Universität Regensburg WS 2013/14

Vorlesungsmitschrift (*kein* offizielles Skript)

Funktionalanalysis

Prof. Dr. Harald Garcke

Version vom 4. April 2014

Gesetzt in LaTeX von Johannes Prem

Inhaltsverzeichnis

1 Einführung: Wovon handelt die Funktionalanalysis?

Zum Beispiel von der *Analysis auf Banachräumen* (vollständigen normierten Vektorräumen)

1.1. Auf $\mathbb{R}^n$ definiere

$$\forall x \in \mathbb{R}^n: \quad \|x\|_2 \coloneqq |x| = \Big(\sum_{i=1}^{n} x_i^2\Big)^{\frac{1}{2}}$$

1.2 (Funktionen auf kompakten Teilmengen des $\mathbb{R}^n$)**.**
Zum Beispiel: $K \subset \mathbb{R}^n$ kompakt, z. B. $K = [0,1]$.

$$C^0(K) \coloneqq \{f \mid f\colon K \to \mathbb{R} \text{ stetig}\}$$

wird Banachraum mit der Norm:

$$\|f\|_\infty \coloneqq \sup_{x\in K} |f(x)| \ < \infty$$

1.3 (Operatoren auf $C^0([0,1])$)**.**
Definiere

$$L\big(C^0([0,1]),\, C^0([0,1])\big) \coloneqq \big\{T\colon C^0([0,1]) \to C^0([0,1]) \bigm| T \text{ ist linear und stetig}\big\}$$

Beispiele:

$$(Tf)(x) \coloneqq g(x)\, f(x) \qquad \text{wobei } g \in C^0([0,1])$$

$$(Tf)(x) \coloneqq \sum_{i=1}^{n} f(x_i)\, L_i(x) \qquad \text{wobei } 0 \le x_0 < x_1 < \dots < x_n \le 1$$

$$L_i\colon \text{ Lagrange-Basis-Fkt.:} \quad L_i(x) = \prod_{\substack{j=0\\ j\neq i}}^{n} \frac{x - x_j}{x_i - x_j}$$

$$(Tf)(x) \coloneqq \int_0^1 K(x,y)\, f(y)\, \mathrm{d}y \qquad \text{wobei } K \in C^0\left([0,1]^2\right)$$

Bemerkung: $L\big(C^0([0,1]),\, C^0([0,1])\big)$ wird zu einem Banachraum mit der Operatornorm

$$\|T\|_{L(C^0,C^0)} \coloneqq \sup_{f\neq 0} \frac{\|Tf\|_{C^0}}{\|f\|_{C^0}}$$

1.4. Welche Besonderheiten ergeben sich in unendlich-dimensionalen Räumen?

(1) Problem in ∞-dimensionalen Vektorräumen: Wenig sinnvolle Aussagen ohne Topologie möglich

(2) Für $T\colon \mathbb{R}^n \to \mathbb{R}^n$ linear gilt:

$$T \text{ surjektiv} \quad \iff \quad T \text{ injektiv}$$

Im ∞-dim. ist dies i. A. falsch.
Beispiel:

$$c_* := \left\{ x = (x_k)_{k\in\mathbb{N}} \;\middle|\; x_k \in \mathbb{R},\ \exists\, \bar{k} \in \mathbb{N}\ \forall\, \ell > \bar{k}\colon\ x_\ell = 0 \right\}$$

c_* modelliert „Folgen, die irgendwann abbrechen". Außerdem enthält c_* den $\mathbb{R}^n$ für $n \in \mathbb{N}$ beliebig groß.

Definiere die sog. *Shift-Abbildung* wie folgt:

$$T(x_1, x_2, x_3, \dots) := (0, x_1, x_2, x_3, \dots)$$

Dann ist T injektiv, aber nicht surjektiv.

