|[x]\|_U = d_U(x) = \inf_{u \in U} \|x - u\| = 0$. Dann existiert eine Folge $\{u_n\}_{n \in \mathbb{N}} \subset U$ mit $\|u_n - x\| \to 0$, d. h. $u_n \to x$. Aus der Abgeschlossenheit von U folgt nun $x \in U$ und damit $[x] = 0$. Also ist $\|\cdot\|_U$ sogar eine Norm auf X/U.

Zu (iii): Wir verwenden nun Satz 6.1 und zeigen, dass für einen Banachraum X jede Cauchy-Folge bezüglich d_U konvergiert. Sei dafür $\{x_n\}_{n \in \mathbb{N}}$ eine Cauchy-Folge bezüglich d_U. Zunächst konstruieren wir daraus eine Cauchy-Folge bezüglich der Norm in X wie folgt: Wie oben finden wir zu $x_{n+1} - x_n$ und $\varepsilon := 2^{-n}$ ein $u_{n+1} \in U$ mit

$$\|x_{n+1} - x_n - u_{n+1}\| < d_U(x_{n+1} - x_n) + 2^{-n} \qquad \text{für alle } n \in \mathbb{N}.$$

Da $\{x_n\}_{n \in \mathbb{N}}$ eine Cauchy-Folge ist, können wir eine Teilfolge auswählen (die wir der Übersichtlichkeit halber ebenfalls mit $\{x_n\}_{n \in \mathbb{N}}$ bezeichnen) mit

$$d_U(x_{n+1} - x_n) \leq 2^{-n} \qquad \text{für alle } n \in \mathbb{N}.$$

Setzen wir nun $z_n := x_n - \sum_{k=2}^{n} u_k$, dann gilt für alle $m = n + p \geq n \in \mathbb{N}$

$$\begin{aligned} \|z_m - z_n\| = \left\|x_{n+p} - x_n - \textstyle\sum_{k=n+1}^{n+p} u_k\right\| &\leq \sum_{k=0}^{p-1} \|x_{n+k+1} - x_{n+k} - u_{n+k+1}\| \\ &\leq \sum_{k=0}^{p-1} 2^{-n-k+1} \leq 2^{-n+2}. \end{aligned}$$

Also ist $\{z_n\}_{n\in\mathbb{N}}$ eine Cauchy-Folge im Banachraum X und konvergiert daher gegen ein $z \in X$. Wegen $\sum_{k=2}^{n} u_k \in U$ für alle $n \in \mathbb{N}$ folgt daraus aber auch

$$d_U(x_n - z) = \inf_{u\in U} \|x_n - z - u\| \leq \|x_n - z - \textstyle\sum_{k=2}^{n} u_k\| = \|z_n - z\| \to 0,$$

d. h. $\{x_n\}_{n\in\mathbb{N}}$ konvergiert bezüglich d_U entlang der eingangs gewählten Teilfolge und damit (als Cauchy-Folge) auch entlang der gesamten Folge.

Da U abgeschlossen ist, gilt schließlich $U = \operatorname{cl} U = \{x \in X : d_U(x) = 0\}$, und die Behauptung folgt aus Satz 6.1. □

Ein Spezialfall, den wir im weiteren Verlauf öfters benötigen werden, ist $U = \ker T$ für ein $T \in L(X, Y)$. Anschaulich wird uns das erlauben, einen nichtinjektiven Operator „injektiv zu machen“.

Lemma 6.3. *Seien X, Y normierte Räume, $T \in L(X, Y)$, und $U \subset \ker T \subset X$ ein abgeschlossener Unterraum. Dann existiert genau ein $S \in L(X/U, Y)$ mit*

(i) $S[x] = Tx$ für alle $x \in X$;

(ii) $\|S\|_{L(X/U,Y)} = \|T\|_{L(X,Y)}$.

Für $U = \ker T$ ist S injektiv.

Beweis. Wir definieren

$$S : X/U \to Y, \qquad [x] \mapsto Tx.$$

Der Operator S ist wohldefiniert, da für $y \in [x]$ gilt $y - x \in U \subset \ker T$, d. h. $Ty - Tx = T(y - x) = 0$. Direkt aus der Definition folgt, dass S linear ist und $Tx = S[x]$ für alle $x \in X$ gilt. Weiter ist für alle $x \in X$ und $y \in [x]$

$$\|S[x]\|_Y = \|Tx\|_Y = \|Ty\|_Y \leq \|T\|_{L(X,Y)} \|y\|_X.$$

Da $y \in [x]$ genau dann gilt, wenn $y = x - v$ für ein $v \in U$ ist, folgt

$$\|S[x]\|_Y \leq \inf_{y\in[x]} \|T\|_{L(X,Y)} \|y\|_X = \|T\|_{L(X,Y)} \inf_{v\in U} \|x - v\|_X = \|T\|_{L(X,Y)} \|[x]\|_U$$

und damit $\|S\|_{L(X/U,Y)} \leq \|T\|_{L(X,Y)}$. Insbesondere ist $S \in L(X/U, Y)$. Umgekehrt ist für alle $x \in X$

$$\begin{aligned} \|Tx\|_Y = \|S[x]\|_{L(X,Y)} &\leq \|S\|_{L(X/U,Y)} \|[x]\|_{X/U} = \|S\|_{L(X/U,Y)} \inf_{u\in U} \|x - u\|_X \\ &\leq \|S\|_{L(X/U,Y)} \|x\|_X, \end{aligned}$$

d. h. $\|T\|_{L(X,Y)} \leq \|S\|_{L(X/U,Y)}$.

Ist $U = \ker T$ (welcher stets abgeschlossen ist), so folgt aus $0 = S[x] = Tx$ sofort $x \in \ker T = U$ und damit nach Definition der Äquivalenzklassen $[x] = [0]$. Also ist $\ker S = \{[0]\}$ und damit S injektiv. □

Aus dem Beweis (für $T = \mathrm{Id}_X$) folgt auch, dass für einen abgeschlossenen Unterraum $U \subsetneq X$ die *Quotientenabbildung*

$$Q : X \to X/U, \qquad x \mapsto [x],$$

linear und stetig ist mit Operatornorm $\|Q\|_{L(X,X/U)} = 1$. Es gilt dann $T = S \circ Q$.

Analog können wir die fehlende Surjektivität durch Einschränkung auf das Bild von T herstellen; für die *stetige* Invertierbarkeit ist natürlich wieder die Abgeschlossenheit des Bildes wesentlich.

Satz 6.4. *Seien X, Y Banachräume und $T \in L(X, Y)$ mit $\operatorname{ran} T$ abgeschlossen. Dann ist die durch $S \circ Q = T$ definierte Abbildung $S : X/\ker T \to \operatorname{ran} T$ ein Isomorphismus, d. h.*

$$X/\ker T \simeq \operatorname{ran} T.$$

Beweis. Nach Lemma 6.3 ist $S : X/\ker T \to \operatorname{ran} T = \operatorname{ran} S$ linear, stetig und (wegen der Einschränkung des Bildraums) bijektiv. Weiter ist $X/\ker T$ nach Satz 6.2 (iii) ein Banachraum. Da Y ein Banachraum und $\operatorname{ran} T$ nach Voraussetzung abgeschlossener Unterraum ist, ist $\operatorname{ran} T$ wegen Lemma 3.5 ebenfalls ein Banachraum. Aus dem Satz 5.6 von der stetigen Inversen folgt dann, dass S^{-1} ebenfalls stetig und damit S ein Isomorphismus ist. □

Folgerung 6.5. *Seien X, Y Banachräume und $T \in L(X, Y)$ surjektiv. Dann ist $X/\ker T \simeq Y$.*

Teil III

DUALRÄUME UND SCHWACHE KONVERGENZ

7 LINEARE FUNKTIONALE UND DUALRÄUME

Wir haben bereits gesehen, dass eine wesentliche Schwierigkeit in unendlichdimensionalen Räumen darin besteht, dass die Konvergenz in der Norm nicht äquivalent zu einer komponentenweisen Konvergenz ist; aus diesem Grund gelten zum Beispiel die nützlichen Folgerungen des Satz 2.5 von Heine–Borel (insbesondere der Satz von Bolzano–Weierstraß, Folgerung 2.6) nicht mehr. Da man diese aber nicht kampflos aufgeben will, suchen wir in diesem Teil einen Konvergenzbegriff, der die komponentenweise Konvergenz auf unendlichdimensionale Räume verallgemeinert.

Die Grundidee ist die folgende: Für einen Vektor $x = (x_1, \dots, x_n) \in \mathbb{K}^n$ ist die Komponentenabbildung $x \mapsto x_k \in \mathbb{K}$ für ein $1 \le k \le n$ linear (offensichtlich) und stetig (offensichtlich bezüglich $\|\cdot\|_1$ und damit bezüglich jeder Norm nach Satz 3.8). Analog betrachten wir daher für einen unendlichdimensionalen normierten $\mathbb{K}$-Vektorräume X stetige (das ist hier eine zusätzliche Forderung!) lineare Abbildungen von X nach $\mathbb{K}$. Den Raum $X^* := L(X, \mathbb{K})$ aller solcher Abbildungen nennt man *Dualraum* von X; die Elemente $x^* \in X^*$ heißen *stetige lineare Funktionale*. Da $\mathbb{K}$ vollständig ist, folgt aus Satz 4.5 sofort, dass X^*, versehen mit der Operatornorm

$$\|x^*\|_{X^*} = \sup_{x \in B_X} |x^*(x)|,$$

ein Banachraum ist. Statt $x^*(x)$ schreibt man auch oft

$$\langle x^*, x\rangle_X := x^*(x)$$

um zu betonen, dass die *duale Paarung* $\langle \cdot, \cdot \rangle_X : X^* \times X \to \mathbb{K}$ bilinear ist, d. h.

$$\langle \alpha x_1^* + x_2^*, \beta x_1 + x_2\rangle_X = \alpha\beta\langle x_1^*, x_1\rangle_X + \alpha\langle x_1^*, x_2\rangle_X + \beta\langle x_2^*, x_1\rangle + \langle x_2^*, x_2\rangle_X$$

für alle $x_1^*, x_2^* \in X^*$, $x_1, x_2 \in X$ und $\alpha, \beta \in \mathbb{K}$ gilt (selbst im Fall $\mathbb{K} = \mathbb{C}$!) Wir erinnern, dass nach Lemma 4.3 (iii) gilt

$$|\langle x^*, x\rangle_X| \le \|x^*\|_{X^*}\|x\|_X \qquad \text{für alle } x^* \in X^*,\ x \in X.$$

Offensichtlich ist ein Vektor $x \in \mathbb{K}^n$ eindeutig durch seine Komponenten bestimmt; die stetigen linearen Funktionale charakterisieren also in gewisser Weise den Raum $\mathbb{K}^n$ vollständig. Die Frage ist nun, ob unendlichdimensionale normierte Räume durch ihren Dualraum auf ähnliche Weise charakterisiert werden (insbesondere, ob der entsprechende – noch zu definierende – Konvergenzbegriff aussagekräftig ist). Die Antwort ist nicht offensichtlich und benötigt einen weiteren Hauptsatz der Funktionalanalysis, den wir im nächsten Kapitel beweisen werden.

Der Rest dieses Kapitels ist Beispielen gewidmet. Dies wird kompliziert dadurch, dass Funktionale durch ihre Wirkung auf Elemente von X festgelegt sind und in der Regel nicht explizit (d. h. ohne Bezug auf ein $x \in X$) angegeben werden können. Für einige typische Räume ist aber eine konkretere Darstellung möglich. Dafür zeigt man üblicherweise, dass der Dualraum (isometrisch) isomorph zu einem bereits bekannten Banachraum ist. Zum Beispiel ist aus der linearen Algebra bekannt, dass der algebraische Dualraum (d. h. der Vektorraum aller – nicht notwendigerweise stetigen – linearen Funktionale) von $\mathbb{K}^n$ wieder die Dimension n hat. Da nach Satz 4.8 alle endlichdimensionalen Räume gleicher Dimension isomorph sind, ist $(\mathbb{K}^n)^* \simeq \mathbb{K}^n$ (unabhängig von der gewählten Norm); einen isometrischen Isomorphismus erhalten wir als Nebeneffekt des nächsten Beispiels.

In unendlichdimensionalen Räumen ist die Angelegenheit natürlich komplizierter. Wir betrachten als erstes Beispiel wieder die Folgenräume $\ell^p(\mathbb{K})$ und $c_0(\mathbb{K})$.

Satz 7.1. *Sei $1 \leq p < \infty$ und q mit $\frac{1}{p} + \frac{1}{q} = 1$ (wobei $\frac{1}{\infty} := 0$). Dann ist die Abbildung*

$$T : \ell^q(\mathbb{K}) \to \ell^p(\mathbb{K})^*, \qquad \langle Tx, y\rangle_{\ell^p} = \sum_{k=1}^{\infty} x_k y_k, \tag{7.1}$$

ein isometrischer Isomorphismus.

Die selbe Abbildungsvorschrift vermittelt einen isometrischen Isomorphismus $\ell^1(\mathbb{K}) \cong c_0(\mathbb{K})^*$.

Beweis. Wir zeigen zunächst, dass T einen stetigen linearen Operator definiert. Sei dafür $x \in \ell^q(\mathbb{K})$ und $y \in \ell^p(\mathbb{K})$. Für $1 < p < \infty$ folgt aus der *Hölderschen Ungleichung*

$$\sum_{k=1}^{N} |x_k||y_k| \leq \left(\sum_{k=1}^{N} |x_k|^q\right)^{\frac{1}{q}} \left(\sum_{k=1}^{N} |y_k|^p\right)^{\frac{1}{p}} \leq \|x\|_q \|y\|_p \qquad \text{für alle } N \in \mathbb{N}.$$

Für $p = 1$ und $q = \infty$ folgt diese Ungleichung direkt aus der Definition der Supremumsnorm. Also ist die Reihe $\sum_{k=1}^{\infty} x_k y_k$ absolut konvergent und

$$|\langle Tx, y\rangle_{\ell^p}| \leq \sum_{k=1}^{\infty} |x_k y_k| \leq \|x\|_q \|y\|_p.$$

Für festes $x \in \ell^q(\mathbb{K})$ ist also $Tx \in L(\ell^p(\mathbb{K}), \mathbb{K}) = \ell^p(\mathbb{K})^*$ mit

$$\|Tx\|_{\ell^p(\mathbb{K})^*} \leq \|x\|_q \tag{7.2}$$

definiert. Insbesondere ist $\|Tx\|_{\ell^p(\mathbb{K})^*}$ endlich und damit T als Operator von $\ell^q(\mathbb{K})$ nach $\ell^p(\mathbb{K})^*$ wohldefiniert. Aus der Abbildungsvorschrift (7.1) folgt, dass T linear ist, und aus (7.2), dass T stetig ist.

Nun zur Bijektivität. Dafür verwenden wir die *Einheitsvektoren* $e_k \in \ell^p(\mathbb{K})$ (für $1 \le p \le \infty$ beliebig)

$$[e_k]_j = \begin{cases} 1 & \text{falls } k = j, \\ 0 & \text{sonst.} \end{cases}$$

Man sieht leicht, dass $\|e_k\|_p = 1$ für alle $1 \le p \le \infty$, was den Namen erklärt. Gilt nun $Tx = 0 \in \ell^p(\mathbb{K})^*$, so auch $0 = \langle Tx, e_k\rangle_{\ell^p} = x_k$ für alle $k \in \mathbb{N}$ und damit $x = 0$. Also ist T injektiv.

Für die Surjektivität sei $y^* \in \ell^p(\mathbb{K})^*$ gegeben; wir müssen nun ein $x \in \ell^q(\mathbb{K})$ finden mit $Tx = y^*$. Dafür gehen wir schrittweise vor, indem wir die Wirkung von Tx und y^* auf gegebene Vektoren vergleichen. Zunächst gilt für die Einheitsvektoren $\langle Tx, e_k\rangle_{\ell^p} = x_k$, weshalb auch $x_k = \langle y^*, e_k\rangle_{\ell^p}$ für alle $k \in \mathbb{N}$ gelten muss. Damit ist die Kandidaten-Folge $x = \{x_k\}_{k\in\mathbb{N}}$ eindeutig festgelegt; es bleibt zu zeigen, dass $x \in \ell^q(\mathbb{K})$ gilt. Dafür betrachte eine endliche Folge $y \in c_e(\mathbb{K})$, d. h. $y = \sum_{k=1}^N y_k e_k$ für ein $N \in \mathbb{N}$ und $y_k \in \mathbb{K}$, $1 \le k \le N$. Dann gilt wegen der Linearität von y^*

$$\langle y^*, y\rangle_{\ell^p} = \langle y^*, \textstyle\sum_{k=1}^N y_k e_k\rangle_{\ell^p} = \displaystyle\sum_{k=1}^N y_k \langle y^*, e_k\rangle_{\ell^p} = \sum_{k=1}^N y_k x_k. \tag{7.3}$$

Für $p = 1$ und $q = \infty$ folgt aus $x_k = \langle y^*, e_k\rangle_{\ell^1}$ sofort

$$\|x\|_\infty = \sup_{k\in\mathbb{N}} |x_k| = \sup_{k\in\mathbb{N}} |\langle y^*, e_k\rangle_{\ell^1}| \le \sup_{y\in B_{\ell^1}} |\langle y^*, y\rangle_{\ell^1}| = \|y^*\|_{\ell^1(\mathbb{K})^*} \tag{7.4}$$

wegen $\|e_k\|_1 = 1$, und damit $x \in \ell^\infty(\mathbb{K})$. Für $1 < p < \infty$ wählen wir konkret eine Folge $y \in c_e(\mathbb{K})$ durch

$$y_k := \begin{cases} |x_k|^{q-1}\sigma_k & \text{für } k \le N, \\ 0 & \text{sonst,} \end{cases}$$

mit $\sigma_k x_k = |x_k|$ und $|\sigma_k| = 1$, so dass für alle $1 \le k \le N$ gilt

$$x_k y_k = |x_k|^q = |x_k|^{p(q-1)} = |y_k|^p.$$

Daraus folgt

$$\sum_{k=1}^N |x_k|^q = \sum_{k=1}^N x_k y_k = \langle y^*, y\rangle_{\ell^p} \le \|y^*\|_{\ell^p(\mathbb{K})^*}\|y\|_p = \|y^*\|_{\ell^p(\mathbb{K})^*}\left(\sum_{k=1}^N |x_k|^q\right)^{\frac{1}{p}},$$

und durch Division durch den zweiten Term auf der rechten Seite und Verwenden von $1 - \frac{1}{p} = \frac{1}{q}$ erhalten wir

$$\left(\sum_{k=1}^N |x_k|^q\right)^{\frac{1}{q}} \le \|y^*\|_{\ell^p(\mathbb{K})^*}.$$

Grenzübergang $N \to \infty$ liefert nun

$$\|x\|_q \leq \|y^*\|_{\ell^p(\mathbb{K})^*} \tag{7.5}$$

und damit $x \in \ell^q(\mathbb{K})$. Bleibt zu zeigen, dass $\langle Tx, y\rangle_{\ell^p} = \langle y^*, y\rangle_{\ell^p}$ für alle $y \in \ell^p(\mathbb{K})$ gilt. Vergleich von (7.3) mit (7.1) ergibt $\langle Tx, y\rangle_{\ell^p} = \langle y^*, y\rangle_{\ell^p}$ für alle $y \in c_e(\mathbb{K})$. Da $c_e(\mathbb{K})$ ein dichter Unterraum von ℓ^p für $1 \leq p < \infty$ ist (siehe Satz 3.15), folgt mit Satz 4.7 auch $Tx = y^*$ (sonst wären Tx und y^* zwei unterschiedliche Fortsetzungen von $y^*|_{c_e}$, im Widerspruch zur Eindeutigkeit der Fortsetzung). Also ist T surjektiv und damit bijektiv.

Da $\ell^p(\mathbb{K})$ für $1 \leq p \leq \infty$ ein Banachraum ist, ist T nach dem Satz 5.6 von der stetigen Inversen sogar stetig invertierbar. Aus (7.2) und (7.4) bzw. (7.5) folgt dann

$$\|Tx\|_{\ell^p(\mathbb{K})^*} \leq \|x\|_q \leq \|y^*\|_{\ell^p(\mathbb{K})^*} = \|Tx\|_{\ell^p(\mathbb{K})^*}.$$

Also ist T ein isometrischer Isomorphismus.

Analog zeigt man (mit Hilfe der Wahl $y_k = \sigma_k$), dass $\ell^1(\mathbb{K}) \cong c_0(\mathbb{K})^*$ ist. □

Der obige Beweis (ohne Grenzübergang) zeigt auch die Isometrie $(\mathbb{R}^N, \|\cdot\|_p) \cong (\mathbb{R}^N, \|\cdot\|_q)^*$, da in diesem Fall stets $x \in \mathbb{R}^N$ unabhängig von der gewählten Norm gilt. Dagegen funktioniert der Beweis für $\ell^p(\mathbb{K})$ mit $p = \infty$ nicht (denn hier ist $c_e(\mathbb{K})$ *nicht* dicht!) Tatsächlich kann man zeigen, dass $\ell^\infty(\mathbb{K})^*$ strikt größer als $\ell^1(\mathbb{K})$ ist.

Mit einer ähnlichen Konstruktion sowie dem Einsatz von Resultaten aus der Maßtheorie zeigt man den folgenden Darstellungssatz; für den (technischen) Beweis sei auf die Literatur verwiesen, z. B. auf [Brokate & Kersting 2019, Satz 13.4].

Satz 7.2. *Sei $1 \leq p < \infty$ und q mit $\frac{1}{p} + \frac{1}{q} = 1$ (wobei $\frac{1}{\infty} := 0$). Dann ist die Abbildung*

$$T : L^q(\Omega) \to L^p(\Omega)^*, \qquad \langle Tf, g\rangle_{L^p} = \int_\Omega f(t)g(t)\,dt,$$

ein isometrischer Isomorphismus.