(3) Grundproblem der linearen Algebra: Finde Normalformen für lineare Abbildungen.
Ziel: Verallgemeinerung auf ∞-dim. Räume.

$$\left.\begin{array}{l}\text{Diagonalisierbarkeit}\\ \text{symmetrischer Matrizen}\end{array}\right\} \rightsquigarrow \left\{\begin{array}{l}\text{Spektralsatz für}\\ \text{kompakte, normale Operatoren}\end{array}\right.$$

$$\left.\begin{array}{l}\text{Jordansche}\\ \text{Normalform}\end{array}\right\} \rightsquigarrow \left\{\begin{array}{l}\text{Spektralsatz für}\\ \text{kompakte Operatoren}\end{array}\right.$$

(4) Kompaktheit
In ∞-dim. Banachräumen ist die abgeschlossene Einheitskugel *nicht* kompakt.
Beispiel c_*: Nutze die Norm

$$\|x\|_{c_*} := \max_{n\in\mathbb{N}} |x_n|$$

und die Einheitsvektoren $e_i = (\delta_{ij})_{j\in\mathbb{N}} = (0, \dots, 0, 1, 0, \dots)$ (wobei die 1 an der i-ten Stelle steht). Dann gilt:

$$\|e_i\|_{c_*} = 1 \qquad \text{und} \qquad \|e_i - e_k\|_{c_*} = 1 \text{ für } i \neq k$$

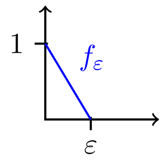
Also hat $(e_i)_{i\in\mathbb{N}}$ *keine* konvergente Teilfolge, woraus folgt, dass die Einheitskugel nicht kompakt ist.

(5) Nicht alle Normen sind zueinander äquivalent.
Beispiel: Betrachte auf $C^0([0,1])$ die Normen

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

$$\|f\|_{L^2} := \sqrt{\int_0^1 \big(f(x)\big)^2 \,\mathrm{d}x}$$

Es gilt $\|f\|_{L^2} \leq \|f\|_\infty$. Aber: Es gibt keine Konstante $c \in \mathbb{R}_{>0}$, so dass für alle $f \in C^0([0,1])$ gilt: $\|f\|_\infty \leq c\,\|f\|_{L^2}$. Betrachte dazu:



Es gilt: $\|f_\varepsilon\|_\infty = 1,\ \|f_\varepsilon\|_{L^2} \leq \sqrt{\varepsilon}$.

Außerdem gilt:

$$\big(C^0([0,1]),\ \|\cdot\|_\infty\big) \text{ ist Banachraum}$$
$$\big(C^0([0,1]),\ \|\cdot\|_{L^2}\big) \text{ ist normierter Vektorraum (aber nicht vollständig)}$$

Funktionalanalysis lässt sich sinnvoll nur in vollständigen Räumen entwickeln. Deshalb werden wir nicht vollständige Räume vervollständigen (siehe 2.18).

2 Grundstrukturen der Funktionalanalysis

2.1 (Topologie)**.** Sei X eine Menge, $\mathcal{T}$ ein System von Teilmengen. Dann heißt $\mathcal{T}$ *Topologie (auf X)*, falls gilt:

(T1) $\emptyset \in \mathcal{T},\ X \in \mathcal{T}$

(T2) $\mathcal{T}' \subset \mathcal{T} \implies \bigcup \mathcal{T}' \in \mathcal{T}$

(T3) $T_1, T_2 \in \mathcal{T} \implies T_1 \cap T_2 \in \mathcal{T}$

Ein topologischer Raum $(X, \mathcal{T})$ heißt *Hausdorff-Raum*, falls er zusätzlich das Hausdorffsche Trennungsaxiom erfüllt:

(T4) $\forall\, x_1, x_2 \in X, x_1 \neq x_2\ \ \exists\, U_1, U_2 \in \mathcal{T}\colon\ U_1 \cap U_2 = \emptyset \wedge x_i \in U_i$

Mengen in $\mathcal{T}$ heißen *offene Mengen*. Komplemente offener Mengen heißen *abgeschlossene Mengen*.

Eine Menge $W \subset X$ mit $x \in W$ für die eine offene Menge U mit $x \in U$ und $U \subset W$ existiert, heißt *Umgebung von x*.