Wieder gilt, dass $L^\infty(\Omega)^*$ strikt größer als $L^1(\Omega)$ ist.[1]

Zum Abschluss betrachten wir noch den Dualraum von Quotientenräumen. Dafür benötigen wir etwas Notation. Für einen normierten Vektorraum X und Teilmengen $A \subset X$ sowie $B \subset X^*$ definieren wir die *Annihilatoren*

$$\begin{aligned} A^\perp &:= \{x^* \in X^* : \langle x^*, x\rangle_X = 0 \quad \text{für alle } x \in A\},\\ B_\perp &:= \{x \in X : \quad \langle x^*, x\rangle_X = 0 \quad \text{für alle } x^* \in B\}. \end{aligned}$$

[1]Ein verwandtes Resultat aus der Maßtheorie ist der Satz von Radon–Riesz, der den Dualraum von $C(\Omega)$ mit dem Raum der regulären Borelmaße identifiziert; siehe z. B. [Alt 2012, Satz 4.23].

Dies sind stets abgeschlossene Unterräume von X^* bzw. X, denn für jede Folge $\{x_n^*\}_{n\in\mathbb{N}} \subset A^\perp$ mit $x_n^* \to x^* \in X^*$ gilt für beliebiges $x \in A$

$$|\langle x^*, x\rangle_X| = |\langle x^* - x_n^*, x\rangle_X| \leq \|x_n^* - x^*\|_{X^*}\|x\|_X \to 0$$

und damit $x^* \in A^\perp$ (und analog für $B_\perp$). Weiterhin ist $X^\perp = \{0\}$ und $\{0\}^\perp = X^*$.

Wir zeigen nun, dass für abgeschlossene Unterräume $(X/U)^* \cong U^\perp$ ist.

Satz 7.3. *Sei X ein Banachraum und $U \subset X$ ein abgeschlossener Unterraum. Dann ist die Abbildung*

$$T : (X/U)^* \to U^\perp, \qquad \langle Tu^*, x\rangle_X = \langle u^*, [x]\rangle_{X/U} \quad \text{für alle } x \in X,$$

ein isometrischer Isomorphismus.

Beweis. Wir zeigen zuerst, dass T wohldefiniert ist. Sei $u^* \in (X/U)^*$ und $x \in U$ beliebig, dann gilt

$$\langle Tu^*, x\rangle_X = \langle u^*, [x]\rangle_{X/U} = \langle u^*, [0]\rangle_{X/U} = 0$$

und damit $Tu^* \in U^\perp$. Analog folgt die Injektivität von T.

Sei nun $x^* \in U^\perp$ gegeben, und definiere $u^* \in (X/U)^*$ durch $\langle u^*, [x]\rangle_{X/U} = \langle x^*, x\rangle_X$ für alle $x \in X$. Wegen $x^* \in U^\perp$ ist diese Definition unabhängig von der Wahl des Repräsentanten $x \in [x]$, und offensichtlich ist u^* ein lineares Funktional. Bleibt für die Surjektivität zu zeigen, dass u^* stetig ist. Sei dafür $[x] \in X/U$ beliebig. Dann gilt

$$|\langle u^*, [x]\rangle_{X/U}| = |\langle x^*, y\rangle_X| \leq \|x^*\|_{X^*}\|y\|_X \qquad \text{für alle } y \in [x],$$

und daher

$$|\langle u^*, [x]\rangle_{X/U}| \leq \|x^*\|_{X^*} \inf_{y\in[x]} \|y\|_X = \|x^*\|_{X^*} \inf_{u\in U} \|x-u\|_X = \|x^*\|_X \|[x]\|_{X/U}.$$

Also ist $u^* \in (X/U)^*$, und es gilt $\|u^*\|_{(X/U)^*} \leq \|x^*\|_X = \|Tu^*\|_X$. Damit ist T eine bijektive Abbildung zwischen Banachräumen (da $U^\perp \subset X^*$ abgeschlossen ist) und damit stetig invertierbar, d. h. ein Isomorphismus.

Auf der anderen Seite gilt für alle $x \in X$

$$|\langle Tu^*, x\rangle_X| = |\langle u^*, [x]\rangle_{X/U}| \leq \|u^*\|_{(X/U)^*}\|[x]\|_{X/U} \leq \|u^*\|_{(X/U)^*}\|x\|_X,$$

da die Quotientenabbildung $x \mapsto [x]$ Norm 1 hat. Zusammen folgt $\|u^*\|_{(X/U)^*} = \|Tu^*\|_X$, d. h. T ist eine Isometrie. $\square$

8 DER SATZ VON HAHN–BANACH

Wir kommen nun zu einem zweiten Fundamentalprinzip der Funktionalanalysis, das algebraische und topologische Begriffe verknüpft: Linearität und Stetigkeit sind verträglich in dem Sinn, dass man ein auf einem Unterraum definiertes Funktional derart auf den gesamten Raum fortsetzen kann, dass gleichzeitig Linearität und Beschränktheit erhalten bleiben. Ähnlich wie das Prinzip der gleichmäßigen Beschränktheit beruht auch dieses Prinzip auf einem abstrakten Resultat, diesmal über reellwertige sublineare Funktionen (oder äquivalent dazu, wie wir später sehen werden, über konvexe Mengen).

Satz 8.1 (Hahn–Banach). *Sei X ein Vektorraum über $\mathbb{R}$ und $p : X \to \mathbb{R}$* sublinear*, d. h.*

(i) $p(\lambda x) = \lambda p(x)$ für alle $x \in X$ und $\lambda \geq 0 \in \mathbb{R}$;

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

Sei weiterhin $U \subset X$ ein Unterraum und $f_0 : U \to \mathbb{R}$ linear mit $f_0(x) \leq p(x)$ für alle $x \in U$. Dann existiert eine lineare Fortsetzung *$f : X \to \mathbb{R}$ mit*

(i) $f(x) = f_0(x)$ für alle $x \in U$;

(ii) $f(x) \leq p(x)$ für alle $x \in X$.

Beweis. Der Beweis verläuft in zwei Schritten. Zunächst zeigen wir induktiv, dass f_0 auf einen Unterraum mit um eins größerer Dimension wie gewünscht erweitert werden kann. Im zweiten Schritt zeigen wir, dass dieser Prozess auch für unendlichdimensionale Räume zum Abschluss kommt.

Seien dafür U und f_0 wie vorausgesetzt gegeben. Wir wählen $x_1 \in X \setminus U$ und erweitern f_0 von U auf

$$U_1 := \{x + \lambda x_1 : x \in U, \lambda \in \mathbb{R}\}$$

durch $f_1(x + \lambda x_1) := f_0(x) + \lambda\alpha$ für ein geeignetes $\alpha \in \mathbb{R}$. Mit dieser Definition ist f_1 auf jeden Fall linear; bleibt noch α so zu wählen, dass $f_1 \leq p$ gilt. Aus der Linearität von f_0 und der Sublinearität von p folgt zunächst für alle $x, y \in U$

$$f_0(x) + f_0(y) = f_0(x + y) \leq p(x + y) \leq p(x - x_1) + p(x_1 + y).$$

Umsortieren ergibt

$$f_0(x) - p(x - x_1) \le p(x_1 + y) - f_0(y) \qquad \text{für alle } x, y \in U.$$

Diese Ungleichung bleibt erhalten, wenn wir das Supremum über alle $x \in U$ und das Infimum über alle $y \in U$ bilden. Wir finden daher ein $\alpha \in \mathbb{R}$ so, dass

$$\sup_{x \in U} f_0(x) - p(x - x_1) \le \alpha \le \inf_{y \in U} p(x_1 + y) - f_0(y) \tag{8.1}$$

gilt. Wir zeigen nun dass für dieses α in der Tat $f_1(y) \le p(y)$ für alle $y \in U_1$ gilt, d. h. für alle $y = x + \lambda x_1$ mit $x \in U$ und $\lambda \in \mathbb{R}$. Für $\lambda > 0$ erhalten wir zunächst aus der zweiten Ungleichung in (8.1) und Abschätzung des Infimums nach oben durch $x \in U$ beliebig

$$f_0(x) + \alpha \le p(x_1 + x) \qquad \text{für alle } x \in U.$$

Wegen $\lambda > 0$ folgt daraus mit der Sublinearität von p

$$f_1(x + \lambda x_1) = f_0(x) + \lambda\alpha = \lambda \left(f_0\left(\tfrac{x}{\lambda}\right) + \alpha\right) \le \lambda p\left(x_1 + \tfrac{x}{\lambda}\right) = p(x + \lambda x_1).$$

Für $\lambda < 0$ erhalten wir analog aus der ersten Ungleichung in (8.1) durch Abschätzung des Supremums nach unten

$$f_0(x) - \alpha \le p(x - x_1) \qquad \text{für alle } x \in U,$$

und damit wegen $-\lambda > 0$

$$f_1(x + \lambda x_1) = -\lambda \left(f_0\left(\tfrac{x}{-\lambda}\right) - \alpha\right) \le (-\lambda) p\left(\tfrac{x}{-\lambda} - x_1\right) = p(x + \lambda x_1).$$

Für $\lambda = 0$ ist schließlich $f_1(x) = f_0(x) \le p(x)$. Also gilt $f_1(y) \le p(y)$ für alle $y \in U_1$.

Ist X endlichdimensional, so können wir mit dieser Prozedur fortfahren, bis $U_n = X$ gilt; für separable Räume verwendet man zusätzlich vollständige Induktion. Im allgemeinen Fall brauchen wir aber schweres Gerät: Das *Zornsche Lemma*, welches äquivalent zum Auswahlaxiom (und zum Wohlordnungssatz) ist, und garantiert, dass eine nichtleere halbgeordnete Menge, für die jede vollständig geordnete Teilmenge nach oben beschränkt ist, ein maximales Element enthält.[1] Um es anzuwenden, definieren wir die Menge aller Fortsetzungen (im Sinne der Aussage des Satzes)

$$\mathcal{A} := \{(W, f) : W \supset U \text{ Unterraum}, \ f : W \to \mathbb{R} \text{ linear mit } f|_U = f_0, \ f \le p\}$$

und versehen $\mathcal{A}$ mit der Halbordnung

$$(W_1, f_1) \le (W_2, f_2) \qquad \text{genau dann, wenn} \qquad W_1 \subset W_2, \quad f_2|_{W_1} = f_1.$$

[1] Eine etwas unhandlich zu formulierende Aussage, was die Grundlage für den oft zitierten – Jerry L. Bona zugeschriebenen – Spruch liefert: „Das Auswahlaxiom ist offensichtlich wahr, das Wohlordnungsprinzip ist offensichtlich falsch, und wer weiß das schon beim Zornschen Lemma?“

Offensichtlich ist $(U, f_0) \in \mathcal{A}$ und damit $\mathcal{A} \neq \emptyset$. Sei nun $\mathcal{B} \subset \mathcal{A}$ eine vollständig geordnete Teilmenge, d. h. für alle $a, b \in \mathcal{B}$ gilt entweder $a \leq b$ oder $b \leq a$. Wir setzen

$$W_* := \bigcup_{(W,f)\in\mathcal{B}} W$$

sowie

$$f_* : W_* \to \mathbb{R}, \qquad f_*(x) = f(x) \quad \text{für alle } x \in W,\ (W, f) \in \mathcal{B}.$$

Da $\mathcal{B}$ vollständig geordnet bezüglich der Fortsetzung, ist diese Vorschrift nicht widersprüchlich (was nicht offensichtlich ist). Weiter ist W_* ein Unterraum (ebenfalls aufgrund der vollständigen Ordnung von $\mathcal{B}$) und f_* linear. Nach Konstruktion ist dann $(W_*, f_*) \geq (W, f)$ für alle $(W, f) \in \mathcal{B}$ und damit eine obere Schranke. Das Zornsche Lemma garantiert daher die Existenz eines maximalen Elements (U_*, f). Dann muss $U_* = X$ gelten, denn sonst könnten wir wie im ersten Schritt eine weitere Fortsetzung von f auf $U_1 \supsetneq U_*$ konstruieren, im Widerspruch zur Maximalität von (U_*, f). □

Die im Beweis garantierte Fortsetzung ist im Allgemeinen nicht eindeutig; für verschiedene Wahlen von α, die (8.1) erfüllen, erhalten wir unterschiedliche Fortsetzungen.

Der Nutzen dieses Resultats für die Funktionalanalysis besteht darin, dass Normen konvex und damit insbesondere sublinear sind.[2] Wir können also ein stetiges lineares Funktional durch Festlegung der Werte auf einem Unterraum definieren und *stetig* auf X fortsetzen. Dies garantiert dann, dass der Dualraum für unsere Zwecke genügend Elemente enthält.[3]

Satz 8.2 (Fortsetzungssatz von Hahn–Banach). *Sei X ein normierter Raum über $\mathbb{K}$, sei $U \subset X$ ein Unterraum und $u^* : U \to \mathbb{K}$ linear und stetig, d. h. $u^* \in U^*$. Dann existiert ein $x^* \in X^*$ mit $\langle x^*, x\rangle_X = \langle u^*, x\rangle_U$ für alle $x \in U$ und $\|x^*\|_{X^*} = \|u^*\|_{U^*}$.*

Beweis. Sei zunächst $\mathbb{K} = \mathbb{R}$. Wir definieren für gegebenes $u^* \in U^*$

$$p : X \to \mathbb{R}, \qquad p(x) = \|u^*\|_{U^*}\|x\|_X.$$

Dann ist p sublinear, und es gilt

$$\langle u^*, x\rangle_U \leq \|u^*\|_{U^*}\|x\|_X = p(x) \qquad \text{für alle } x \in U.$$

Nach dem Satz 8.1 von Hahn–Banach existiert daher eine lineare Fortsetzung $x^* : X \to \mathbb{R}$ mit $\langle x^*, x\rangle_X \leq p(x)$ für alle $x \in X$. Es bleibt zu zeigen, dass x^* stetig ist und die behauptete Norm besitzt. Dafür verwenden wir die Linearität von x^*: Für alle $x \in X$ gilt

$$-\langle x^*, x\rangle_X = \langle x^*, -x\rangle_X \leq p(-x) = \|u^*\|_{U^*}\|-x\|_X = \|u^*\|_{U^*}\|x\|_X,$$

[2] Da die Konvexität der Norm äquivalent zur Konvexität der durch sie definierten Kugeln ist, kann man den Satz von Hahn–Banach auch in *lokalkonvexen Räumen* verwenden; dies sind topologische Vektorräume, in denen – grob gesprochen – die durch die Topologie definierten Umgebungen konvex sind. Diese besitzen viele der in Folge gezeigten Eigenschaften; siehe etwa [Werner 2018, Kapitel VIII].

[3] In [Luxemburg & Väth 2001] wurde sogar gezeigt, dass der Satz von Hahn–Banach äquivalent ist zur Aussage, dass für jeden Banachraum $X \neq \{0\}$ auch $X^* \neq \{0\}$ ist.

woraus zusammen mit $\langle x^*, x\rangle_X \leq p(x)$ für alle $x \in X$ folgt

$$|\langle x^*, x\rangle_X| \leq \|u^*\|_{U^*}\|x\|_X \qquad \text{für alle } x \in X,$$

d. h. x^* ist stetig und es gilt $\|x^*\|_{X^*} \leq \|u^*\|_{U^*}$. Da x^* Fortsetzung ist, folgt sofort

$$\|u^*\|_{U^*} = \sup_{x\in U\setminus\{0\}} \frac{|\langle u^*, x\rangle_U|}{\|x\|_X} = \sup_{x\in U\setminus\{0\}} \frac{|\langle x^*, x\rangle_X|}{\|x\|_X} \leq \sup_{x\in X\setminus\{0\}} \frac{|\langle x^*, x\rangle_X|}{\|x\|_X} = \|x^*\|_{X^*}$$

und damit $\|x^*\|_{X^*} = \|u^*\|_{U^*}$.

Den Fall $\mathbb{K} = \mathbb{C}$ führen wir auf den ersten Fall zurück. Für gegebenes $u^* \in X^*$ nehmen wir den Realteil $u^*_{\mathbb{R}} := \operatorname{Re} u^*$, definiert durch $\langle u^*_{\mathbb{R}}, x\rangle_{U_{\mathbb{R}}} := \operatorname{Re}\langle u^*, x\rangle_U$ für alle $x \in U$. Damit ist $u^*_{\mathbb{R}}$ ein $\mathbb{R}$-lineares Funktional auf U und wegen

$$\|u^*_{\mathbb{R}}\|_{U^*_{\mathbb{R}}} := \sup_{u\in U\setminus\{0\}} \frac{|\operatorname{Re}\langle u^*, u\rangle_U|}{\|u\|_X} \leq \sup_{u\in U\setminus\{0\}} \frac{|\langle u^*, u\rangle_U|}{\|u\|_X} = \|u^*\|_{U^*}$$

stetig. Weiter gilt wegen der $\mathbb{C}$-Linearität von u^* und $\operatorname{Im}(x) = -\operatorname{Re}(ix)$

$$\langle u^*, x\rangle_U = \operatorname{Re}\langle u^*, x\rangle_U + i\operatorname{Im}\langle u^*, x\rangle_U = \langle u^*_{\mathbb{R}}, x\rangle_{U_{\mathbb{R}}} - i\langle u^*_{\mathbb{R}}, ix\rangle_{U_{\mathbb{R}}} \qquad \text{für alle } x \in U.$$

Sei nun $x^*_{\mathbb{R}}$ die $\mathbb{R}$-lineare Fortsetzung von $u^*_{\mathbb{R}}$ wie im ersten Fall konstruiert, und setze

$$\langle x^*, x\rangle_X := \langle x^*_{\mathbb{R}}, x\rangle_{X_{\mathbb{R}}} - i\langle x^*_{\mathbb{R}}, ix\rangle_{X_{\mathbb{R}}} \qquad \text{für alle } x \in X.$$

Dann ist $x^* : X \to \mathbb{C}$ eine $\mathbb{C}$-lineare Fortsetzung von u^*. Für die Stetigkeit wähle zu $x \in X$ ein $\sigma \in \mathbb{C}$ mit $|\sigma| = 1$, so dass $|\langle x^*, x\rangle_X| = \sigma\langle x^*, x\rangle_X \in \mathbb{R}$ gilt. Dann ist

$$|\langle x^*, x\rangle_X| = \sigma\langle x^*, x\rangle_X = \langle x^*, \sigma x\rangle_X = \langle x^*_{\mathbb{R}}, \sigma x\rangle_{X_{\mathbb{R}}} \leq \|x^*_{\mathbb{R}}\|_{X^*_{\mathbb{R}}}\|x\|_X,$$

woraus $\|x^*\|_{X^*} = \|x^*_{\mathbb{R}}\|_{X^*_{\mathbb{R}}} = \|u^*_{\mathbb{R}}\|_{U^*_{\mathbb{R}}} \leq \|u^*\|_{U^*}$ und damit wegen $\|u^*\|_{U^*} \leq \|x^*\|_{X^*}$ wie im ersten Fall $\|x^*\|_{X^*} = \|u^*\|_{U^*}$ folgt. □

Beachten Sie, dass U weder als dicht noch als abgeschlossen vorausgesetzt war, im Gegensatz zu Satz 4.7 (welcher dafür Eindeutigkeit der Fortsetzung garantiert).

Aus dem Satz von Hahn–Banach erhalten wir nun elegant eine Reihe sehr nützlicher Aussagen. Die nächste Resultat garantiert, dass der Dualraum X^* stets „fein" genug ist, um einzelne Elemente von X zu unterscheiden; dies erlaubt, einen normierten Raum vollständig durch seinen Dualraum zu charakterisieren.

Satz 8.3. *Sei X ein normierter Raum und $x \in X \setminus \{0\}$. Dann existiert ein* normierendes Funktional *$x^* \in X^*$ mit $\|x^*\|_{X^*} = 1$ und $\langle x^*, x\rangle_X = \|x\|_X$.*

Beweis. Sei $U = \{\lambda x : \lambda \in \mathbb{K}\}$ und $u^* : U \to \mathbb{K}$ definiert durch

$$\langle u^*, \lambda x\rangle_U = \lambda\|x\|_X \qquad \text{für alle } \lambda \in \mathbb{K}.$$

Insbesondere ist für $\lambda = 1$ dann $\langle u^*, x\rangle_U = \|x\|_X$. Weiterhin ist $|\langle u^*, \lambda x\rangle_X| = \|\lambda x\|_X$ für alle $\lambda \in \mathbb{K}$ und damit $\|u^*\|_{U^*} = 1$, und wir erhalten das gewünschte normierende Funktional, indem wir u^* mit Hilfe des Fortsetzungssatzes 8.2 auf X fortsetzen. □

Das normierende Funktional kann man auch als Verallgemeinerung des Vorzeichens in normierten Räumen auffassen.

Diese Charakterisierung illustrieren die folgenden Resultate.

Folgerung 8.4. *Sei X ein normierter Raum. Dann gilt*

$$\|x\|_X = \max_{x^* \in B_{X^*}} |\langle x^*, x\rangle_X| \qquad \textit{für alle } x \in X.$$

Beweis. Dies folgt wegen

$$|\langle x^*, x\rangle_X| \leq \|x^*\|_{X^*}\|x\|_X \leq \|x\|_X \qquad \text{für alle } x^* \in B_{X^*},\ x \in X,$$

mit Gleichheit für das normierende Funktional aus Satz 8.3. □

Im Gegensatz zur Definition der Operatornorm $\|x^*\|_{X^*}$ wird hier das Supremum also stets angenommen.

Folgerung 8.5. *Seien X ein normierter Raum, $U \subset X$ ein abgeschlossener Unterraum und $x_0 \in X \setminus U$. Dann existiert ein $x^* \in X^*$ mit $\langle x^*, x\rangle_X = 0$ für alle $x \in U$ und $\langle x^*, x_0\rangle_X \neq 0$.*

Beweis. Da U abgeschlossener Unterraum ist, ist X/U nach Satz 6.1 (ii) ein normierter Raum. Sei $Q : X \to X/U$, $x \mapsto [x]$, die zugehörige Quotientenabbildung. Wegen $x_0 \notin U$ gilt $Qx_0 \neq [0]$, und Satz 8.3 liefert ein $q^* \in (X/U)^*$ mit $\langle q^*, Qx_0\rangle_{X/U} = \|Qx_0\|_{X/U} \neq 0$. Dann hat $x^* := q^* \circ Q \in L(X, \mathbb{K})$ wegen $Qx = [0]$ für alle $x \in U$ die gewünschten Eigenschaften. □

Die Aussage $\langle x^*, x\rangle_X = 0$ für alle $x \in U$ kann man nach Definition des Annihilators auch kurz als $x^* \in U^\perp$ schreiben.

Folgerung 8.6. *Sei X ein normierter Raum und $U \subset X$ ein Unterraum. Dann sind äquivalent:*

(i) U ist dicht in X;

(ii) $U^\perp = \{0\}$.

Beweis. *(i) ⇒ (ii)*: Sei $x^* \in U^\perp$. Nach Voraussetzung ist dann $\langle x^*, x\rangle_X = 0$ für alle $x \in U$. Also ist x^* eine Fortsetzung des Nulloperators $0 \in L(U,\mathbb{K}) = U^*$ auf X. Nach Satz 4.7 ist für U dicht in X diese Fortsetzung eindeutig und normerhaltend; also folgt $\|x^*\|_{X^*} = \|x^*\|_{U^*} = 0$ und damit $x^* = 0$.

(ii) ⇒ (i): Angenommen, $\operatorname{cl} U \neq X$. Dann existiert ein $x_0 \in X \setminus (\operatorname{cl} U)$ und daher nach Folgerung 8.5 ein $x^* \in X^*$ mit $x^* \neq 0$ und $\langle x^*, x\rangle_X = 0$ für alle $x \in U \subset \operatorname{cl} U$, d. h. $0 \neq x^* \in U^\perp$. Durch Kontraposition folgt die Behauptung. □

Folgerung 8.7. *Sei X ein normierter Raum. Ist X^* separabel, dann ist auch X separabel.*

Beweis. Sei $U^* := \{x_n^* : n \in \mathbb{N}\} \subset X^*$ dicht in X^*. Nach Definition der Operatornorm existiert dann für jedes x_n^* ein $x_n \in B_X$ mit $\langle x_n^*, x_n\rangle_X \geq \frac{1}{2}\|x_n^*\|_{X^*}$. Wir zeigen nun mit Hilfe von Folgerung 8.6, dass der durch die x_n aufgespannte Unterraum $U \subset X$ dicht in X ist. Sei dafür $x^* \in U^\perp$. Dann folgt für alle x_n^* und x_n aus der Definition der Operatornorm, $x_n \in U$, und der umgekehrten Dreiecksungleichung

$$\begin{aligned} \|x^* - x_n^*\|_{X^*} &\geq |\langle x^* - x_n^*, x_n\rangle_X| = |\langle x_n^*, x_n\rangle_X| \geq \frac{1}{2}\|x_n^*\|_{X^*} \\ &\geq \frac{1}{2}\big|\|x^*\|_{X^*} - \|x^* - x_n^*\|_{X^*}\big|. \end{aligned} \tag{8.2}$$

Da U^* dicht in X^* liegt, gilt $\inf_{x_n^* \in U^*} \|x^* - x_n^*\|_{X^*} = 0$. Gehen wir daher auf beiden Seiten von (8.2) zum Infimum über alle $n \in \mathbb{N}$ über, erhalten wir $0 \geq \frac{1}{2}\|x^*\|_{X^*}$ und damit $x^* = 0$. Also ist U dicht in X. Da die rationalen Linearkombinationen der x_n eine abzählbare und dichte Teilmenge von U bilden, ist X separabel. □

Da nach Satz 3.15 der Raum $\ell^1(\mathbb{K})$, aber nicht $\ell^\infty(\mathbb{K})$, separabel ist, kann daher $\ell^1(\mathbb{K})$ nicht isomorph zu $\ell^\infty(\mathbb{K})^*$ sein; ebenso ist $L^1(\Omega)$ nicht isomorph zu $L^\infty(\Omega)^*$. Allerdings ist der Satz 8.1 von Hahn–Banach (wegen der Verwendung des Auswahlaxioms in Form des Zornschen Lemmas) nicht konstruktiv und ermöglicht damit nicht, ein explizites Element aus $\ell^\infty(\mathbb{K})^*$ anzugeben, das nicht in $\ell^1(\mathbb{K})$ repräsentierbar ist.[4]

Folgerung 8.8. *Sei X ein normierter Raum und $U \subset X$ ein Unterraum. Dann gilt $(U^\perp)_\perp = \operatorname{cl} U$.*

Beweis. Für beliebiges $x \in U$ gilt $\langle x^*, x\rangle_X = 0$ für alle $x^* \in U^\perp$, und damit ist nach Definition $x \in (U^\perp)_\perp$. Für $x \in \operatorname{cl} U \setminus U$ betrachte eine Folge $\{x_n\}_{n\in\mathbb{N}} \subset U$ mit $x_n \to x$. Da $U \subset (U^\perp)_\perp$ wie eben gezeigt ist und Annihilatoren stets abgeschlossen sind, gilt $x \in (U^\perp)_\perp$ und damit $\operatorname{cl} U \subset (U^\perp)_\perp$.

Sei nun $x \notin \operatorname{cl} U$. Nach Folgerung 8.5 existiert dann ein $x^* \in U^\perp$ mit $\langle x^*, x\rangle_X \neq 0$, d. h. $x \notin (U^\perp)_\perp$. □

[4]Tatsächlich ist dies beweisbar unmöglich, denn ersetzt man das Auswahlaxiom – bzw. die schwächste Version davon, die für den Satz von Hahn–Banach ausreicht – durch schwächere Axiome, kann man $\ell^1(\mathbb{K}) \cong \ell^\infty(\mathbb{K})^*$ zeigen, siehe [Schechter 1997, § 29.37–38].

Dagegen gilt für $U \subset X^*$ in der Regel nur $\operatorname{cl} U \subset (U_\perp)^\perp$.

Anfangs wurde bereits angedeutet, dass der Satz von Hahn–Banach auch eine geometrische Interpretation als Aussage über konvexe Mengen besitzt; diese betrachten wir nun näher. Folgerung 8.5 besagt insbesondere, dass für alle $x, y \in X$ mit $x \neq y$ ein $x^* \in X^*$ existiert mit $\langle x^*, x - y\rangle_X \neq 0$; man sagt, x^* *trennt* x und y. Anschaulich können wir für $x^* \in X^*$ und $\alpha \in \mathbb{R}$ (im Fall $\mathbb{K} = \mathbb{R}$) die *Hyperebene*

$$H_\alpha := \{x \in X : \langle x^*, x\rangle_X = \alpha\}$$

definieren. Die Aussage ist dann, dass für gegebene x und y mit $x \neq y$ stets eine Hyperebene existiert, die X in zwei *Halbräume* H_α^- und H_α^+ zerteilt mit

$$x \in H_\alpha^- := \{x \in X : \langle x^*, x\rangle_X < \alpha\}, \qquad y \in H_\alpha^+ := \{x \in X : \langle x^*, x\rangle_X > \alpha\}.$$

Wir verallgemeinern dies nun auf die Trennung von Mengen. Wie bei der Fortsetzung ist auch hier die Konvexität wesentlich.

Satz 8.9 (Trennungssatz von Hahn–Banach). *Seien X ein normierter Vektorraum, $A \subset X$ nichtleer, offen und konvex, und $x_0 \in X \setminus A$. Dann existiert ein $x^* \in X^*$ mit*

$$\operatorname{Re}\langle x^*, x\rangle_X < \operatorname{Re}\langle x^*, x_0\rangle_X \qquad \text{für alle } x \in A.$$

Beweis. Wir betrachten zuerst den Fall $\mathbb{K} = \mathbb{R}$, und führen die Aussage zurück auf den Satz 8.1 von Hahn–Banach. Wir benötigen also ein geeignetes sublineares Funktional. Für $A \subset X$ definieren wir dafür das *Minkowski-Funktional*

$$m_A : X \to [0, \infty], \qquad x \mapsto \inf\left\{t > 0 : \tfrac{1}{t}x \in A\right\}.$$

Das Funktional gibt an, wie weit man $x \in X$ in Richtung 0 „zurückziehen“ muss, bis $x \in A$ liegt. (Für $A = B_X$ ist $m_A(x) = \|x\|_X$, womit wir bereits den Zusammenhang zum Fortsetzungssatz sehen.) Erstmal ist nicht klar, ob überhaupt so ein t existiert, d. h. ob das Infimum endlich ist. Wir nehmen daher zuerst an, dass $0 \in A$ liegt, denn dann ist für alle $x \in X$ und t groß genug $\frac{1}{t}x \in A$, da A offen ist (und damit eine Kugel mit Radius $\varepsilon > \frac{1}{t}\|x\|_X$ enthält). Wir zeigen nun die Sublinearität. Direkt aus der Definition folgt $m_A(\lambda x) = \lambda m_A(x)$ für alle $x \in X$ und $\lambda > 0$. Seien nun $x, y \in X$ und $\varepsilon > 0$ beliebig. Aufgrund der Definition des Infimums existieren dann $t, s > 0$ mit

$$t \leq m_A(x) + \varepsilon, \quad \tfrac{1}{t}x \in A, \qquad s \leq m_A(y) + \varepsilon, \quad \tfrac{1}{s}y \in A.$$

Aus der Konvexität von A folgt dann

$$\frac{1}{t+s}(x+y) = \frac{t}{t+s}\left(\tfrac{1}{t}x\right) + \frac{s}{t+s}\left(\tfrac{1}{s}y\right) \in A$$

und damit

$$m_A(x+y) \le t+s \le m_A(x)+m_A(y)+2\varepsilon.$$

Da $\varepsilon > 0$ beliebig war, folgt $m_A(x+y) \le m_A(x)+m_A(y)$.

Wir zeigen nun, dass m_A selber x_0 und A trennt; im nächsten Schritt konstruieren wir daraus ein *lineares* Funktional auf einem geeigneten Unterraum. Zunächst gilt $\frac{1}{t}x_0 \notin A$ für alle $t < 1$ (sonst wäre $x_0 = t\frac{x_0}{t} + (1-t)0 \in A$ für ein $t \in (0,1)$, denn $0 \in A$ und A ist konvex) und damit

$$m_A(x_0) \ge 1. \tag{8.3}$$

Umgekehrt folgt aus $m_A(x) \ge 1$, dass $\frac{1}{t}x \in X \setminus A$ für alle $t \in (0,1)$ gilt. Wählen wir nun eine Folge $\{t_n\}_{n\in\mathbb{N}} \subset (0,1)$ mit $t_n \to 1$, so erhalten wir $x = \lim_{n\to\infty} \frac{1}{t_n}x \in X \setminus A$, denn $X \setminus A$ ist als Komplement einer offenen Menge abgeschlossen. Also gilt

$$m_A(x) < 1 \qquad \text{für alle } x \in A. \tag{8.4}$$

Wir definieren nun wie im Beweis von Satz 8.3 den Unterraum $U := \{\lambda x_0 : \lambda \in \mathbb{R}\}$ und auf U ein lineares Funktional $u^* : U \to \mathbb{R}$ durch

$$\langle u^*, \lambda x_0\rangle_U = \lambda m_A(x_0) \qquad \text{für alle } \lambda \in \mathbb{R}.$$

Dann gilt wegen der Sublinearität von m_A sowie $m_A \ge 0$ für

$$\begin{aligned} \lambda > 0: &\qquad \langle u^*, \lambda x_0\rangle_X = \lambda m_A(x_0) = m_A(\lambda x_0),\\ \lambda \le 0: &\qquad \langle u^*, \lambda x_0\rangle_X = \lambda m_A(x_0) \le 0 \le m_A(\lambda x_0). \end{aligned}$$

Also ist $u^* \le m_A$ auf U. Wir erhalten daher durch den Satz 8.1 von Hahn–Banach eine lineare Fortsetzung $x^* : X \to \mathbb{R}$ mit $x^* \le m_A$ auf ganz X. Bleibt zu zeigen, dass x^* stetig ist und ebenfalls trennt.

Da $0 \in A$ und A als offen vorausgesetzt sind, existiert für $\varepsilon > 0$ klein genug eine abgeschlossene Kugel $B_\varepsilon(0) \subset A$. Also ist $\varepsilon\frac{x}{\|x\|_X} \in B_\varepsilon(0) \subset A$ für alle $x \in X$ und damit

$$\langle x^*, x\rangle_X \le m_A(x) \le \frac{1}{\varepsilon}\|x\|_X \qquad \text{für alle } x \in X,$$

sowie analog für $-x$, d. h. $x^* \in X^*$. Aus (8.3) und (8.4) folgt nun

$$\langle x^*, x\rangle_X \le m_A(x) < 1 \le m_A(x_0) = \langle u^*, x_0\rangle_U = \langle x^*, x_0\rangle_X \qquad \text{für alle } x \in A,$$

woraus die Trennungseigenschaft folgt.

Für den allgemeinen Fall $0 \notin A$ wählen wir $\tilde{x} \in A$ und betrachten $\tilde{A} := A - \tilde{x} := \{x - \tilde{x} : x \in A\}$. Dann ist $\tilde{A}$ ebenfalls offen und konvex, $0 \in \tilde{A}$, und $x_0 - \tilde{x} \notin \tilde{A}$. Wie oben erhalten wir nun ein $x^* \in X^*$ mit $\langle x^*, y\rangle_X < \langle x^*, x_0 - \tilde{x}\rangle_X$ für alle $y \in \tilde{A}$. Für alle $x \in A$ ist dann nach Definition $x - \tilde{x} \in \tilde{A}$ und daher

$$\langle x^*, x\rangle_X = \langle x^*, x - \tilde{x}\rangle_X + \langle x^*, \tilde{x}\rangle_X < \langle x^*, x_0 - \tilde{x}\rangle_X + \langle x^*, \tilde{x}\rangle_X = \langle x^*, x_0\rangle_X,$$

was zu zeigen war.

Den Fall $\mathbb{K} = \mathbb{C}$ führt man wie im Beweis vom Fortsetzungssatz 8.2 auf den reellen Fall zurück. □

Aus Satz 8.9 kann man nun weitere Trennungssätze ableiten, die von fundamentaler Bedeutung für die konvexe Optimierung sind. Zuerst betrachten wir die Trennung zweier disjunkter konvexer Mengen.

Satz 8.10. *Sei X normierter Vektorraum, $A_1, A_2 \subset X$ nichtleer und konvex mit $A_1 \cap A_2 = \emptyset$. Ist A_1 offen, so existiert ein $x^* \in X^*$ und ein $\alpha \in \mathbb{R}$ mit*

$$\operatorname{Re}\langle x^*, x_1\rangle_X < \alpha \leq \operatorname{Re}\langle x^*, x_2\rangle_X \qquad \textit{für alle } x_1 \in A_1,\ x_2 \in A_2.$$

Beweis. Setze $A := A_1 - A_2 = \{x_1 - x_2 : x_1 \in A_1, x_2 \in A_2\}$. Dann ist A offen, denn für $x \in A$ folgt $x \in A_1 - x_2 \subset A$ für ein $x_2 \in A_2$, und $A_1 - x_2$ ist offen. Man rechnet auch leicht nach, dass A als Differenz konvexer Mengen ebenfalls konvex ist. Weiterhin ist $0 \notin A$ wegen $A_1 \cap A_2 = \emptyset$. Aus dem Trennungssatz 8.9 erhalten wir also ein $x^* \in X^* \setminus \{0\}$ mit $\operatorname{Re}\langle x^*, x\rangle_X < \operatorname{Re}\langle x^*, 0\rangle_X = 0$ für alle $x \in A$, d. h.

$$\operatorname{Re}\langle x^*, x_1\rangle_X - \operatorname{Re}\langle x^*, x_2\rangle_X = \operatorname{Re}\langle x^*, x_1 - x_2\rangle_X < 0 \qquad \text{für alle } x_1 \in A_1,\ x_2 \in A_2.$$

Mit $\alpha := \sup_{x_1 \in A_1} \operatorname{Re}\langle x^*, x_1\rangle_X$ folgt nun die Behauptung, zunächst noch möglicherweise mit Gleichheit. Um zu zeigen, dass die erste Ungleichung strikt ist, nehmen wir an, dass ein $x_1 \in A$ existiert mit $\langle x^*, x_1\rangle_X = \alpha$. Wegen $x^* \neq 0$ existiert dann ein $x_0 \in X \setminus \{0\}$ mit $\operatorname{Re}\langle x^*, x_0\rangle_X =: c \neq 0$; ohne Beschränkung der Allgemeinheit können wir $c > 0$ annehmen. Da A_1 offen ist, existiert weiter ein $\varepsilon > 0$ mit

$$\tilde{x}_1 := x_1 + \varepsilon \frac{x_0}{\|x_0\|_X} \in B_\varepsilon(x_1) \subset A_1.$$

Damit gilt

$$\operatorname{Re}\langle x^*, \tilde{x}_1\rangle_X = \operatorname{Re}\langle x^*, x_1\rangle_X + \frac{\varepsilon}{\|x_0\|_X} \operatorname{Re}\langle x^*, x_0\rangle_X = \alpha + \frac{\varepsilon c}{\|x_0\|_X} > \alpha,$$

im Widerspruch zu $\operatorname{Re}\langle x^*, \tilde{x}_1\rangle_X \leq \alpha$ für $\tilde{x}_1 \in A_1$. □

Eine weitere Variante ist die strikte Trennung eines Punktes von einer *abgeschlossenen* Menge.

Satz 8.11. *Seien X ein normierter Vektorraum, $A \subset X$ nichtleer, abgeschlossen und konvex, und $x_0 \in X \setminus A$. Dann existiert ein $x^* \in X^*$ und ein $\alpha \in \mathbb{R}$ mit*

$$\operatorname{Re} \langle x^*, x\rangle_X \leq \alpha < \operatorname{Re} \langle x^*, x_0\rangle_X \qquad \text{für alle } x \in A.$$

Beweis. Da A abgeschlossen ist, existiert eine offene (konvexe) Kugel $U_\varepsilon(x_0) \subset X \setminus A$. Wir wenden Satz 8.10 an auf $A_1 = U_\varepsilon(x_0)$ und $A_2 = A$ und erhalten ein trennendes Funktional $x^* \in X^*$ sowie $\alpha \in \mathbb{R}$ mit $\operatorname{Re} \langle x^*, x_1\rangle_X < \alpha \leq \operatorname{Re} \langle x^*, x_2\rangle_X$ für alle $x_1 \in U_\varepsilon(x_0)$ und $x_2 \in A$. Multiplikation mit -1 und Einsetzen von $x_1 = x_0$ ergibt dann die Aussage. □

9 ADJUNGIERTE OPERATOREN

Ähnlich wie ein normierter Vektorraum durch seinen Dualraum charakterisiert wird, wird auch ein linearer Operator durch einen „dualen" Operator charakterisiert. Für normierte Räume X, Y und einen stetigen linearen Operator $T \in L(X, Y)$ definieren wir den *adjungierten Operator*

$$T^* : Y^* \to X^*, \qquad \langle T^* y^*, x\rangle_X := \langle y^*, Tx\rangle_Y \qquad \text{für alle } x \in X,\ y^* \in Y^*.$$

Dann ist T^* nach Konstruktion linear. Weiterhin ist T^* stetig.

Lemma 9.1. *Sei $T \in L(X, Y)$. Dann ist $T^* \in L(Y^*, X^*)$ mit $\|T^*\|_{L(Y^*,X^*)} = \|T\|_{L(X,Y)}$.*

Beweis. Es gilt nach Definition der Operatornorm

$$\begin{aligned} \|T^*\|_{L(Y^*,X^*)} &= \sup_{y^*\in B_{Y^*}} \|T^* y^*\|_{X^*} = \sup_{y^*\in B_{Y^*}} \sup_{x\in B_X} |\langle T^* y^*, x\rangle_X| \\ &= \sup_{y^*\in B_{Y^*}} \sup_{x\in B_X} |\langle y^*, Tx\rangle_Y| = \sup_{x\in B_X} \sup_{y^*\in B_{Y^*}} |\langle y^*, Tx\rangle_Y| \\ &= \sup_{x\in B_X} \|Tx\|_Y = \|T\|_{L(X,Y)}, \end{aligned}$$

wobei wir für die vorletzte Gleichung Folgerung 8.4 verwendet haben. □

Beispiel 9.2. Ein triviales Beispiel ist die Identität $\mathrm{Id}_X \in L(X, X)$: Es gilt

$$\langle x^*, \mathrm{Id}_X\, x\rangle_X = \langle x^*, x\rangle_X = \langle \mathrm{Id}_{X^*}\, x^*, x\rangle_X \qquad \text{für alle } x \in X,\ x^* \in X^*,$$

d. h. $(\mathrm{Id}_X)^* = \mathrm{Id}_{X^*}$.

Für ein weniger triviales Beispiel definieren wir für $X = Y = \ell^p(\mathbb{K})$ mit $1 \le p < \infty$ den *Rechts-Shift*

$$S_+ : \ell^p(\mathbb{K}) \to \ell^p(\mathbb{K}), \qquad (x_1, x_2, x_3, \dots) \mapsto (0, x_1, x_2, \dots).$$

Da der adjungierte Operator $(S_+)^* : \ell^p(\mathbb{K})^* \to \ell^p(\mathbb{K})^*$ auf linearen Funktionalen operiert, kann er in der Regel nicht konkreter als über die Definition angegeben werden. Wegen $\ell^p(\mathbb{K})^* \cong \ell^q(\mathbb{K})$ mit $\frac{1}{p} + \frac{1}{q} = 1$ ist es aber möglich, eine explizite Darstellung

$T_p^{-1}(S_+)^*T_p : \ell^q(\mathbb{K}) \to \ell^q(\mathbb{K})$ zu finden, wobei $T_p : \ell^q(\mathbb{K}) \to \ell^p(\mathbb{K})^*$ der isometrische Isomorphismus aus Satz 7.1 ist. Sei dafür $y \in \ell^q(\mathbb{K})$ gegeben und $x \in \ell^p(\mathbb{K})$ beliebig. Dann gilt

$$\langle (S_+)^*T_p y, x\rangle_p = \langle T_p y, S_+x\rangle_p = \sum_{k=2}^{\infty} x_{k-1}y_k = \sum_{k=1}^{\infty} x_k y_{k+1} = \langle T_p z, x\rangle_p$$

für $z := (y_2, y_3, y_4, \dots) \in \ell^q(\mathbb{K})$, d. h. $(S_+)^*T_p y = T_p z \in \ell^p(\mathbb{K})^*$ und damit gilt $T_p^{-1}(S_+)^*T_p y = z$. Der adjungierte Operator kann daher dargestellt werden als der *Links-Shift*

$$S_- : \ell^q(\mathbb{K}) \to \ell^q(\mathbb{K}), \qquad (x_1, x_2, x_3, \dots) \mapsto (x_2, x_3, x_4, \dots).$$

Weitere Beispiele erhält man aus den folgenden Rechenregeln, die direkt aus der Definition folgen.

Lemma 9.3. *Seien X, Y, Z normierte Räume, $T_1, T_2 \in L(X, Y)$ und $S \in L(Y, Z)$. Dann gelten:*

(i) $(T_1 + T_2)^* = T_1^* + T_2^*$;

(ii) $(\lambda T_1)^* = \lambda T_1^*$ *für alle* $\lambda \in \mathbb{K}$;

(iii) $(S \circ T)^* = T^* \circ S^*$.

Eine besonders nützliche Eigenschaft ist, dass Adjungieren und Invertieren kommutieren.

Satz 9.4. *Sei $T \in L(X, Y)$ stetig invertierbar. Dann ist auch $T^* \in L(Y^*, X^*)$ stetig invertierbar, und es gilt*

$$(T^*)^{-1} = (T^{-1})^* =: T^{-*}.$$

Beweis. Ist T stetig invertierbar, so gilt $T^{-1}T = \mathrm{Id}_X$, $TT^{-1} = \mathrm{Id}_Y$, und T^{-1} ist stetig. Adjungieren auf beiden Seiten ergibt dann unter Verwendung von Lemma 9.3 (iii) $T^*(T^{-1})^* = \mathrm{Id}_{X^*}$ und analog für TT^{-1}. Also ist $(T^{-1})^*$ Inverse von T^* und nach Lemma 9.1 stetig. □

Folgerung 9.5. *Sind X und Y (isometrisch) isomorph, dann sind auch X^* und Y^* (isometrisch) isomorph.*

Beweis. Sind X und Y isomorph, so existiert nach Definition eine stetig invertierbare Abbildung $T : X \to Y$. Dann ist nach Satz 9.4 auch $T^* : Y^* \to X^*$ stetig invertierbar und damit Y^* und X^* isomorph.

Ist T zusätzlich isometrisch, dann ist auch $T : B_X \to B_Y$ ein Isomorphismus, und daraus folgt für alle $y^* \in Y^*$

$$\|T^* y^*\|_{X^*} = \sup_{x \in B_X} |\langle T^* y^*, x\rangle_X| = \sup_{x \in B_X} |\langle y^*, Tx\rangle_Y| = \|y^*\|_{Y^*}.$$

Also ist auch T^* isometrisch. □

Wir untersuchen nun, was wir mit Hilfe des adjungierten Operators über die Lösbarkeit der Gleichung $Tx = y$ aussagen können. Diese besitzt eine eindeutige Lösung, die stetig von der rechten Seite abhängt genau dann, wenn T injektiv und surjektiv ist. Für endlichdimensionale Operatoren (d. h. lineare Gleichungssysteme) können wir dies über den Rang von T ausdrücken; mit dem Dimensionssatz erhält man dann, dass T surjektiv ist genau dann, wenn T^* (der in diesem Fall der transponierten Matrix entspricht) injektiv ist (und umgekehrt). Eine ähnliche Aussage wäre auch für unendlichdimensionale Räume nützlich. Anstelle des Dimensionssatzes verwenden wir dafür die für Satz 7.3 eingeführten Annihilatoren.

Satz 9.6. *Seien X, Y normierte Räume und $T \in L(X, Y)$. Dann gelten:*

(i) $(\operatorname{ran} T)^\perp = \ker T^*$;

(ii) $(\ker T^*)_\perp = \operatorname{cl}(\operatorname{ran} T)$;

(iii) $(\operatorname{ran} T^*)_\perp = \ker T$;

(iv) $(\ker T)^\perp \supset \operatorname{cl}(\operatorname{ran} T^*)$.

Beweis. Für (i) betrachte $y^* \in (\operatorname{ran} T)^\perp$, d. h. für alle $x \in X$ gilt

$$0 = \langle y^*, Tx\rangle_X = \langle T^* y^*, x\rangle_X.$$

Dann gilt aber nach Definition $T^* y^* = 0 \in X^*$ und damit $y^* \in \ker T^*$. Umgekehrt folgt aus $y^* \in \ker T^*$ mit der selben Rechnung $\langle y^*, Tx\rangle_X = 0$, d. h. $y^* \in (\operatorname{ran} T)^\perp$.

Aussage (ii) erhalten wir aus (i) zusammen mit Folgerung 8.8:

$$(\ker T^*)_\perp = ((\operatorname{ran} T)^\perp)_\perp = \operatorname{cl}(\operatorname{ran} T).$$

Genauso zeigt man (iii) und (iv) (unter Berücksichtigung der Bemerkung nach Folgerung 8.8). □

Der Operator T ist also genau dann surjektiv, wenn T^* injektiv und $\operatorname{ran} T$ abgeschlossen ist. Auch letztere Bedingung kann über den adjungierten Operator charakterisiert werden; dies besagt ein weiterer Hauptsatz der Funktionalanalysis, der *Satz vom abgeschlossenen Bild*. Für den Beweis, der wesentlich auf den Hahn–Banach-Sätzen beruht, brauchen wir zwei Lemmata.

Lemma 9.7. *Seien X, Y Banachräume und $T \in L(X,Y)$. Dann sind äquivalent:*

(i) $\operatorname{ran} T$ *ist abgeschlossen;*

(ii) es existiert ein $C > 0$, so dass für alle $y \in \operatorname{ran} T$ ein $x \in X$ existiert mit $Tx = y$ und $\|x\|_X \leq C\|y\|_Y$.

Beweis. (i) ⇒ (ii): Wir verwenden, dass $\operatorname{ran} T$ nach Satz 6.4 isomorph zu $X/\ker T$ ist; der Operator $S : X/\ker T \to \operatorname{ran} T$ aus Lemma 6.3 mit $T = S \circ Q$, wobei Q die Quotientenabbildung $x \mapsto [x]$ bezeichnet, ist also stetig invertierbar. Für alle $x \in X$ existiert daher ein $[x] \in X/\ker T$ mit $S[x] = Tx$ und damit

$$\|[x]\|_{X/\ker T} \leq \|S^{-1}\|_{L(\operatorname{ran} T, X/\ker T)}\|Tx\|_Y.$$

Wie im Beweis von Satz 6.2 finden wir nun nach Definition des Infimums zu $\varepsilon := \|[x]\|_{X/\ker T}$ ein $u_\varepsilon \in \ker T$ mit

$$\|x - u_\varepsilon\|_X \leq \inf_{u \in \ker T} \|x - u\|_X + \varepsilon = 2\|[x]\|_{X/\ker T}.$$

Also erfüllt $\tilde{x} := x - u_\varepsilon \in X$ wegen $T\tilde{x} = Tx = S[x] = y$ die gewünschte Abschätzung mit $C := 2\|S^{-1}\|_{L(\operatorname{ran} T, X/\ker T)}$.

(ii) ⇒ (i): Mit Hilfe der Voraussetzung und der obigen Definition von S finden wir für $y \in \operatorname{ran} T$ ein eindeutiges $[x] \in X/\ker T$ mit $S[x] = Tx = y$ und

$$\|[x]\|_{X/\ker T} \leq \|x\|_X \leq C\|y\|_Y.$$

Also ist S stetig invertierbar, und damit $\operatorname{ran} T$ nach Folgerung 1.9 als Urbild des Banachraums $X/\ker T$ unter der stetigen Abbildung S^{-1} abgeschlossen. □

Beachten Sie, dass die Abschätzung in Aussage (ii) nur für *ein* Urbild x gelten muss, nicht für alle! Aus dem Beweis folgt aber, dass (ii) äquivalent ist zu

$$\|[x]\|_{X/\ker T} \leq C\|Tx\|_Y \qquad \text{für alle } x \in X.$$

Oft kann man eine schärfere Variante nachweisen, aus der zusätzlich die Injektivität von T folgt.

Folgerung 9.8. *Seien X, Y Banachräume und $T \in L(X,Y)$. Existiert ein $C > 0$ mit*

$$\|x\|_X \leq C\|Tx\|_Y \qquad \textit{für alle } x \in X,$$

so ist T injektiv und $\operatorname{ran} T$ *abgeschlossen.*

Beweis. Ist $Tx = 0$, so folgt aus der Ungleichung sofort $x = 0$ und damit die Injektivität von T. In diesem Fall ist x das einzige Urbild von Tx, und deshalb ist nach Voraussetzung Lemma 9.7 (ii) erfüllt und damit $\operatorname{ran} T$ abgeschlossen. □

Für die stetige Invertierbarkeit von T fehlt dann lediglich die Surjektivität, die wir mit Hilfe des folgenden Lemmas nachweisen können.

Lemma 9.9. *Seien X, Y Banachräume und $T \in L(X,Y)$. Existiert ein $c > 0$ mit*

$$c\|y^*\|_{Y^*} \leq \|T^*y^*\|_{X^*} \qquad \text{für alle } y^* \in Y^*,$$

so ist T surjektiv.

Beweis. Wir verwenden den Satz 5.5 von der offenen Abbildung und zeigen $\delta U_Y \subset T(U_X)$ für ein $\delta > 0$ für die offenen Einheitskugeln in X bzw. Y, wobei wie im Beweis dieses Satzes genügt, $cU_Y \subset \operatorname{cl} T(U_X) =: A$ zu zeigen. Wir führen einen indirekten Beweis. Sei dafür $y_0 \notin A$. Da A nichtleer, konvex und abgeschlossen ist, existieren nach dem strikten Trennungssatz (Satz 8.11) ein $y^* \in Y^*$ und ein $\alpha \in \mathbb{R}$ mit

$$\operatorname{Re}\langle y^*, y\rangle_Y \leq \alpha < \operatorname{Re}\langle y^*, y_0\rangle_Y \qquad \text{für alle } y \in A.$$

Da A wegen der Linearität von T mit y auch σy für alle $\sigma \in \mathbb{K}$ mit $|\sigma| = 1$ enthält, folgt

$$|\langle y^*, y\rangle_Y| = \operatorname{Re}\langle y^*, \sigma y\rangle_Y < \operatorname{Re}\langle y^*, y_0\rangle_Y \leq |\langle y^*, y_0\rangle_Y| \qquad \text{für alle } y \in A,$$

wobei $\sigma \in \mathbb{K}$ so gewählt ist, dass $\sigma\langle y^*, y\rangle_Y = |\langle y^*, y\rangle_Y|$ und $|\sigma| = 1$ ist.

Nach Voraussetzung und Lemma 4.3 (i) gilt dann

$$\begin{aligned} c\|y^*\|_{Y^*} \leq \|T^*y^*\|_{X^*} &= \sup_{x\in U_X} |\langle T^*y^*, x\rangle_X| = \sup_{x\in U_X} |\langle y^*, Tx\rangle_Y| \\ &\leq |\langle y^*, y_0\rangle_Y| \leq \|y^*\|_{y^*}\|y_0\|_Y \end{aligned}$$

und damit $\|y_0\|_Y \geq c$, d. h. $y_0 \notin cU_Y$. Durch Kontraposition folgt $cU_Y \subset A = \operatorname{cl} T(U_X)$. □

Wir kommen nun zum Hauptresultat dieses Kapitels.

Satz 9.10 (vom abgeschlossenen Bild). *Seien X, Y Banachräume und $T \in L(X,Y)$. Dann sind äquivalent:*

(i) $\operatorname{ran} T$ *ist abgeschlossen;*

(ii) $\operatorname{ran} T = (\ker T^*)_\perp$*;*

(iii) $\operatorname{ran} T^*$ *ist abgeschlossen;*

(iv) $\operatorname{ran} T^* = (\ker T)^\perp$.

Beweis. *(i) ⇔ (ii):* Ist $\operatorname{ran} T$ abgeschlossen, so gilt nach Satz 9.6 (ii)

$$(\ker T^*)_\perp = \operatorname{cl}(\operatorname{ran} T) = \operatorname{ran} T.$$

Die andere Richtung folgt sofort aus der Abgeschlossenheit von Annihilatoren.

(i) ⇒ (iv): Zunächst gilt stets $\langle T^* y^*, x\rangle_X = \langle y^*, Tx\rangle_Y = 0$ für alle $x \in \ker T$ und damit $\operatorname{ran} T^* \subset (\ker T)^\perp$. Für die umgekehrte Inklusion sei $x^* \in (\ker T)^\perp$. Wir konstruieren nun ein $y^* \in Y^*$ mit $x^* = T^* y^*$. Zuerst definieren wir ein lineares Funktional

$$y_0^* : \operatorname{ran} T \to \mathbb{K}, \qquad \langle y_0^*, Tx\rangle_{\operatorname{ran} T} := \langle x^*, x\rangle_X \quad \text{für alle } x \in X.$$

Dieses Funktional ist wohldefiniert, denn für $x_1, x_2 \in X$ mit $Tx_1 = Tx_2$ ist $x_1 - x_2 \in \ker T$ und daher $\langle x^*, x_1 - x_2\rangle_X = 0$. Zum Nachweis der Stetigkeit verwenden wir Lemma 9.7: Da $\operatorname{ran} T$ abgeschlossen ist, existiert ein $C > 0$ mit $\|x\|_X \leq C\|y\|_Y$ für alle $y \in \operatorname{ran} T$ und ein $x \in X$ mit $Tx = y$. Daher gilt für $y \in \operatorname{ran} T$ beliebig und $x \in X$ entsprechend gewählt

$$|\langle y_0^*, y\rangle_{\operatorname{ran} T}| = |\langle y_0^*, Tx\rangle_{\operatorname{ran} T}| = |\langle x^*, x\rangle_X| \leq \|x^*\|_{X^*}\|x\|_X \leq C\|x^*\|_{X^*}\|y\|_Y.$$

Also ist y_0^* stetig. Da $\operatorname{ran} T \subset Y$ ein Unterraum ist, können wir nun y_0^* mit Hilfe des Fortsetzungssatzes 8.2 stetig zu einem Funktional $y^* \in Y^*$ fortsetzen. Dann gilt

$$\langle x^*, x\rangle_X = \langle y_0^*, Tx\rangle_{\operatorname{ran} T} = \langle y^*, Tx\rangle_Y = \langle T^* y^*, x\rangle_X \qquad \text{für alle } x \in X$$

und damit $x^* = T^* y^*$, d. h. $x^* \in \operatorname{ran} T^*$.

(iv) ⇒ (iii) folgt wieder direkt aus der Abgeschlossenheit von Annihilatoren.

(iii) ⇒ (i): Wir zeigen $\operatorname{ran} T = \operatorname{cl}(\operatorname{ran} T) =: U$. Dafür konstruieren wir einen Operator $S \in L(X, U)$ durch $Sx := Tx$ für alle $x \in X$ und zeigen mit Hilfe von $S^* : U^* \to X^*$ und Lemma 9.9, dass S surjektiv (d. h. $\operatorname{ran} T = \operatorname{ran} S = U$) ist. Für beliebige $y^* \in Y^*$ bezeichnen wir die Einschränkung auf $U \subset Y$ mit $y^*|_U \in U^*$; dann gilt

$$\langle T^* y^*, x\rangle_X = \langle y^*, Tx\rangle_Y = \langle y^*|_U, Sx\rangle_U = \langle S^* y^*|_U, x\rangle_X \qquad \text{für alle } x \in X,$$

d. h. $T^* y^* = S^* y^*|_U$ und damit ist $\operatorname{ran} T^* \subset \operatorname{ran} S^*$. Sei umgekehrt $u^* \in U^*$ und damit $S^* u^* \in \operatorname{ran} S^*$ beliebig. Dann kann u^* mit dem Fortsetzungssatz 8.2 stetig zu einem $y^* \in Y^*$ fortgesetzt werden; analog zu oben gilt dann

$$\langle S^* u^*, x\rangle_X = \langle u^*, Sx\rangle_U = \langle y^*, Tx\rangle_Y = \langle T^* y^*, x\rangle_X \qquad \text{für alle } x \in X,$$

d. h. $S^* u^* = T^* y^*$ und damit $\operatorname{ran} S^* \subset \operatorname{ran} T^*$. Weiter erhalten wir aus Satz 9.6 (ii) und der Bemerkung nach Folgerung 8.8

$$\{0\} = U^\perp = (\operatorname{cl}(\operatorname{ran} S))^\perp = ((\ker S^*)_\perp)^\perp \supset \operatorname{cl} \ker S^* = \ker S^*.$$

Also ist $S^* : U^* \to \operatorname{ran} S^*$ injektiv und, da $\operatorname{ran} S^* = \operatorname{ran} T^*$ nach Voraussetzung abgeschlossen ist, nach Folgerung 5.7 stetig invertierbar. Für alle $u^* \in U^*$ gilt also

$$\|u^*\|_{U^*} = \|S^{-*} S^* u^*\|_{U^*} \leq \|S^{-*}\|_{L(X^*, U^*)} \|S^* u^*\|_{X^*}.$$

Damit ist die Voraussetzung von Lemma 9.9 mit $c := \|S^{-*}\|_{L(X^*, U^*)}^{-1}$ erfüllt, also ist S surjektiv. Wir erhalten dadurch $\operatorname{ran} S = U = \operatorname{cl}(\operatorname{ran} T)$, d. h. $\operatorname{ran} T = \operatorname{ran} S$ ist abgeschlossen. □

In der Praxis wird der Satz vom abgeschlossenen Bild oft angewendet, indem man die Voraussetzung von Folgerung 9.8 sowie Injektivität von T^* nachweist; aus ersterem erhält man die Injektivität, aus letzterem zusammen mit der Abgeschlossenheit des Bildes die Surjektivität von T. Damit hat die Gleichung $Tx = y$ für alle $y \in Y$ eine eindeutige Lösung, und es gilt die *a priori Abschätzung* $\|x\|_X \leq C\|y\|_Y$. Dies liefert ein fundamentales Werkzeug in der Theorie partieller Differentialgleichungen.

10 REFLEXIVITÄT

Wie wir in den letzten Kapiteln gesehen haben, ist der Dualraum X^* eines normierten Vektorraums X von Interesse, da er X in gewisser Weise charakterisiert. In der selben Weise wird natürlich X^* durch seinen Dualraum $(X^*)^*$ charakterisiert, und man kann sich nun fragen, inwiefern dies transitiv ist, d. h. was der *Bidualraum* $X^{**} := (X^*)^*$ von X über X aussagt.

Wie schauen die Elemente von X^{**} aus? Zunächst betrachten wir für $x \in X$ und $x^* \in X^*$ die duale Paarung $\langle x^*, x\rangle_X \in \mathbb{K}$. Hält man x^* fest, so erhält man eine stetige lineare Abbildung von X nach $\mathbb{K}$ (nämlich x^*; so war die duale Paarung ja definiert). Man kann aber auch x festhalten und erhält so eine lineare Abbildung x^{**} von X^* nach $\mathbb{K}$; wegen

$$x^{**}(x^*) := \langle x^*, x\rangle_X \leq \|x^*\|_{X^*}\|x\|_X \qquad \text{für alle } x^* \in X^* \tag{10.1}$$

ist diese auch stetig. Dies definiert die *kanonische Einbettung*

$$J_X : X \to X^{**}, \qquad \langle J_X(x), x^*\rangle_{X^*} = \langle x^*, x\rangle_X \quad \text{für alle } x^* \in X^*.$$

Man vergewissert sich leicht, dass J_X linear und wegen (10.1) auch stetig ist. Es gilt sogar noch mehr: Aus Folgerung 8.4 erhalten wir

$$\|x\|_X = \max_{x^* \in B_{X^*}} |\langle x^*, x\rangle_X| = \max_{x^* \in B_{X^*}} |\langle J_X x, x^*\rangle_{X^*}| = \|J_X x\|_{X^{**}},$$

und damit ist J_X eine Isometrie (und daher injektiv). Wir halten dies fest als

Satz 10.1. *Die kanonische Einbettung $J_X : X \to X^{**}$ ist linear, injektiv und isometrisch.*

Das Bild $\operatorname{ran} J_X$ ist ein Unterraum von X^{**}. Da J_X isometrisch und X^{**} als Dualraum stets ein Banachraum ist, ist $\operatorname{ran} J_X$ vollständig genau dann, wenn X vollständig ist; ersteres ist nach Lemma 3.5 genau dann der Fall, wenn $\operatorname{ran} J_X$ abgeschlossen in X^{**} ist. Ist X nicht vollständig, können wir $\operatorname{cl}\operatorname{ran} J_X \subset X^{**}$ als „Vervollständigung" von X ansehen.

Im Allgemeinen ist $\operatorname{ran} J_X$ ein echter Unterraum, da J_X nicht surjektiv sein muss. Ist J_X surjektiv, so heißt X *reflexiv*. In diesem Fall ist J_X sogar ein isometrischer Isomorphismus; es gilt also $X \cong X^{**}$. Beachten Sie, dass $X \cong X^{**}$ alleine noch *nicht* bedeutet, dass X reflexiv ist: Für die Reflexivität muss die Isometrie zwingend durch die kanonische Einbettung vermittelt werden.

Beispiel 10.2.

(i) Alle endlichdimensionalen Räume sind reflexiv, denn es gilt stets

$$\dim(X^{**}) = \dim(X^*) = \dim(X),$$

und daher folgt aus der Injektivität von J_X bereits die Surjektivität.

(ii) Die Folgenräume $\ell^p(\mathbb{K})$ sind reflexiv für $1 < p < \infty$. Dafür verwenden wir die Isometrien $T_p : \ell^q(\mathbb{K}) \to \ell^p(\mathbb{K})^*$ und $T_q : \ell^p(\mathbb{K}) \to \ell^q(\mathbb{K})^*$ aus Satz 7.1. Aus der Definition folgt

$$\langle T_q x, y\rangle_{\ell^q} = \sum_{k=1}^{\infty} x_k y_k = \langle T_p y, x\rangle_{\ell^p} \qquad \text{für alle } x \in \ell^p(\mathbb{K}),\ \ y \in \ell^q(\mathbb{K}).$$

Für $x^{**} \in \ell^p(\mathbb{K})^{**}$ setzen wir nun $x := T_q^{-1}T_p^* x^{**} \in \ell^p(\mathbb{K})$ und zeigen $J_{\ell^p} x = x^{**}$. Für $x^* \in \ell^p(\mathbb{K})^*$ beliebig setze $y := T_p^{-1}x^* \in \ell^q(\mathbb{K})$. Dann gilt

$$\begin{aligned}\langle J_{\ell^p} x, x^*\rangle_{(\ell^p)^*} &= \langle x^*, x\rangle_{\ell^p} = \langle T_p y, x\rangle_{\ell^p} = \langle T_q x, y\rangle_{\ell^q} \\ &= \langle T_p^* x^{**}, y\rangle_{\ell^q} = \langle x^{**}, T_p y\rangle_{(\ell^p)^*} = \langle x^{**}, x\rangle_{(\ell^p)^*}.\end{aligned}$$

Also ist $J_{\ell^p} = T_p^{-*}T_q$ surjektiv.

(iii) Ebenso zeigt man, dass $L^p(\Omega)$ reflexiv ist für $1 < p < \infty$.

(iv) Nicht reflexiv sind $\ell^1(\mathbb{K})$, $\ell^\infty(\mathbb{K})$, $c_0(\mathbb{K})$, $c(\mathbb{K})$ sowie $L^1(\Omega)$, $L^\infty(\Omega)$, $C_b(X)$; dies folgt aus den nächsten drei Sätzen.

Analog zum Bidualraum können wir auch für $T \in L(X, Y)$ einen *biadjungierten Operator* $T^{**} := (T^*)^* \in L(X^{**}, Y^{**})$ definieren. Dieser ist verträglich mit der kanonischen Einbettung.

Lemma 10.3. *Seien X, Y normierte Räume und $T \in L(X, Y)$. Dann gilt*

$$T^{**} \circ J_X = J_Y \circ T.$$

Beweis. Für beliebige $x \in X$ und $y^* \in Y^*$ gilt

$$\langle T^{**} J_X x, y^*\rangle_{Y^*} = \langle J_X x, T^* y^*\rangle_{X^*} = \langle T^* y^*, x\rangle_X = \langle y^*, Tx\rangle_Y = \langle J_Y T x, y^*\rangle_{Y^*}. \qquad \square$$

Wir zeigen nun einige Resultate über die „Vererbung“ der Reflexivität.

Satz 10.4. *Seien X ein normierter Vektorraum und Y ein reflexiver Banachraum mit $X \simeq Y$. Dann ist auch X reflexiv.*

Beweis. Ist $T : X \to Y$ ein Isomorphismus, d. h. stetig invertierbar, so ist nach Satz 9.4 auch T^* und damit dann auch T^{**} stetig invertierbar. Aus Lemma 10.3 folgt nun, dass J_X stetig invertierbar und damit surjektiv ist genau dann, wenn es J_Y ist. □

Satz 10.5. *Sei X ein reflexiver Banachraum und $U \subset X$ ein abgeschlossener Unterraum. Dann ist auch U reflexiv.*

Beweis. Sei $u^{**} \in U^{**}$ beliebig. Da wir jedes stetige lineare Funktional $x^* \in X^*$ durch Einschränkung auf U auch als Funktional $x^*|_U \in U^*$ auffassen dürfen, können wir ein $x^{**} \in X^{**}$ definieren durch

$$\langle x^{**}, x^* \rangle_{X^*} = \langle u^{**}, x^*|_U \rangle_{U^*} \qquad \text{für alle } x^* \in X^*. \tag{10.2}$$

Nun ist X reflexiv und damit existiert ein $x := J_X^{-1} x^{**} \in X$. Wir zeigen nun $x \in U$. Angenommen, $x \notin U$. Dann gibt es nach Folgerung 8.5 ein $x^* \in U^\perp$ mit $\langle x^*, x \rangle_X \neq 0$. Insbesondere ist dann $x^*|_U = 0$, und zusammen mit (10.2) erhalten wir daraus den Widerspruch

$$0 = \langle u^{**}, x^*|_U \rangle_{U^*} = \langle x^{**}, x^* \rangle_{X^*} = \langle x^*, x \rangle_X \neq 0.$$

Also gilt $x \in U$, und es bleibt $J_U x = u^{**}$ zu zeigen. Sei dafür $u^* \in U^*$ gegeben. Nach dem Fortsetzungssatz 8.2 von Hahn–Banach existiert dann ein $x^* \in X^*$ mit $x^*|_U = u^*$. Daraus folgt, wieder mit (10.2),

$$\begin{aligned} \langle J_U x, u^* \rangle_{U^*} &= \langle u^*, x \rangle_U = \langle x^*, x \rangle_X = \langle J_X x, x^* \rangle_{X^*} \\ &= \langle x^{**}, x^* \rangle_{X^*} = \langle u^{**}, x^*|_U \rangle_{U^*} = \langle u^{**}, u^* \rangle_{U^*}, \end{aligned}$$

d. h. $J_U x = u^{**}$. □

Satz 10.6. *Sei X ein Banachraum. Dann ist X reflexiv genau dann, wenn X^* reflexiv ist.*

Beweis. Sei X reflexiv. Wir müssen zeigen, dass $J_{X^*} : X^* \to (X^{**})^*$ surjektiv ist. Sei dafür $x^{***} \in (X^{**})^*$ beliebig, und definiere $x^* \in X^*$ durch $\langle x^*, x \rangle_X = \langle x^{***}, J_X x \rangle_{X^{**}}$ für alle $x \in X$. Da X reflexiv ist, existiert für jedes $x^{**} \in X^{**}$ ein $x := J_X^{-1} x^{**} \in X$. Dann gilt

$$\langle x^{***}, x^{**} \rangle_{X^{**}} = \langle x^{***}, J_X x \rangle_{X^{**}} = \langle x^*, x \rangle_X = \langle J_X x, x^* \rangle_{X^*} = \langle x^{**}, x^* \rangle_{X^*} = \langle J_{X^*} x^*, x^{**} \rangle_{X^{**}}$$

und daher $x^{***} = J_{X^*} x^*$.

Sei nun umgekehrt X^* reflexiv. Dann ist nach dem eben Gezeigten X^{**} reflexiv. Weiter ist $\operatorname{ran} J_X \subset X^{**}$ ein abgeschlossener Unterraum (denn J_X ist eine Isometrie – bildet also Cauchy-Folgen auf Cauchy-Folgen ab – und X ist vollständig) und damit nach Satz 10.5 ebenfalls reflexiv. Wegen $X \simeq \operatorname{ran} J_X$ (denn $J_X : X \to \operatorname{ran} X$ ist nach Konstruktion surjektiv und damit nach Satz 10.1 zusammen mit Satz 5.6 ein Isomorphismus), ist X nach Satz 10.4 auch reflexiv. □

Aus diesen Resultaten folgt wie schon behauptet, dass $\ell^1(\mathbb{K})$ nicht reflexiv sein kann: Ansonsten wäre $\ell^\infty(\mathbb{K})^* \cong (\ell^1(\mathbb{K})^*)^* \cong \ell^1(\mathbb{K})$ nach Satz 3.15 separabel und damit nach Folgerung 8.7 auch $\ell^\infty(\mathbb{K})$, was aber (ebenfalls nach Satz 3.15) nicht der Fall ist. Dann können aber nach Satz 10.6 auch $\ell^\infty(\mathbb{K})$ und $c_0(\mathbb{K})$ nicht reflexiv sein, denn $\ell^\infty(\mathbb{K}) \cong \ell^1(\mathbb{K})^*$ und $c_0(\mathbb{K})^* \cong \ell^1(\mathbb{K})$. Schließlich ist $c_0(\mathbb{K})$ ein nicht-reflexiver abgeschlossener Unterraum von $c(\mathbb{K})$, welcher damit nach Satz 10.5 auch nicht reflexiv sein kann. Analog argumentiert man für die Funktionenräume.

11 SCHWACHE KONVERGENZ

Wir kommen nun zu der versprochenen Verallgemeinerung der komponentenweisen Konvergenz auf unendlichdimensionale Vektorräume, mit deren Hilfe wir ein analoges Resultat zum Satz 2.5 von Heine–Borel beweisen können. Wie wir in Kapitel 7 argumentiert haben, sind stetige Funktionale das unendlichdimensionale Pendant zur Komponentenauswertung; entsprechend definieren wir unseren Konvergenzbegriff.

Sei X ein normierter Raum und $\{x_n\}_{n\in\mathbb{N}} \subset X$ eine Folge. Wir sagen, $\{x_n\}_{n\in\mathbb{N}}$ *konvergiert schwach* in X gegen ein $x \in X$ falls gilt

$$\lim_{n\to\infty} \langle x^*, x_n\rangle_X = \langle x^*, x\rangle_X \qquad \text{für alle } x^* \in X^*$$

und schreiben $x_n \rightharpoonup x$. Analog sagen wir, die Folge $\{x_n^*\}_{n\in\mathbb{N}} \subset X^*$ *konvergiert schwach-∗* in X^*, falls

$$\lim_{n\to\infty} \langle x_n^*, x\rangle_X = \langle x^*, x\rangle_X \qquad \text{für alle } x \in X$$

und schreiben $x_n^* \rightharpoonup^* x^*$. Diese Grenzwerte sind stets eindeutig. Für die schwach-∗ Konvergenz folgt dies direkt aus der Definition: Gilt $x_n^* \rightharpoonup^* x^*$ und $x_n^* \rightharpoonup^* y^*$, so ist

$$\langle x^*, x\rangle_X = \lim_{n\to\infty} \langle x_n^*, x\rangle_X = \langle y^*, x\rangle_X \qquad \text{für alle } x \in X$$

und damit nach Definition $x^* = y^*$. Ist $x_n \rightharpoonup x$ und $x_n \rightharpoonup y$, so gilt analog

$$\langle x^*, x\rangle_X = \langle x^*, y\rangle_X \qquad \text{für alle } x^* \in X^*.$$

Wäre nun aber $x \neq y$, so existiert nach Folgerung 8.5 ein $x^* \in X^*$ mit $\langle x^*, x-y\rangle_X \neq 0$, im Widerspruch zur obigen Gleichung. Ist X reflexiv, so stimmen schwache Konvergenz (in X) und schwach-∗ Konvergenz (in X^{**}) überein; in endlichdimensionalen Räumen fallen alle Konvergenzbegriffe zusammen.

Zur besseren Unterscheidung nennen wir Folgen *stark konvergent*, wenn sie bezüglich der durch die Norm induzierten Metrik konvergieren. Jede Folge, die stark in X konvergiert, konvergiert auch schwach; dies folgt sofort aus

$$\langle x^*, x_n - x\rangle_X \leq \|x^*\|_{X^*}\|x_n - x\|_X \to 0.$$

Genauso folgt aus der starken Konvergenz in X^* die schwach-∗ Konvergenz. Die Umkehrung gilt dagegen nicht.

Beispiel 11.1. Betrachten wir die Folge der Einheitsvektoren $\{e_n\}_{n\in\mathbb{N}} \subset \ell^p(\mathbb{K}), 1 \le p \le \infty$, so gilt $\|e_n\|_p = 1$ für alle $n \in \mathbb{N}$; weiter ist $\{e_n\}_{n\in\mathbb{N}}$ für kein p eine Cauchy-Folge.

Für $1 < p < \infty$ können wir nun die Darstellung des Dualraums $\ell^p(\mathbb{K})^* \cong \ell^q(\mathbb{K})$ verwenden: Für jedes $x^* \in \ell^p(\mathbb{K})^*$ existiert ein $y \in \ell^q(\mathbb{K})$ mit

$$\langle x^*, e_n\rangle_{\ell^p} = \sum_{k=1}^{\infty} y_k [e_n]_k = y_n \to 0,$$

denn wenn $\|y\|_q$ endlich ist, muss $\{y_k\}_{k\in\mathbb{N}}$ eine Nullfolge sein. Also gilt $e_n \rightharpoonup 0$ in $\ell^p(\mathbb{K})$ für $1 < p < \infty$ und damit auch $e_n \rightharpoonup^* 0$.

Da $\ell^1(\mathbb{K})$ nicht reflexiv ist, haben wir hier die Wahl: Fassen wir $e_n \in \ell^1(\mathbb{K})$ mit Hilfe des Isomorphismus T aus Satz 7.1 als Element des Dualraums von $c_0(\mathbb{K})$ auf, so haben wir

$$\langle Te_n, y\rangle_{c_0} = \sum_{k=1}^{\infty} [e_n]_k y_k = y_n \to 0 \qquad \text{für alle } y \in c_0(\mathbb{K})$$

und damit $e_n \rightharpoonup^* 0$ in $\ell^1(\mathbb{K})$. Dagegen gilt *nicht* $e_n \rightharpoonup 0$, denn für die konstante Folge $y := \{1\}_{k\in\mathbb{N}} \in \ell^\infty(\mathbb{K}) \cong \ell^1(\mathbb{K})^*$ gilt

$$\langle Ty, e_n\rangle_{\ell^1} = \sum_{k=1}^{\infty} [e_n]_k y_k = y_n = 1 \qquad \text{für alle } n \in \mathbb{N}.$$

Da wegen $c_0(\mathbb{K}) \subset \ell^\infty(\mathbb{K})$ als Grenzwert nur 0 in Frage kommt, kann $\{e_n\}_{n\in\mathbb{N}}$ nicht schwach konvergieren.[1]

Schwach konvergente Folgen konvergieren also nicht unbedingt bezüglich der Norm; es gilt noch nicht einmal $\|x_n\|_X \to \|x\|_X$. Wir haben aber die folgenden schwächeren Aussagen.

Satz 11.2. *Sei X ein normierter Raum und seien $\{x_n\}_{n\in\mathbb{N}} \subset X$ und $\{x_n^*\}_{n\in\mathbb{N}} \subset X^*$.*

(i) Aus $x_n \rightharpoonup x$ folgt $\|x\|_X \le \liminf_{n\to\infty} \|x_n\|_X$.

(ii) Aus $x_n^* \rightharpoonup^* x^*$ folgt $\|x^*\|_{X^*} \le \liminf_{n\to\infty} \|x_n^*\|_{X^*}$.

Beweis. Wir zeigen zuerst (ii). Sei $\{x_n^*\}_{n\in\mathbb{N}} \subset X^*$ eine schwach-$*$ konvergente Folge mit $x_n^* \rightharpoonup^* x^* \in X^*$. Dann gilt auch (Satz 3.6 (iii) für $(\mathbb{R}, |\cdot|)$) $|\langle x_n^*, x\rangle_X| \to |\langle x^*, x\rangle_X|$ für alle $x \in X$ und daher

$$|\langle x^*, x\rangle_X| = \lim_{n\to\infty} |\langle x_n^*, x\rangle_X| \le \liminf_{n\to\infty} \|x_n^*\|_{X^*}\|x\|_X.$$

[1] In der Tat ist die schwache Konvergenz in $\ell^1(\mathbb{K})$ äquivalent zur starken Konvergenz; dies ist eine besondere Eigenschaft dieses Raums, bekannt als *Lemma von Schur*, siehe z. B. [Kaballo 2018, Satz 10.15].

Nehmen wir das Supremum über alle $x \in B_X$, folgt daraus nach Definition der Operatornorm die gewünschte Aussage.

Aussage (i) zeigt man analog mit Hilfe von Folgerung 8.4. □

Die Eigenschaft in (i) bezeichnet man als *schwache Unterhalbstetigkeit*, die in (ii) als *schwach-∗ Unterhalbstetigkeit.*

Weiterhin sind schwach konvergente Folgen beschränkt.

Satz 11.3. *Sei X ein normierter Raum. Dann sind schwach konvergente Folgen in X beschränkt. Ist X vollständig, so sind auch schwach-∗ konvergente Folgen in X^* beschränkt.*

Beweis. Wir zeigen wieder zuerst die Aussage für X^*. Aus $x_n^* \rightharpoonup^* x^*$ folgt wieder $|\langle x_n^*, x\rangle_X| \to |\langle x^*, x\rangle_X|$ für alle $x \in X$ und daher

$$\sup_{n\in\mathbb{N}} |\langle x_n^*, x\rangle_X| < \infty \qquad \text{für alle } x \in X.$$

Da X nach Annahme ein Banachraum ist und die Abbildung $x^* \in X^* = L(X, \mathbb{K})$ linear und stetig ist, folgt aus dem Satz 5.3 von Banach–Steinhaus

$$\sup_{n\in\mathbb{N}} \|x_n^*\|_{X^*} < \infty,$$

d. h. $\{x_n^*\}_{n\in\mathbb{N}}$ ist beschränkt.

Die Aussage für die schwache Konvergenz führen wir mit Hilfe der kanonischen Einbettung $J_X : X \to X^{**}$ zurück auf die schwach-∗ Konvergenz (in X^{**}): Konvergiert $\{x_n\}_{n\in\mathbb{N}} \subset X$ schwach mit $x_n \rightharpoonup x$, so folgt daraus

$$\langle J_X x_n - J_X x, x^*\rangle_{X^*} = \langle J_X(x_n - x), x^*\rangle_{X^*} = \langle x^*, x_n - x\rangle_X \to 0 \qquad \text{für alle } x^* \in X^*,$$

d. h. $\{J_X x_n\}_{n\in\mathbb{N}}$ konvergiert schwach-∗ in X^{**}. Da J_X nach Satz 10.1 eine Isometrie ist, folgt aus der zuerst gezeigten Aussage (denn der Dualraum X^* ist stets vollständig)

$$\sup_{n\in\mathbb{N}} \|x_n\|_X = \sup_{n\in\mathbb{N}} \|J_X x_n\|_{X^{**}} < \infty,$$

d. h. $\{x_n\}_{n\in\mathbb{N}}$ ist beschränkt. □

Aus Satz 11.3 folgt zum Beispiel, dass die duale Paarung von schwach konvergenten Folgen noch konvergiert, solange wenigstens eine davon stark konvergiert.

Folgerung 11.4. *Sei X ein normierter Raum und seien $\{x_n\}_{n\in\mathbb{N}} \subset X$ und $\{x_n^*\}_{n\in\mathbb{N}} \subset X^*$. Gilt entweder*

(i) $x_n \rightharpoonup x$ und $x_n^* \to x^*$ oder

(ii) $x_n \to x$ *und* $x_n^* \rightharpoonup^* x^*$ *und* X *ist vollständig,*

so konvergiert $\langle x_n^*, x_n\rangle_X \to \langle x^*, x\rangle$.

Beweis. Gilt (i), so schätzen wir mit der produktiven Null $-\langle x^*, x_n\rangle_X + \langle x^*, x_n\rangle_X$ ab

$$\begin{aligned}|\langle x^*, x\rangle_X - \langle x_n^*, x_n\rangle_X| &\le |\langle x^*, x - x_n\rangle_X| + |\langle x_n^* - x^*, x_n\rangle_X|\\ &\le |\langle x^*, x - x_n\rangle_X| + \|x_n\|_X\|x_n^* - x^*\|_{X^*}.\end{aligned}$$

Für $n \to \infty$ verschwindet der erste Summand wegen $x_n \rightharpoonup x$; nach Satz 11.3 ist $\{\|x_n\|_X\}_{n\in\mathbb{N}}$ beschränkt, und wegen $x_n^* \to x^*$ verschwindet daher auch der zweite Term.

Analog beweist man die Aussage unter Voraussetzung (ii). □

Dagegen folgt im Allgemeinen aus $x_n \rightharpoonup x$ und $x_n^* \rightharpoonup x^*$ *nicht* $\langle x_n^*, x_n\rangle_X \to \langle x^*, x\rangle_X$. Als Beispiel betrachte wieder die Einheitsvektoren in $\ell^2(\mathbb{K}) \cong \ell^2(\mathbb{K})^*$, für die nach Beispiel 11.1 $e_n \rightharpoonup 0$ und $e_n \rightharpoonup^* 0$, aber $\langle Te_n, e_n\rangle_{\ell^2} = \sum_{k=1}^\infty [e_n]_k^2 = 1$ gelten.

Das Beispiel zeigt auch, dass die abgeschlossene Menge $S_{\ell^2} := \left\{x \in \ell^2 : \|x\|_2 = 1\right\}$ nicht *schwach (folgen-)abgeschlossen* ist, d. h. für $\{x_n\}_{n\in\mathbb{N}} \subset S_{\ell^2}$ mit $x_n \rightharpoonup x$ nicht unbedingt $x \in S_{\ell^2}$ gelten muss (vergleiche Satz 11.2). Dies liegt daran, dass diese Menge nicht konvex ist.

Satz 11.5. *Sei X ein normierter Raum und $U \subset X$ konvex. Dann ist U genau dann abgeschlossen, wenn U schwach folgenabgeschlossen ist.*

Beweis. Da eine konvergente Folge stets auch schwach konvergiert, ist jede schwach abgeschlossene Menge auch abgeschlossen.[2] Sei daher $U \subset X$ konvex und abgeschlossen, und betrachte eine Folge $\{x_n\}_{n\in\mathbb{N}} \subset U$ mit $x_n \rightharpoonup x \in X$. Angenommen, $x \in X \setminus U$. Dann finden wir nach dem strikten Trennungssatz (Satz 8.11) ein $x^* \in X^*$ und ein $\alpha \in \mathbb{R}$ mit insbesondere

$$\operatorname{Re}\langle x^*, x_n\rangle_X \le \alpha < \operatorname{Re}\langle x^*, x\rangle_X \quad \text{für alle } n \in \mathbb{N}.$$

Wegen $x_n \rightharpoonup x$ können wir in der ersten Ungleichung zum Grenzwert $n \to \infty$ übergehen und erhalten dadurch den Widerspruch

$$\operatorname{Re}\langle x^*, x\rangle_X \le \alpha < \operatorname{Re}\langle x^*, x\rangle_X.$$ □

Dagegen sind abgeschlossene konvexe Teilmengen von X^* in der Regel nicht schwach-∗ abgeschlossen. (Eine Ausnahme bildet wegen Satz 11.2 (ii) die Einheitskugel B_{X^*}.)

Wir kommen nun zu den versprochenen schwachen Kompaktheitsresultaten. Wir nennen eine Teilmenge $U \subset X$ *schwach folgenkompakt*, wenn jede Folge in U eine schwach konvergente Teilfolge mit Grenzwert in U enthält. Analog definieren wir *schwach-∗ folgenkompakte* Teilmengen von X^*.

[2] Schwache Abgeschlossenheit ist also eine *stärkere* Eigenschaft als Abgeschlossenheit.

Satz 11.6 (Banach–Alaoglu[3]). *Sei X ein normierter Raum. Ist X separabel, so ist die Einheitskugel B_{X^*} in X^* schwach-$*$ folgenkompakt.*

Beweis. Die Aussage erinnert an den Satz 2.12 von Arzelà–Ascoli: stetige lineare Funktionale sind Elemente in $C(X)$, die Beschränktheit in der Operatornorm entspricht der gleichgradigen Stetigkeit, und anstelle der Kompaktheit können wir gleich die Separabilität von X nutzen. Der Beweis verläuft daher analog.

Sei $\{x_n^*\}_{n\in\mathbb{N}} \subset B_{X^*}$ eine Folge und $\{x_m : m \in \mathbb{N}\} \subset X$ eine dichte Teilmenge. Für alle $m \in \mathbb{N}$ gilt $|\langle x_n^*, x_m\rangle_X| \leq \|x_m\|_X$, daher sind die Folgen $\{\langle x_n^*, x_m\rangle_X\}_{n\in\mathbb{N}} \subset \mathbb{K}$ beschränkt und enthalten wegen der Vollständigkeit von $\mathbb{K}$ jeweils eine konvergente Teilfolge, deren Grenzwert wir mit $\langle x^*, x_m\rangle_X$ bezeichnen (dies impliziert noch nicht die Existenz eines Grenzwertes $x^* \in X^*$!) Genau wie im Beweis von Satz 2.12 konstruiert man daraus eine Diagonalfolge $\{z_n^*\}_{n\in\mathbb{N}} \subset \{x_n^*\}_{n\in\mathbb{N}} \subset B_{X^*}$ mit $\langle z_n^*, x_m\rangle_X \to \langle x^*, x_m\rangle_X$ für alle $m \in \mathbb{N}$. Wir suchen nun den schwach-$*$ Grenzwert $x^* \in X^*$ dieser Diagonalfolge.

Dafür betrachten wir zuerst den Unterraum $Z = \operatorname{span}\{x_m : m \in \mathbb{N}\}$ und definieren ein lineares Funktional

$$z^* : Z \to \mathbb{K}, \qquad \langle z^*, z\rangle_X := \lim_{n\to\infty} \langle z_n^*, z\rangle_X \quad \text{für alle } z \in Z.$$

Dieses Funktional ist wohldefiniert, da alle $z \in Z$ Linearkombinationen der x_m sind und wir daher den rechten Grenzwert durch die entsprechende Linearkombination der $\langle x^*, x_m\rangle_X$ ausdrücken können. Außerdem ist $|\langle z_n^*, z\rangle_X| \leq \|z\|_X$ und daher auch $|\langle z^*, z\rangle_X| \leq \|z\|_X$ für alle $z \in Z$, also ist z^* stetig mit $\|z^*\|_{Z^*} \leq 1$. Dieses Funktional setzen wir nun mit Hilfe des Fortsetzungssatzes 8.2 von Hahn–Banach zu einem Funktional $x^* \in B_{X^*}$ fort.

Bleibt zu zeigen, dass $z_n^* \rightharpoonup^* x^*$ gilt. Seien dafür $x \in X$ und $\varepsilon > 0$ beliebig. Da Z dicht in X liegt, finden wir ein $z \in Z$ mit $\|z - x\|_X \leq \varepsilon$. Aus der Konvergenz $\langle z_n^*, z\rangle_X \to \langle z^*, z\rangle_X$ erhalten wir weiterhin ein $N \in \mathbb{N}$ mit $|\langle z_n^* - z^*, z\rangle_X| \leq \varepsilon$ für alle $n \geq N$. Daraus folgt nun

$$\begin{aligned}
|\langle x^*, x\rangle_X - \langle z_n^*, x\rangle_X| &\leq |\langle x^* - z_n^*, x - z\rangle_X| + |\langle x^* - z_n^*, z\rangle_X| \\
&\leq (\|x^*\|_{X^*} + \|z_n^*\|_{X^*})\|x - z\|_X + |\langle z^* - z_n^*, z\rangle_X| \\
&\leq 2\varepsilon + \varepsilon = 3\varepsilon
\end{aligned}$$

für $n \geq N$. Also gilt $\langle z_n^*, x\rangle_X \to \langle x^*, x\rangle_X$ für alle $x \in X$ und damit $z_n^* \rightharpoonup^* x^*$. □

[3]Diese Namensgebung ist streng genommen nicht korrekt: Das von Alaoglu bewiesene Resultat garantiert die *Überdeckungskompaktheit* der Einheitskugel in der schwach-$*$ Topologie, und sein Beweis verwendet daher auch schwereres Gerät. Die hier zitierte Aussage hat dagegen schon Banach bewiesen. Für metrische Räume ist die schwach-$*$ Topologie auf der Einheitskugel aber ebenfalls metrisierbar, so dass beide Aussagen äquivalent sind. In der Literatur wird daher auch dieser Spezialfall zumeist unter dem allgemeineren Namen zitiert.

Die Separabilität von X ist dabei unverzichtbar: Betrachte zum Beispiel die Funktionale $e_n^* \in \ell^\infty(\mathbb{K})^*$ mit $\langle e_n^*, x\rangle_{\ell^\infty} = x_n$. Dann gilt $\|e_n^*\|_{(\ell^\infty)^*} = 1$, aber da Folgen in $\ell^\infty(\mathbb{K})$ nicht konvergieren müssen, enthält $\{e_n^*\}_{n\in\mathbb{N}}$ keine schwach-$*$ konvergente Teilfolge.

Durch Skalierung erhält man aus Satz 11.6, dass jede abgeschlossene Kugel in X^* schwach-$*$ folgenkompakt ist. Insbesondere gilt der „schwach-$*$ Satz von Bolzano–Weierstraß".

Folgerung 11.7. *Ist X ein separabler normierter Raum, so hat jede beschränkte Folge in X^* eine schwach-$*$ konvergente Teilfolge.*

Da ein reflexiver Raum isomorph zum Dualraum seines Dualraums ist, erhalten wir daraus auch die schwache Folgenkompaktheit der entsprechenden Einheitskugel.

Satz 11.8 (Eberlein–Šmulian[4]). *Sei X ein normierter Raum. Ist X reflexiv, so ist die Einheitskugel B_X schwach folgenkompakt.*

Beweis. Wir führen die Aussage zurück auf Satz 11.6 von Banach–Alaoglu, brauchen dafür aber die Separabilität, weshalb wir nicht direkt für $X^{**} \cong X$ argumentieren können. Sei $\{x_n\}_{n\in\mathbb{N}} \subset B_X$ und betrachte $U := \operatorname{cl}(\operatorname{span}\{x_n : n \in \mathbb{N}\})$. Dann ist U als abgeschlossener Unterraum von X nach Satz 10.5 ebenfalls reflexiv und nach Definition separabel. Also ist $U^{**} \cong U$ separabel (denn $\{J_U x_n\}_{n\in\mathbb{N}}$ ist dicht in $U^{**} = J_U(U)$) und damit nach Folgerung 8.7 auch U^*. Die Folge $\{J_U x_n\}_{n\in\mathbb{N}} \subset U^{**}$ ist nun beschränkt in U^{**} (denn J_U ist eine Isometrie) und hat daher nach Folgerung 11.7 eine schwach-$*$ konvergente Teilfolge $\{J_U x_{n_k}\}_{k\in\mathbb{N}}$ mit $J_U x_{n_k} \rightharpoonup^* u^{**} \in U^{**}$. Da U reflexiv ist, existiert ein $x := J_U^{-1} u^{**} \in U \subset X$. Dann gilt für alle $x^* \in X^*$ wegen $x_n \in U$

$$\begin{aligned}\langle x^*, x_{n_k}\rangle_X = \langle x^*|_U, x_{n_k}\rangle_U &= \langle J_U x_{n_k}, x^*|_U\rangle_{U^*}\\ &\to \langle u^{**}, x^*|_U\rangle_{U^*} = \langle x^*|_U, x\rangle_U = \langle x^*, x\rangle_X\end{aligned}$$

und damit $x_{n_k} \rightharpoonup x$. □

Damit erhalten wir auch einen „schwachen Satz von Bolzano–Weierstraß".

Folgerung 11.9. *Ist X ein reflexiver normierter Raum, so hat jede beschränkte Folge in X eine schwach konvergente Teilfolge.*

[4]Auch diese Zuordnung ist eigentlich inkorrekt: Der Satz von Eberlein–Šmulian besagt, dass in einem Banachraum die schwache Folgenkompaktheit und die Überdeckungskompaktheit in der schwachen Topologie äquivalent sind. (Da die schwache Topologie im Gegensatz zur schwach-$*$ Topologie nicht metrisierbar ist, ist das eine nichttriviale Aussage.) Zusammen mit einem Satz von Goldstine, dass die Reflexivität von X äquivalent zur schwachen Überdeckungskompaktheit von B_X ist (siehe z. B. [Werner 2018, Satz VIII.3.18]), folgt daraus der hier bewiesene Spezialfall (der – für separable Räume – ebenfalls bereits von Banach bewiesen wurde). Auch dieser wird trotzdem öfter unter dem angegebenen Namen angewendet.

Wir haben nun alles zur Hand, um den Satz von Weierstraß (Folgerung 2.10) auf unendlich-dimensionale Räume zu übertragen.

Satz 11.10. *Sei X ein reflexiver normierter Raum, $U \subset X$ nichtleer, beschränkt, konvex und abgeschlossen, und $f : X \to \mathbb{R}$ schwach unterhalbstetig. Dann existiert ein $\bar{x} \in U$ mit*

$$f(\bar{x}) = \min_{x \in U} f(x).$$

Beweis. Da die Menge $\{f(x) : x \in U\} \subset \mathbb{R}$ nichtleer ist, existiert $M := \inf_{x \in U} f(x) < \infty$ (wobei der Fall $M = -\infty$, d. h. f ist nicht nach unten beschränkt, hier noch nicht ausgeschlossen ist). Aus den Eigenschaften des Infimums folgt dann, dass eine Folge $\{y_n\}_{n \in \mathbb{N}} \subset f(U) \subset \mathbb{R}$ existiert mit $y_n \to M$, d. h. es existiert eine Folge $\{x_n\}_{n \in \mathbb{N}} \subset U$ mit

$$f(x_n) = y_n \to M = \inf_{x \in U} f(x).$$

Eine solche Folge wird *Minimalfolge* genannt. Beachten Sie, dass wir aus der Konvergenz von $\{f(x_n)\}_{n \in \mathbb{N}}$ noch nicht auf die Konvergenz von $\{x_n\}_{n \in \mathbb{N}}$ schließen können!

Aus der Beschränktheit von U folgt insbesondere, dass die Minimalfolge beschränkt ist und daher nach Folgerung 11.9 eine schwach konvergente Teilfolge $\{x_{n_k}\}_{k \in \mathbb{N}}$ mit Grenzwert $\bar{x} \in X$ besitzt. Da U konvex und abgeschlossen ist, folgt nach Satz 11.5 aus $\{x_{n_k}\}_{k \in \mathbb{N}} \subset U$ auch $\bar{x} \in U$. Dieser Grenzwert ist Kandidat für einen Minimierer.

Aus der Definition der Minimalfolge folgt nun, dass auch für die Teilfolge $f(x_{n_k}) \to M$ gilt. Mit der schwachen Unterhalbstetigkeit von f und der Definition des Infimums erhalten wir daher

$$\inf_{x \in U} f(x) \le f(\bar{x}) \le \liminf_{k \to \infty} f(x_{n_k}) = M = \inf_{x \in U} f(x).$$

Das Infimum wird also in $\bar{x}$ angenommen, d. h. $-\infty < f(\bar{x}) = \min_{x \in U} f(x)$. □

Analog können wir für X^* mit Hilfe der schwach-∗-Konvergenz argumentieren, wenn X separabel ist. Diese (und ähnliche) Resultate sind fundamental für die Variationsrechnung.

Teil IV

HILBERTRÄUME

12 SKALARPRODUKTE UND ORTHOGONALITÄT

Besonders weitreichende Aussagen über lineare Operatoren sind in Hilberträumen möglich, in denen zu der algebraischen und topologischen Struktur von normierten Vektorräumen eine weitere geometrische Struktur hinzukommt: das Skalarprodukt. Wie wir im nächsten Kapitel sehen werden, erlaubt dies eine Charakterisierung, die komplett ohne Dualräume auskommt, und damit das Übertragen der Strukturtheorie der linearen Algebra vom euklidischen Vektorraum auf unendlichdimensionale (Hilbert-)Räume.

SKALARPRODUKTE

Die grundlegende Definition ist nicht anders als in der linearen Algebra.

Definition 12.1. Sei X ein Vektorraum über $\mathbb{K}$. Ein *Skalarprodukt* auf X ist eine Abbildung $(\cdot,\cdot)_X : X \times X \to \mathbb{K}$ mit den folgenden Eigenschaften:

(i) $(\lambda x_1 + x_2, y)_X = \lambda\,(x_1, y)_X + (x_2, x)_X$ für alle $x_1, x_2, y \in X$ und $\lambda \in \mathbb{K}$;

(ii) $(x, y)_X = \overline{(y, x)_X}$ für alle $x, y \in \mathbb{K}$;

(iii) $(x, x)_X \geq 0$ für alle $x \in X$ mit $(x, x)_X = 0$ genau dann, wenn $x = 0 \in X$.

In diesem Fall heißt das Paar $(X, (\cdot,\cdot)_X)$ *Prä-Hilbertraum*. Ist das Skalarprodukt aus dem Kontext offensichtlich, bezeichnen wir den Prä-Hilbertraum auch kurz mit X.

Aus (i) und (ii) folgt sofort

$$(x, \lambda y_1 + y_2)_X = \overline{\lambda}\,(x, y_1)_X + (x, y_2)_X \quad \text{für alle } x, y_1, y_2 \in X \text{ und } \lambda \in \mathbb{K}.$$

Das Skalarprodukt ist damit *sesquilinear* („anderthalbfach linear"); beachte, dass dies in der Literatur nicht einheitlich (d. h. manchmal als linear im zweiten Argument) definiert wird. Insbesondere gilt wegen $\lambda + \overline{\lambda} = 2\operatorname{Re}\lambda$ die *binomische Formel*

$$(x + y, x + y)_X = (x, x)_X + 2\operatorname{Re}\,(x, y)_X + (y, y)_X \quad \text{für alle } x, y \in X. \tag{12.1}$$

Aus (ii) folgt auch $(x, x)_X \in \mathbb{R}$ für alle $x \in X$ sogar im Fall $\mathbb{K} = \mathbb{C}$, so dass Eigenschaft (iii) sinnvoll ist. Trotz der formalen Ähnlichkeit sollte man das (sesquilineare, symmetrische) Skalarprodukt in X nicht mit der (bilinearen, nicht symmetrischen) dualen Paarung zwischen X und X^* verwechseln.

Eine fundamentale Eigenschaft des Skalarprodukts ist die *Cauchy–Schwarz-Ungleichung*.

Satz 12.2. *Sei* $(X, (\cdot, \cdot)_X)$ *ein Prä-Hilbertraum. Dann gilt*

$$|\,(x, y)_X\,| \leq \sqrt{(x, x)_X}\sqrt{(y, y)_X} \qquad \text{für alle } x, y \in X.$$

Beweis. Für alle $x, y \in X$ und $\lambda \in \mathbb{K}$ folgt aus den Eigenschaften des Skalarprodukts und $\lambda\bar{\lambda} = |\lambda|^2 \in \mathbb{R}$

$$0 \leq (x + \lambda y, x + \lambda y)_X = (x, x)_X + 2\,\mathrm{Re}\,\bar{\lambda}\,(x, y)_X + |\lambda|^2\,(y, y)_X\,.$$

Sei nun $y \neq 0$ (sonst ist die Aussage trivialerweise erfüllt). Dann folgt speziell für $\lambda = -\frac{(x,y)_X}{(y,y)_X}$

$$0 \leq (x, x)_X - 2\frac{|\,(x, y)_X\,|^2}{(y, y)_X} + \frac{|\,(x, y)_X\,|^2}{(y, y)_X} = (x, x)_X - \frac{|\,(x, y)_X\,|^2}{(y, y)_X}$$

und damit die Behauptung. □

Daraus folgt ein wesentliches Resultat: Die neue Struktur ist verträglich mit den bereits eingeführten Strukturen, genauso wie die Norm verträglich ist mit der (durch sie induzierten) metrischen Struktur.

Satz 12.3. *Sei* $(X, (\cdot, \cdot)_X)$ *ein Prä-Hilbertraum. Dann wird durch*

$$\|x\|_X := \sqrt{(x, x)_X}$$

eine Norm auf X *induziert. Ist* $(X, \|\cdot\|_X)$ *vollständig, so heißt* X Hilbertraum.

Beweis. Die Normeigenschaften folgen direkt aus denen des Skalarprodukts: Aus $\|x\|_X = 0$ folgt $(x, x)_X = 0$ und damit $x = 0$. Für $\lambda \in \mathbb{K}$ und $x \in X$ gilt

$$\|\lambda x\|_X^2 = (\lambda x, \lambda x)_X = \lambda\bar{\lambda}\,(x, x)_X = |\lambda|^2\|x\|_X^2$$

und damit die positive Homogenität. Für die Dreiecksungleichung verwenden wir die Cauchy–Schwarz-Ungleichung: Für alle $x, y \in X$ gilt nach Definition der Norm und wegen $\mathrm{Re}\,\lambda \leq |\lambda|$

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

Zu jedem Prä-Hilbertraum gehört also stets ein kanonischer normierter Vektorraum, zwischen denen wir in Folge nicht unterscheiden; wenn wir also von Normen, Umgebungen oder konvergenten Folgen in Hilberträumen reden, sind stets die zur induzierten Norm gehörenden gemeint.[1] Insbesondere folgt aus der Cauchy–Schwarz-Ungleichung in Satz 12.2 sofort, dass das Skalarprodukt in jeder Komponente stetig ist (bezüglich der induzierten Norm).

Umgekehrt lässt sich das Skalarprodukt durch die induzierte Norm ausdrücken: Für $\mathbb{K} = \mathbb{R}$ gilt

$$(x, y)_X = \frac{1}{4}\left(\|x + y\|_X^2 - \|x - y\|_X^2\right), \tag{12.2}$$

für $\mathbb{K} = \mathbb{C}$,

$$(x, y)_X = \frac{1}{4}\left(\|x + y\|_X^2 - \|x - y\|_X^2 + i\|x + iy\|_X^2 - i\|x - iy\|_X^2\right), \tag{12.3}$$

wie man mit Hilfe der binomischen Formel (12.1) recht einfach nachrechnet; die Gleichung (12.2) bzw. (12.3) wird als *Polarisationsformel* bezeichnet. Tatsächlich funktioniert dies nur für die induzierte Norm.

Satz 12.4. *Ein normierter Raum $(X, \|\cdot\|_X)$ ist genau dann ein Prä-Hilbertraum, wenn die* Parallelogramm-Identität

$$\|x + y\|_X^2 + \|x - y\|_X^2 = 2\left(\|x\|_X^2 + \|y\|_X^2\right) \qquad \textit{für alle } x, y \in X$$

gilt.

Beweis. Ist X ein Prä-Hilbertraum, so folgt die Parallelogramm-Identität direkt aus (12.1). Umgekehrt kann man mit Hilfe der Parallelogramm-Identität nachrechnen, dass durch (12.2) bzw. (12.3) ein Skalarprodukt definiert wird; die zum Teil recht aufwendige Rechnung ist z. B. in [Werner 2018, Satz V.1.7] zu finden. □

Mit Hilfe der Parallelogramm-Identität kann man nun recht einfach Beispiele finden.

Beispiel 12.5. Hilberträume sind zum Beispiel

(i) $\mathbb{K}^n$ mit dem Skalarprodukt $(x, y)_{\mathbb{K}^n} := \sum_{k=1}^n x_n\overline{y_n}$,

(ii) $\ell^2(\mathbb{K})$ mit dem Skalarprodukt $(x, y)_{\ell^2} := \sum_{k=1}^\infty x_n\overline{y_n}$,

(iii) $L^2(\Omega)$ mit dem Skalarprodukt $(x, y)_{L^2} := \int_\Omega x(t)\overline{y(t)}\,dt$,

jeweils mit der kanonischen Norm. Prä-Hilbertraum, aber nicht vollständig, ist

(iv) $c_c(\mathbb{K}) \subset \ell^2(\mathbb{K})$ mit dem Skalarprodukt aus (ii);

[1] Es gibt aber durchaus Situationen, in denen es sinnvoll ist, *nicht* die induzierte Norm zu verwenden.

keine Prä-Hilberträume sind

(v) $\ell^p(\mathbb{K})$ und $L^p(\Omega)$ für $p \neq 2$,

(vi) $C(K)$ für einen kompakten metrischen Raum $K \neq \{0\}$.

So, wie die Norm den geometrischen Begriff der „Länge" verallgemeinert, ist das Skalarprodukt eine Verallgemeinerung des „Winkels" – von besonderer Bedeutung ist auch hier der rechte Winkel. Wir sagen, $x, y \in X$ sind *orthogonal*, falls $(x, y)_X = 0$ gilt. In diesem Fall gilt der *Satz von Pythagoras*

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

Weiter heißt für eine beliebige Menge $A \subset X$

$$A^\perp := \left\{x \in X : (x, y)_X = 0 \text{ für alle } y \in A\right\}$$

das *orthogonale Komplement* von A in X. Auch hier sollte man trotz der formalen Ähnlichkeit nicht das orthogonale Komplement (als Teilmenge von X) mit dem Annihilator (als Teilmenge von X^*) verwechseln. Mit den selben Argumenten wie für letzteren zeigt man jedoch, dass $A^\perp$ stets abgeschlossen ist und $\operatorname{cl} A \subset (A^\perp)^\perp$ gilt.

PROJEKTIONEN

Wir kommen nun zu einem zentralen Satz der Hilbertraumtheorie, der eine *eindeutige* Projektion auf konvexe Mengen garantiert. Dabei ist sowohl die Vollständigkeit als auch die Parallelogramm-Identität wesentlich. (Vergleiche auch Satz 11.10.)

Satz 12.6 (Projektionssatz). *Sei X ein Hilbertraum und sei $C \subset X$ nichtleer, konvex und abgeschlossen. Dann existiert für jedes $x \in X$ ein eindeutiges $z \in C$ mit*

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

Beweis. Wir zeigen zuerst die Existenz mit Hilfe der Vollständigkeit von X. Setze dafür $d := \inf_{y\in C} \|y - x\|_X$ für gegebenes $x \in X$; dieses Infimum existiert, da C nichtleer und die Norm nicht-negativ ist. Aus den Eigenschaften des Infimums folgt dann die Existenz einer Folge $\{y_n\}_{n\in\mathbb{N}} \subset C$ mit $\|y_n - x\|_X \to d$. Wir zeigen nun, dass $\{y_n\}_{n\in\mathbb{N}}$ eine Cauchy-Folge ist. Aus der Parallelogramm-Identität folgt für alle $n, m \in \mathbb{N}$

$$2\left(\|y_n - x\|_X^2 + \|y_m - x\|_X^2\right) = \|(y_n + y_m) - 2x\|_X^2 + \|y_n - y_m\|_X^2$$

und damit

$$\|y_n - y_m\|_X^2 = 2\left(\|y_n - x\|_X^2 + \|y_m - x\|_X^2\right) - 4\|\tfrac{y_n+y_m}{2} - x\|_X^2. \tag{12.4}$$

Da C konvex ist, ist mit $y_n, y_m \in C$ auch $\frac{1}{2}y_n + \frac{1}{2}y_m \in C$, und daher folgt nach Definition von d

$$0 \leq \|y_n - y_m\|_X^2 \leq 2\left(\|y_n - x\|_X^2 + \|y_m - x\|_X^2\right) - 4d^2.$$

Nach Definition der Folge $\{y_n\}_{n\in\mathbb{N}}$ geht nun die rechte Seite gegen Null für $n, m \to \infty$. Damit ist $\{y_n\}_{n\in\mathbb{N}}$ eine Cauchy-Folge, die wegen der Vollständigkeit von X gegen ein $z \in X$ konvergiert. Da C abgeschlossen ist, gilt sogar $z \in C$. Aus der Stetigkeit der Norm folgt dann

$$\|z - x\|_X = \lim_{n\to\infty} \|y_n - x\|_X = d = \inf_{y\in C} \|y - x\|_X.$$

Für die Eindeutigkeit seien $z, \tilde{z} \in C$ mit $\|z - x\|_X = d = \|\tilde{z} - x\|_X$. Dann ist wegen der Konvexität von C auch $\frac{1}{2}z + \frac{1}{2}\tilde{z} \in C$, und daraus folgt wie in (12.4)

$$\|z - \tilde{z}\|_X^2 = 2\left(\|z - x\|_X^2 + \|\tilde{z} - x\|_X^2\right) - 4\|\tfrac{z+\tilde{z}}{2} - x\|_X^2 = 4d^2 - 4\|\tfrac{z+\tilde{z}}{2} - x\|_X^2 \leq 0,$$

d. h. $z = \tilde{z}$. □

Durch die Zuordnung $x \mapsto z$ wird also eine Abbildung $P_C : X \to C$ definiert, die man als *(metrische) Projektion* auf C bezeichnet. Diese lässt sich mit Hilfe des Skalarprodukts charakterisieren.

Lemma 12.7. *Sei X ein Hilbertraum und sei $C \subset X$ nichtleer, konvex und abgeschlossen. Dann sind für $x \in X$ und $z \in C$ äquivalent:*

(i) $z = P_C(x)$;

(ii) $\operatorname{Re}(z - x, y - z)_X \geq 0$ *für alle* $y \in C$.

Beweis. *(ii) ⇒ (i):* Aus der binomischen Formel und (ii) folgt für alle $y \in C$

$$\begin{aligned}\|y - x\|_X^2 &= \|(z - x) + (y - z)\|_X^2 = \|z - x\|_X^2 + 2\operatorname{Re}(z - x, y - z)_X + \|y - z\|_X^2 \\ &\geq \|z - x\|_X^2,\end{aligned}$$

d. h. $z = P_C(x)$ nach dem Projektionssatz 12.6.

(i) ⇒ (ii): Sei $z = P_C(x) \in C$ und $y \in C$ beliebig. Da C konvex ist, ist für alle $t \in (0, 1)$ auch $y_t := (1 - t)z + ty \in C$. Also gilt nach dem Projektionssatz 12.6

$$\begin{aligned}\|z - x\|_X^2 &\leq \|y_t - x\|_X^2 = \|(z - x) + t(y - z)\|_X^2 \\ &= \|z - x\|_X^2 + 2t\operatorname{Re}(z - x, y - z)_X + t^2\|y - z\|_X^2,\end{aligned}$$

woraus durch Subtraktion von $\|z - x\|_X^2$ und Division durch $2t > 0$ folgt

$$0 \leq \operatorname{Re}(z - x, y - z)_X + \frac{t}{2}\|y - z\|_X^2.$$

Grenzübergang $t \to 0$ ergibt nun (ii). □

Von besonderer Bedeutung ist dabei der Fall, dass C ein abgeschlossener Unterraum ist.

Folgerung 12.8. *Sei X ein Hilbertraum und sei $U \subset X$ ein abgeschlossener Unterraum. Dann sind für $x \in X$ und $z \in U$ äquivalent:*

(i') $z = P_U(x)$;

(ii') $(z - x, u)_X = 0$ *für alle* $u \in U$.

Beweis. Da jeder Unterraum konvex ist, können wir Lemma 12.7 anwenden und müssen nur noch zeigen, dass im Spezialfall eines Unterraums die Bedingung (ii) äquivalent zu (ii') ist.

(ii') ⟹ (ii): Sei $y \in U$ beliebig. Dann ist wegen $z \in U$ auch $u := y - z \in U$, so dass mit (ii') insbesondere (ii) gilt.

(ii) ⟹ (ii'): Sei $u \in U$ beliebig. Dann ist auch $y := u + z \in U$, womit aus (ii) folgt

$$\operatorname{Re}(z - x, u)_X = \operatorname{Re}(z - x, y - z)_X \geq 0 \qquad \text{für alle } u \in U.$$

Durch Einsetzen von $-u \in U$ sieht man, dass sogar $\operatorname{Re}(z - x, u)_X = 0$ für alle $u \in U$ gelten muss. Analog folgt durch Einsetzen von $-iu \in U$ wegen der Sesquilinearität des Skalarprodukts und $\operatorname{Re}(-ix) = \operatorname{Im}(x)$ auch $\operatorname{Im}(z - x, u)_X = 0$ und damit (ii'). □

In diesem Fall hat die Projektion zusätzliche nützliche Eigenschaften; man spricht dann auch von einer *orthogonalen Projektion.*

Satz 12.9. *Sei X ein Hilbertraum und sei $U \subset X$ ein abgeschlossener Unterraum. Dann gelten:*

(i) $P_U \in L(X, X)$;

(ii) $\|P_U\|_{L(X,X)} = 1$ *falls* $U \neq \{0\}$;

(iii) $\ker P_U = U^\perp$;

(iv) $P_{U^\perp} = \operatorname{Id} - P_U$.

Beweis. Die Aussagen folgen alle aus der Tatsache, dass nach Folgerung 12.8 genau dann $z = P_U(x)$ ist, wenn $z \in U$ und $z - x \in U^\perp$ gilt.

Zu (i): Da $U^\perp$ ein Unterraum ist, folgt für alle $\lambda_1, \lambda_2 \in \mathbb{K}$, $x_1, x_2 \in X$ und $z_1 = P_U(x_1)$, $z_2 = P_U(x_2)$ dass gilt

$$(\lambda_1 z_1 + \lambda_2 z_2) - (\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1(z_1 - x_1) + \lambda_2(z_2 - x_2) \in U^\perp,$$

d. h. $\lambda_1 P_U(x_1) + \lambda_2 P_U(x_2) = P_U(\lambda_1 x_1 + \lambda_2 x_2)$ und damit die Linearität. Aus $P_U(x) - x \in U^\perp$ für alle $x \in X$ folgt nun für $z := P_U(x)$

$$\|x\|_X^2 = \|x - z + z\|_X^2 = \|x - z\|_X^2 + 2\,\mathrm{Re}\,(x - z, z)_X + \|z\|_X^2 \geq \|z\|_X^2, \tag{12.5}$$

d. h. $\|P_U(x)\|_X = \|z\|_X \leq \|x\|_X$ und damit die Stetigkeit von P_U.

Zu (ii): Wegen (12.5) gilt zunächst $\|P_U\|_{L(X,X)} \leq 1$. Für $x \in U \setminus \{0\}$ ist nun $z := x \in U$ und $z - x = 0 \in U^\perp$, woraus $P_U(x) = x$ und damit $\|P_U\|_{L(X,X)} = 1$ folgt.

Zu (iii): Es gilt $P_U(x) = 0 \in U$ genau dann, wenn $0 - x = -x \in U^\perp$ ist. Da U ein Unterraum ist, ist letzteres äquivalent zu $x \in U^\perp$.

Zu (iv): Wir müssen zeigen, dass für $x \in X$ beliebig und $w := x - P_U(x)$ gilt $w \in U^\perp$ und $w - x \in (U^\perp)^\perp$. Ersteres folgt aus der Eigenschaft $P_U(x) - x \in U^\perp$ der Projektion auf U, letzteres aus

$$w - x = (x - P_U(x)) - x = -P_U(x) \in U \subset (U^\perp)^\perp. \qquad \square$$

Aus Satz 12.9 (iv) zusammen mit $U \cap U^\perp = \{0\}$ erhalten wir insbesondere die folgende *orthogonale Zerlegung*:

Folgerung 12.10. *Sei X ein Hilbertraum und $U \subset X$ ein nichtleerer abgeschlossener Unterraum. Dann gilt für alle $x \in X$*

$$x = u + u_\perp \qquad \text{für} \qquad u = P_U(x) \in U, \quad u_\perp = P_{U^\perp}(x) \in U^\perp,$$

und für $x \neq 0$ ist diese Zerlegung eindeutig.

Damit können wir – ganz ohne Hahn–Banach – eine zu Folgerung 8.8 analoge Aussage zeigen.

Folgerung 12.11. *Sei X ein Hilbertraum und sei $U \subset X$ ein Unterraum. Dann gilt $(U^\perp)^\perp = \mathrm{cl}\, U$.*

Beweis. Wörtlich wie im Beweis von Folgerung 8.8 zeigt man zunächst $\mathrm{cl}\, U \subset (U^\perp)^\perp$. Sei nun $x \in (U^\perp)^\perp$. Wir betrachten nun den abgeschlossenen Unterraum $V := \mathrm{cl}\, U$ und können daher nach Folgerung 12.10 schreiben $x = v + v^\perp$ mit $v \in V$ und $v^\perp \in V^\perp$. Aus $U \subset \mathrm{cl}\, U$ und der Definition des orthogonalen Komplements folgt sofort $V^\perp \subset U^\perp$ und damit $v^\perp \in U^\perp$.[2] Andererseits ist auch $v^\perp = x - v \in (U^\perp)^\perp$ wegen $v \in V \subset (U^\perp)^\perp$ wie bereits gezeigt und $x \in (U^\perp)^\perp$ nach Annahme, und daraus folgt

$$\|v^\perp\|_X^2 = \left(v^\perp, v^\perp\right)_X = \left(x - v, v^\perp\right)_X = 0.$$

Also ist $v^\perp = 0$ und daher $x = v \in V = \mathrm{cl}\, U$. $\square$

[2] Ein Grenzwertargument analog zu dem im Beweis von Folgerung 8.8 zeigt, dass sogar $V^\perp = U^\perp$ gilt.

Daraus erhalten wir ein Kriterium für die Invertierbarkeit von Operatoren auf Hilberträumen, welches (in einer etwas komplizierteren Form) den ersten Grundstein der modernen Theorie der partiellen Differentialgleichungen bildet.

Satz 12.12 (Lax–Milgram). *Sei X ein Hilbertraum und sei $T \in L(X,X)$. Existiert ein $\gamma > 0$ mit*

$$|\,(Tx,x)_X\,| \geq \gamma \|x\|_X^2 \qquad \textit{für alle } x \in X, \tag{12.6}$$

so ist T invertierbar mit $\|T^{-1}\|_{L(X,X)} \leq \gamma^{-1}$.

Beweis. Aus (12.6) folgt zusammen mit der Cauchy–Schwarz-Ungleichung

$$\|Tx\|_X \geq \gamma \|x\|_X \qquad \text{für alle } x \in X, \tag{12.7}$$

und damit nach Folgerung 9.8 mit $C = \gamma^{-1}$ sowohl die Injektivität von T als auch die Abgeschlossenheit von $\operatorname{ran} T$. Sei nun $x \in (\operatorname{ran} T)^\perp$. Dann ist $(Tx,x)_X = 0$, und aus (12.6) folgt $x = 0$, d. h. $(\operatorname{ran} T)^\perp = \{0\}$. Nach Folgerung 12.11 gilt dann

$$\operatorname{ran} T = ((\operatorname{ran} T)^\perp)^\perp = \{0\}^\perp = X,$$

und damit ist T surjektiv und daher invertierbar. Für beliebige $y \in X$ können wir daher $x := T^{-1}y$ in (12.7) einsetzen und erhalten

$$\|T^{-1}y\|_X \leq \frac{1}{\gamma}\|TT^{-1}y\|_X = \frac{1}{\gamma}\|y\|_X. \qquad \square$$

ORTHONORMALBASEN

In endlichdimensionalen Vektorräumen lässt sich die orthogonale Projektion auf einen Unterraum explizit mit Hilfe der Basisvektoren berechnen. Wir übertragen dies nun auf Hilberträume. Eine Teilmenge $S \subset X$ eines Prä-Hilbertraums nennen wir dafür *Orthonormalsystem*, wenn für alle $u, v \in S$ gilt

$$(u,v)_X = \begin{cases} 0 & \text{falls } u \neq v, \\ 1 & \text{falls } u = v. \end{cases}$$

(Ist nur die erste Bedingung erfüllt, d. h. gilt $\|u\|_X \neq 1$ für ein $u \in S$, so spricht man von einem *Orthogonalsystem.*) Die Frage nach der Existenz von Orthonormalsystemen lassen wir zunächst offen und untersuchen erst einmal ihre Eigenschaften. Als erstes betrachten wir die Projektion auf einen *endlichdimensionalen* Unterraum.

Lemma 12.13. *Sei X ein Prä-Hilbertraum, $S \subset X$ ein Orthonormalsystem, und $e_1, \dots, e_n \in S$. Dann gilt für $U = \operatorname{span}\{e_1, \dots, e_n\}$*

$$P_U x = \sum_{k=1}^n (x, e_k)_X\, e_k \qquad \textit{für alle } x \in X.$$

Beweis. Zu $x \in X$ betrachten wir $z := \sum_{k=1}^{n} (x, e_k)_X\, e_k$. Nach Konstruktion ist dann $z \in U$, und aus der Orthogonalität der e_j folgt für alle $1 \leq j \leq n$

$$\left(x - z, e_j\right)_X = \left(x, e_j\right)_X - \sum_{k=1}^{n} (x, e_k)_X \left(e_k, e_j\right)_X = \left(x, e_j\right)_X - \left(x, e_j\right)_X = 0.$$

Also ist auch $(x - z, u)_X = 0$ für alle $u \in U$ und damit $z = P_U x$ nach Folgerung 12.8. □

Folgerung 12.14. *Sei X ein Prä-Hilbertraum und $\{e_1, \ldots, e_n\} \subset X$ ein endliches Orthonormalsystem. Dann gilt die* Besselsche Ungleichung

$$\sum_{k=1}^{n} |\, (x, e_k)_X\,|^2 \leq \|x\|_X^2 \qquad \textit{für alle } x \in X.$$

Beweis. Betrachte $U = \operatorname{span}\{e_1, \ldots, e_n\}$. Dann gilt für alle $x \in X$

$$\|P_U x\|_X^2 = \left(\sum_{k=1}^{n} (x, e_k)_X\, e_k, \sum_{j=1}^{n} \left(x, e_j\right)_X e_j\right)_X = \sum_{k=1}^{n} \sum_{j=1}^{n} (x, e_k)_X \overline{\left(x, e_j\right)_X} \left(e_k, e_j\right)_X = \sum_{k=1}^{n} |\, (x, e_k)_X\,|^2. \tag{12.8}$$

Die Besselsche Ungleichung folgt nun aus $\|P_U x\|_X \leq \|x\|_X$, siehe Satz 12.9 (ii). □

Die Frage ist nun, wann das auch für (abzählbar) unendlichdimensionale Unterräume funktioniert, d. h. wir in den obigen Summen den Grenzübergang $n \to \infty$ machen können.

Satz 12.15. *Sei X ein Hilbertraum und $S = \{e_n : n \in \mathbb{N}\}$ ein Orthonormalsystem. Dann sind äquivalent:*

(i) $\operatorname{span} S$ *ist dicht in X;*

(ii) für alle $x \in X$ gilt

$$x = \sum_{k=1}^{\infty} (x, e_k)_X\, e_k;$$

(iii) für alle $x \in X$ gilt die Parseval-Identität

$$\|x\|_X^2 = \sum_{k=1}^{\infty} |\, (x, e_k)_X\,|^2.$$

Ist eine dieser Eigenschaften erfüllt, so nennen wir S eine Orthonormalbasis.

Beweis. *(i) ⇒ (ii):* Sei $U_m := \operatorname{span}\{e_1, \dots, e_m\}$ und $P_m := P_{U_m}$. Sei weiterhin $x \in X$ beliebig und $\{x_n\}_{n\in\mathbb{N}} \subset \operatorname{span} S$ mit $x_n \to x$. Da Linearkombinationen nach Definition stets endlich sind, finden wir für jedes $x_n \in \operatorname{span} S$ ein $m_n \in \mathbb{N}$ mit $x_n \in U_{m_n}$, wobei wir ohne Beschränkung der Allgemeinheit annehmen dürfen, dass $\{m_n\}_{n\in\mathbb{N}}$ monoton wachsend ist. Dann folgt nach Definition der Projektion

$$0 \leq \|x - P_{m_n}x\|_X = \inf_{u\in U_{m_n}} \|x-u\|_X \leq \|x - x_n\|_X \to 0 \quad \text{für } n \to \infty.$$

Da die U_m geschachtelt sind, ist weiterhin $\{\|x - P_m x\|_X\}_{m\in\mathbb{N}}$ monoton fallend; also muss die gesamte Folge gegen Null konvergieren. Aus Lemma 12.13 folgt dann

$$0 \leq \|x - \textstyle\sum_{k=1}^m (x, e_k)_X\, e_k\|_X = \|x - P_m x\|_X \to 0 \quad \text{für } m \to \infty.$$

(ii) ⇒ (iii): Sei $x \in X$ beliebig. Für die Partialsummen $s_n := \sum_{k=1}^n (x, e_k)_X\, e_k$ gilt nach (12.8)

$$\|s_n\|_X^2 = (s_n, s_n)_X = \sum_{k=1}^n |\,(x, e_k)_X\,|^2.$$

Nach (ii) konvergiert nun $s_n \to x$ und damit $\|s_n\|_X \to \|x\|_X$, sodass durch Grenzübergang $n \to \infty$ auf beiden Seiten (iii) folgt.

(iii) ⇒ (ii): Analog rechnet man nach, dass für alle $x \in X$ und die Partialsummen s_n gilt

$$(x, s_n)_X = \sum_{k=1}^n |\,(x, e_k)_X\,|^2 = \|s_n\|_X^2.$$

Daraus folgt

$$\|x - s_n\|_X^2 = \|x\|_X^2 - 2\operatorname{Re}(x, s_n)_X + \|s_n\|_X^2 = \|x\|_X^2 - \|s_n\|_X^2 \to 0$$

und damit (ii).

(ii) ⇒ (i): Ist $\operatorname{span} S$ nicht dicht in X, so existiert ein $x \in X$ und $\varepsilon > 0$ mit $\|x - x_n\|_X > \varepsilon$ für alle Folgen $\{x_n\}_{n\in\mathbb{N}} \subset \operatorname{span} S$. Dann gilt dies insbesondere für die Partialsummenfolge $\{s_n\}_{n\in\mathbb{N}}$, und damit kann (ii) nicht gelten. □

Beachten Sie, dass (i) bereits impliziert, dass X separabel ist.

Zum Beispiel bildet in $\ell^2(\mathbb{K})$ die Folge der Einheitsvektoren $\{e_n\}_{n\in\mathbb{N}}$ eine Orthonormalbasis. Etwas komplizierter nachzuweisen ist, dass in $L^2((-\pi, \pi))$ für $\mathbb{K} = \mathbb{C}$ die Funktionen

$$e_k(t) = \frac{1}{\sqrt{2\pi}} e^{ikt}, \quad k \in \mathbb{Z},$$

eine Orthonormalbasis bilden. Damit lässt sich jede Funktion $f \in L^2((-\pi, \pi))$ schreiben als

$$f(t) = \sum_{k\in\mathbb{Z}} c_k e_k(t), \qquad c_k := (f, e_k)_{L^2} = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} f(t) e^{-ikt}\, dt.$$

Diese Reihe wird *Fourier-Reihe* genannt. (Analog nennt man die Reihe in Satz 12.15 (ii) manchmal auch *(verallgemeinerte) Fourier-Reihe* sowie $(x, e_k)_X$ *(verallgemeinerten) Fourier-Koeffizient.*)

Allgemeiner haben wir das folgende Resultat.

Satz 12.16. *Sei X ein unendlichdimensionaler Hilbertraum. Dann sind äquivalent:*

(i) X ist separabel;

(ii) X besitzt eine abzählbare Orthonormalbasis.

Beweis. (i) $\Rightarrow$ (ii): Sei $\{x_n : n \in \mathbb{N}\}$ dicht in X. Wir definieren nun rekursiv

$$\tilde{e}_n := x_n - \sum_{k=1}^{n-1} (x_n, e_k)_X \, e_k,$$

$$e_n := \begin{cases} \frac{\tilde{e}_n}{\|\tilde{e}_n\|_X} & \text{falls } \tilde{e}_n \neq 0, \\ 0 & \text{falls } \tilde{e}_n = 0. \end{cases}$$

Dann ist $\|e_n\|_X = 1$ für alle $n \in \mathbb{N}$ sowie $(e_n, e_k)_X = 0$ für alle $k < n \in \mathbb{N}$, d. h. $\{e_n : n \in \mathbb{N}\}$ ist ein Orthonormalsystem. Weiter ist $\operatorname{span}\{e_n : n \in \mathbb{N}\} = \operatorname{span}\{x_n : n \in \mathbb{N}\}$ dicht in X und damit $\{e_n : n \in \mathbb{N}\}$ sogar eine Orthonormalbasis.

(ii) $\Rightarrow$ (i): Ist $\{e_n : n \in \mathbb{N}\}$ eine abzählbare Orthonormalbasis, so ist die Menge aller endlichen, rationalen Linearkombinationen abzählbar sowie dicht in $\operatorname{span}\{e_n : n \in \mathbb{N}\}$ und damit auch in X. □

Die Konstruktion im ersten Schritt entspricht natürlich genau der *Gram–Schmidt-Orthonormalisierung* aus der linearen Algebra.[3]

Aus Satz 12.16 erhalten wir das folgende erstaunliche Resultat.

Folgerung 12.17 (Satz von Fischer–Riesz). *Jeder unendlichdimensionale separable $\mathbb{K}$-Hilbertraum ist isometrisch isomorph zu $\ell^2(\mathbb{K})$.*

[3] Ist X nicht separabel, kann man die Existenz einer (dann überabzählbaren) Orthonormalbasis stattdessen mit Hilfe des Zornschen Lemmas zeigen. Dabei muss man verwenden, dass eine Orthonormalbasis ein maximales Orthonormalsystem ist, d. h. in keinem größeren Orthonormalsystem enthalten ist; siehe [Kaballo 2018, Satz 6.6, 6.8]. Auch Satz 12.15 kann man auf diese Situation übertragen, da auch für ein überabzählbares Orthonormalsystem höchstens abzählbar viele Skalarprodukte $(x, e)_X$ von Null verschieden sind; siehe [Werner 2018, Lemma V.4.5].

Beweis. Sei X ein unendlichdimensionaler separabler Hilbertraum und $\{e_n : n \in \mathbb{N}\}$ eine Orthonormalbasis. Wir konstruieren nun einen isometrischen Isomorphismus $T : X \to \ell^2(\mathbb{K})$, indem wir jedem $x \in X$ eine Folge $Tx := \{y_k\}_{k\in\mathbb{N}}$ durch $y_k := (x, e_k)_X$ zuordnen. Aus der Parseval-Identität folgt dann $Tx = y \in \ell^2(\mathbb{K})$ sowie $\|Tx\|_{\ell^2} = \|x\|_X$, also ist T insbesondere stetig. Weiterhin ist T linear und injektiv. Für die Surjektivität sei $y \in \ell^2(\mathbb{K})$ gegeben. Da y quadratsummierbar ist, muss $\{\sum_{k=1}^n |y_k|^2\}_{n\in\mathbb{N}}$ eine Cauchy-Folge (in $\mathbb{R}$) sein. Damit ist auch $\{\sum_{k=1}^n y_k e_k\}_{n\in\mathbb{N}}$ eine Cauchy-Folge (in X), denn für alle $m < n \in \mathbb{N}$ gilt

$$\|\textstyle\sum_{k=m+1}^n y_k e_k\|_X^2 = \left(\displaystyle\sum_{k=m+1}^n y_k e_k, \sum_{j=m+1}^n y_j e_j\right)_X = \sum_{k=m+1}^n |y_k|^2.$$

Da X vollständig ist, konvergiert also die Reihe $\sum_{k=1}^\infty y_k e_k = x \in X$. Aus der Stetigkeit des Skalarprodukts folgt schließlich

$$[Tx]_k = (x, e_k)_X = \lim_{n\to\infty} \left(\sum_{j=1}^n y_j e_j, e_k\right)_X = y_k \qquad \text{für alle } k \in \mathbb{N}$$

und damit $Tx = y$. □

Alle unendlichdimensionalen separablen Hilberträume sind also isometrisch isomorph!

13 DER SATZ VON RIESZ

Wir betrachten nun, wie sich die in Teil III beschriebene Dualitätstheorie in Hilberträumen verhält. Jeder Hilbertraum X wird ja durch die induzierte Norm zu einem normierten Raum, dem ein Dualraum X^* zugeordnet ist. Formal ähnelt die duale Paarung $\langle \cdot, \cdot \rangle_X$ zwischen X und X^* dem Skalarprodukt $(\cdot, \cdot)_X$ auf X (vergleiche etwa Annihilator und orthogonales Komplement). Tatsächlich besteht zwischen beiden eine enge Beziehung.

Satz 13.1 (Darstellungssatz von Fréchet–Riesz). *Sei X ein Hilbertraum. Dann existiert zu jedem $x^* \in X^*$ genau ein* Riesz-Repräsentant $x \in X$ *mit*

$$\langle x^*, z \rangle_X = (z, x)_X \qquad \text{für alle } z \in X.$$

Weiterhin gilt $\|x\|_X = \|x^*\|_{X^*}$.

Beweis. Zuerst stellen wir fest, dass für festes $x \in X$ die Abbildung $T_x : z \mapsto (z, x)_X$ ein stetiges lineares Funktional auf X ist. Wir zeigen nun, dass die Abbildung

$$R_X : X \to X^*, \qquad x \mapsto T_x,$$

bijektiv und isometrisch ist, d. h. dass wir jedes $x^* \in X^*$ in der Form T_x für genau ein $x \in X$ darstellen können. Direkt aus der Definition folgt zunächst

$$|\langle R_X(x), z \rangle_X| = |\,(z, x)_X\,| \leq \|z\|_X \|x\|_X \qquad \text{für alle } z \in X,$$

mit Gleichheit für $z = x$. Division durch $\|z\|_X$ und Supremum über alle $z \in X \setminus \{0\}$ ergibt $\|R_X(x)\|_{X^*} = \|x\|_X$ für alle $x \in X$. Also ist R_X isometrisch und damit injektiv.

Für die Surjektivität sei $x^* \in X^*$ beliebig. Für $x^* = 0$ wählen wir $x = 0$. Ist $x^* \neq 0$, so ist $\ker x^*$ wegen Satz 8.3 ein echter, abgeschlossener Unterraum von X. Damit ist das orthogonale Komplement $(\ker x^*)^\perp \neq \{0\}$ (denn sonst würden wir mit Hilfe von Folgerung 12.11 erhalten, dass $\ker x^* = X$ ist). Also existiert ein $x \in (\ker x^*)^\perp \setminus \{0\}$. Insbesondere gilt $\langle x^*, x \rangle_X \neq 0$, denn sonst wäre $x \in (\ker x^*)^\perp \cap \ker x^*$ und damit nach Satz 12.9 (iv)

$$x = P_{\ker x^*}(x) = x - P_{(\ker x^*)^\perp}(x) = x - x = 0,$$

im Widerspruch zu $x \neq 0$. Sei nun $z \in X$ beliebig. Dann gilt für alle $\lambda \in \mathbb{K}$

$$\langle x^*, z - \lambda x \rangle_X = \langle x^*, z \rangle_X - \lambda \langle x^*, x \rangle_X.$$

Setzen wir $\lambda_z := \frac{\langle x^*, z\rangle_X}{\langle x^*, x\rangle_X}$, so ist daher $z - \lambda_z x \in \ker x^*$. Wegen $x \in (\ker x^*)^\perp$ folgt daraus $(z - \lambda_z x, x)_X = 0$. Zusammen erhalten wir

$$\frac{\langle x^*, z\rangle_X}{\langle x^*, x\rangle_X} = \lambda_z = \frac{(z, x)_X}{(x, x)_X}.$$

Auflösen ergibt dann

$$\langle x^*, z\rangle_X = \left(z, \frac{\overline{\langle x^*, x\rangle_X}}{\|x\|_X^2} x\right)_X,$$

d. h. $x^* = R_X\left(\frac{\overline{\langle x^*, x\rangle_X}}{\|x\|_X^2} x\right)$. Damit ist das gesuchte Urbild gefunden. □

Die Abbildung $R_X : X \to X^*$ nennt man *Riesz-Isomorphismus*, obwohl sie nur im Fall $\mathbb{K} = \mathbb{R}$ linear und damit wirklich ein isometrischer Isomorphismus ist. Im Fall $\mathbb{K} = \mathbb{C}$ ist sie zumindest *konjugiert linear*: Direkt aus der Sesquilinearität des Skalarprodukts folgt für alle $x, y \in X$ und $\alpha \in \mathbb{K}$ und für beliebiges $z \in X$

$$\begin{aligned}\langle R_X(\alpha x + y), z\rangle_X &= (z, \alpha x + y)_X = \overline{\alpha}\,(z, x)_X + (z, y)_X \\ &= \langle \overline{\alpha} R_X(x) + R_X(y), z\rangle_X.\end{aligned}$$

Mit Hilfe des Riesz-Isomorphismus kann man nun Eigenschaften zwischen einem Hilbertraum und seinem Dualraum übertragen.

Folgerung 13.2. *Sei X ein Hilbertraum. Dann gelten:*

(i) X^* ist ein Hilbertraum;

(ii) X ist reflexiv.

Beweis. Für (i) vergewissert man sich leicht, dass wegen der konjugierten Linearität und der Bijektivität von R_X auf X^* durch

$$(x^*, y^*)_{X^*} := \left(R_X^{-1} y^*, R_X^{-1} x^*\right)_X \qquad \text{für alle } x^*, y^* \in X^*$$

ein Skalarprodukt definiert wird. Also ist X^* ein Prä-Hilbertraum. Aus der Isometrie von R_X folgt auch

$$\|x^*\|_{X^*}^2 = \|R_X^{-1} x^*\|_X^2 = \left(R_X^{-1} x^*, R_X^{-1} x^*\right)_X = (x^*, x^*)_{X^*},$$

d. h. die Operatornorm wird durch dieses Skalarprodukt induziert. Da Dualräume bezüglich dieser Norm vollständig sind, ist X^* ein Hilbertraum.

Für (ii) müssen wir nachweisen, dass die kanonische Einbettung $J_X : X \to X^{**}$ surjektiv ist. Dafür zeigen wir, dass $J_X = R_{X^*} \circ R_X$ gilt. Seien dazu $x \in X$ und $x^* \in X^*$ beliebig. Dann gilt nach Definition der Riesz-Isomorphismen, des Skalarprodukts in X^* und der kanonischen Einbettung

$$\langle R_{X^*} R_X x, x^*\rangle_{X^*} = (x^*, R_X x)_{X^*} = \left(x, R_X^{-1} x^*\right)_X = \langle x^*, x\rangle_X = \langle J_X x, x^*\rangle_{X^*}.$$

Also ist J_X als Komposition zweier bijektiver Abbildungen bijektiv und damit surjektiv. □

Der Riesz-Isomorphismus erlaubt es auch, die schwache Konvergenz mit Hilfe des Skalarprodukts auszudrücken. Aus der Bijektivität von R_X folgt sofort, dass gilt

$$x_n \rightharpoonup x \qquad \text{genau dann, wenn} \qquad (x_n, z)_X \to (x, z)_X \quad \text{für alle } z \in X.$$

Da wir bei der Definition der schwachen Konvergenz also ohne Dualraum auskommen, ist der Unterschied zur starken Konvergenz nicht mehr so groß.

Folgerung 13.3. *Sei X ein Hilbertraum und $\{x_n\}_{n\in\mathbb{N}} \subset X$. Dann sind äquivalent:*

(i) $x_n \to x$;

(ii) $x_n \rightharpoonup x$ *und* $\|x_n\|_X \to \|x\|_X$.

Beweis. Wir wissen bereits, dass in normierten Räumen jede stark konvergente Folge auch schwach konvergiert und dass die Norm stetig ist. Umgekehrt folgt im Hilbertraum aus der schwachen Konvergenz und der Konvergenz der Norm sofort

$$\|x_n - x\|_X^2 = \|x_n\|_X^2 - 2\,\mathrm{Re}\,(x_n, x)_X + \|x\|_X^2 \to \|x\|_X^2 - 2\,(x, x)_X + \|x\|_X^2 = 0. \qquad \square$$

Der Konvergenzbegriff in (ii) kann auch in Banachräumen als unabhängiger Begriff nützlich sein und wird dann als *strikte Konvergenz* bezeichnet.

Ebenso können wir über den Riesz-Isomorphismus den adjungierten Operator „zurück nach X ziehen": Sind X und Y Hilberträume, so definieren wir für $T \in L(X, Y)$ den *Hilbertraum-adjungierten Operator*

$$T^\star : Y \to X, \qquad T^\star y = R_X^{-1} T^* R_Y y,$$

wobei $T^* : Y^* \to X^*$ der übliche (Banachraum-)adjungierte Operator ist. (Gefahr der Verwechslung wird in Folge kaum bestehen.) Dann folgt aus der Definition von Riesz-Isomorphismus und adjungiertem Operator

$$\begin{aligned}(Tx, y)_Y &= \langle R_Y y, Tx\rangle_Y = \langle T^* R_Y y, x\rangle_X = \left(x, R_X^{-1} T^* R_Y\right)_X \\ &= \left(x, T^\star y\right)_X \qquad \text{für alle } x \in X, y \in Y.\end{aligned} \tag{13.1}$$

Unmittelbar aus der Definition erhalten wir auch die folgenden Rechenregeln.

Lemma 13.4. *Seien X, Y, Z Hilberträume und $S, T \in L(X, Y)$, $R \in L(Y, Z)$. Dann gelten:*

(i) $(S + T)^\star = S^\star + T^\star$;

(ii) $(\lambda T)^\star = \bar{\lambda} T^\star$ *für alle* $\lambda \in \mathbb{K}$;

(iii) $(R \circ T)^\star = T^\star \circ R^\star$.

Wörtlich wie im Banachraum-Fall zeigt man auch zu Satz 9.6 analoge Aussagen für die orthogonalen Komplemente von Kern und Bild (unter Verwendung von Folgerung 12.11 anstelle von Folgerung 8.8). Die folgenden Eigenschaften sind oft nützlich, wobei wir hier und in Folge kurz $T^\star T := T^\star \circ T$ schreiben.

Lemma 13.5. *Seien X, Y Hilberträume und $T \in L(X, Y)$. Dann gilt:*

(i) $T^{\star\star} = T$;

(ii) $\|T^\star\|_{L(Y,X)} = \|T\|_{L(X,Y)}$;

(iii) $\|T^\star T\|_{L(X,X)} = \|T\|^2_{L(X,Y)}$.

Beweis. Zu (i): Aus (13.1) folgt sofort

$$\left(y, T^{\star\star}x\right)_Y = \left(T^\star y, x\right)_X = \overline{(x, T^\star y)_X} = \overline{(Tx, y)_Y} = (y, Tx)_Y$$

für alle $x \in X$ und $y \in Y$. Insbesondere gilt daher für $y = T^{\star\star}x - Tx$

$$\|T^{\star\star}x - Tx\|_Y^2 = \left(y, T^{\star\star x - Tx}\right)_Y = 0$$

und damit $T^{\star\star}x = Tx$ für alle $x \in X$ wie behauptet.

Zu (ii): Zunächst gilt für alle $y \in Y$

$$\|T^\star y\|_X = \|R_X^{-1}T^*R_Y y\|_X = \|T^*R_Y y\|_{X^*} \leq \|T^*\|_{L(Y^*,X^*)}\|R_Y y\|_{Y^*} = \|T\|_{L(X,Y)}\|y\|_Y,$$

da sowohl der Riesz-Isomorphismus als auch (nach Lemma 9.1) die Abbildung $T \mapsto T^*$ isometrisch ist. Supremum über alle $y \in B_Y$ ergibt $\|T^\star\|_{L(Y,X)} \leq \|T\|_{L(X,Y)}$. Aus (i) folgt dann auch $\|T\|_{L(X,Y)} = \|T^{\star\star}\|_{L(X,Y)} \leq \|T^\star\|_{L(Y,X)}$ und damit (ii).

Zu (iii): Aus Folgerung 4.6 und (ii) folgt

$$\|T^\star T\|_{L(X,X)} \leq \|T^\star\|_{L(Y,X)}\|T\|_{L(X,Y)} = \|T\|^2_{L(X,Y)}.$$

Die umgekehrte Ungleichung erhalten wir aus

$$\|Tx\|_Y^2 = (Tx, Tx)_Y = \left(x, T^\star Tx\right)_X \leq \|T^\star T\|_{L(X,X)}\|x\|_X^2$$

und Supremum über alle $x \in B_X$. □

Der Riesz-Isomorphismus erlaubt also, die komplette Dualitätstheorie alleine auf Elementen in X aufzubauen. Man unterscheidet daher oft nicht zwischen Elementen $x^* \in X^*$ und ihren Repräsentanten $R_X^{-1}x^* \in X$, d. h. man behandelt R_X wie die Identität – man sagt, X^* wird mit X *identifiziert*. Insbesondere wird nicht zwischen Banachraum- und Hilbertraum-Adjungierten unterschieden. Dies ist aber nicht in jedem Fall sinnvoll! Eine häufig auftauchende Situation ist, wenn man es mit zwei Hilberträumen X und Y zu tun hat, wobei X stetig und dicht in Y eingebettet ist, aber beide mit unterschiedlichen Skalarprodukten versehen sind. In diesem Fall ist Y^* stetig in X^* eingebettet; identifiziert man Y^* mit Y (d. h. betrachtet man R_Y als Identität), so erhält man das *Gelfand-Tripel* $X \hookrightarrow Y \cong Y^* \hookrightarrow X^*$. Würde man nun auch X^* mit X identifizieren, verlören die Einbettungen jede Aussagekraft; man muss sich also entscheiden. (Natürlich hat man trotzdem den Riesz-Isomorphismus R_X, man kann ihn nur nicht wie die Identität behandeln.) Besonders relevant ist diese Situation in der Theorie der partiellen Differentialgleichungen (wo R_X^{-1} oft genau dem Lösen einer Differentialgleichung entspricht).

Literatur

H. W. Alt (2012), *Lineare Funktionalanalysis. Eine anwendungsorientierte Einführung*. 6. Aufl., Springer-Verlag, Berlin, doi: 10.1007/978-3-642-22261-0.

M. Brokate (2013), Funktionalanalysis, Vorlesungsskript, Zentrum Mathematik, TU München, url: http://www-m6.ma.tum.de/~brokate/fun_ws13.pdf.

M. Brokate & G. Kersting (2019), *Maß und Integral*, 2. Aufl., Mathematik Kompakt, Birkhäuser, Basel, doi: 10.1007/978-3-0348-0988-7.

C. Clason (2019), *Einführung in die Funktionalanalysis*, Mathematik Kompakt, Birkhäuser, Basel, doi: 10.1007/978-3-030-24876-5.

J. B. Conway (2014), *A Course in Point Set Topology*, Undergraduate Texts in Mathematics, Springer, Cham, doi: 10.1007/978-3-319-02368-7.

M. Dobrowolski (2010), *Angewandte Funktionalanalysis*, Springer-Verlag, Berlin, doi: 10.1007/978-3-642-15269-6.

H. W. Engl, M. Hanke & A. Neubauer (1996), *Regularization of Inverse Problems*, Bd. 375, Mathematics and Its Applications, Kluwer Academic Publishers Group, Dordrecht, doi: 10.1007/978-94-009-1740-8.

O. Forster (2017), *Analysis 2*, 11. Aufl., Springer Spektrum, Wiesbaden, doi: 10.1007/978-3-658-19411-6.

W. Kaballo (2018), *Grundkurs Funktionalanalysis*, 2. Aufl., Springer Spektrum, Heidelberg, doi: 10.1007/978-3-662-54748-9.

W. A. J. Luxemburg & M. Väth (2001), The existence of non-trivial bounded functionals implies the Hahn-Banach extension theorem, *Z. Anal. Anwendungen* 20(2), 267–279, doi: 10.4171/zaa/1015.

W. Rudin (2008), *Analysis*, 4. Aufl., R. Oldenbourg Verlag, München, doi: 10.1090/stml/001/09.

E. Schechter (1997), *Handbook of Analysis and Its Foundations*, Academic Press, Inc., San Diego, CA.

G. Wachsmuth (2013), Funktionalanalysis, Vorlesungsskript, Fakultät für Mathematik, TU Chemnitz.

D. Werner (2009), *Einführung in die Höhere Analysis*, 2. Aufl., Springer, Berlin, doi: 10.1007/978-3-540-79696-1.

D. Werner (2018), *Funktionalanalysis*, 8. Aufl., Springer Spektrum, Berlin, doi: 10.1007/978-3-662-55407-4.