Seien $(X, \mathcal{T}_X)$ und $(Y, \mathcal{T}_Y)$ topologische Räume, so heißt $f\colon X \to Y$ *stetig*, falls die Urbilder offener Mengen stets offen sind. (Formal: $\forall\, U' \in \mathcal{T}_Y\colon\ f^{-1}(U') \in \mathcal{T}_X$)

Eine Abbildung $f\colon X \to Y$ heißt *stetig in $x \in X$*, falls

$$f(x) \in V \in \mathcal{T}_Y \quad \implies \quad \exists\, U \in \mathcal{T}_X\colon\ x \in U \subset f^{-1}(V)$$

(d. h. $f^{-1}(V)$ ist Umgebung von x).

2.2. Ist X ein $\mathbb{K}$-Vektorraum mit $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$, so heißt $(X, \mathcal{T})$ *topologischer Vektorraum*, falls $(X, \mathcal{T})$ ein topologischer Raum ist und die Abbildungen

$$\begin{aligned} X \times X \to X, &\quad (x, y) \mapsto x + y \\ \mathbb{K} \times X \to X, &\quad (\alpha, x) \mapsto \alpha\, x \end{aligned}$$

stetig sind. („Algebraische und topologische Struktur sind verträglich“)

2.3 (Metrik)**.** Ein Tupel (X, d) heißt *metrischer Raum*, falls X eine Menge ist und $d\colon X \times X \to \mathbb{R}$ folgende Bedingungen für alle $x, y, z \in X$ erfüllt:

(M1) $d(x, y) \geq 0$ und $d(x, y) = 0 \iff x = y$

(M2) $d(x,y) = d(y,x)$

(M3) $d(x,z) \leq d(x,y) + d(y,z)$

Konvergenz:
$(x_n)_{n\in\mathbb{N}}$ heißt Cauchy-Folge, falls:

$$d(x_k, x_\ell) \to 0 \quad \text{für } k, \ell \to \infty$$

x heißt Grenzwert von $(x_n)_{n\in\mathbb{N}}$ (Notation: $x = \lim_{n\to\infty} x_n$ oder: $x_n \to x$ für $n \to \infty$), falls:

$$d(x_n, x) \to 0 \quad \text{für } n \to \infty$$

(X,d) heißt *vollständig*, falls jede Cauchy-Folge einen Grenzwert in X besitzt.

Abstand von Mengen $A, B \subset X$:

$$\operatorname{dist}(A,B) \coloneqq \inf\{d(a,b) \mid a \in A,\ b \in B\}$$

Für $A \subset X$ und $x \in X$ definieren wir: $\operatorname{dist}(x,A) \coloneqq \operatorname{dist}(\{x\},A)$.

Für $r \in \mathbb{R}_{>0}$ sowie $A \subset X,\ x \in X$ definieren wir:

$$\begin{aligned} B_r(A) &\coloneqq \{x \in X \mid \operatorname{dist}(x,A) < r\} \\ B_r(x) &\coloneqq B_r(\{x\}) \\ \operatorname{diam}(A) &\coloneqq \sup\{d(a_1,a_2) \mid a_1, a_2 \in A\} \end{aligned}$$

Wir sagen A ist *beschränkt*, falls $\operatorname{diam}(A) < \infty$.

2.4 (Topologie von Metriken). Sei (X,d) ein metrischer Raum und $A \subset X$.

$$A^\circ \coloneqq \{x \in X \mid \exists\, r \in \mathbb{R}_{>0} \colon\ B_r(x) \subset A\} \quad \text{ist das } \textit{Innere von } A.$$

$$\overline{A} \coloneqq \{x \in X \mid \forall\, r \in \mathbb{R}_{>0} \colon\ B_r(x) \cap A \neq \emptyset\} \quad \text{ist der } \textit{Abschluss von } A.$$

$$\partial A \coloneqq \overline{A} \setminus A^\circ \quad \text{ist der } \textit{Rand von } A.$$

Wir sagen, dass A offen ist, falls $A^\circ = A$ gilt, und dass A abgeschlossen ist, falls $\overline{A} = A$ gilt.

Durch die Definition $\mathcal{T} \coloneqq \{A \subset X \mid A \text{ offen}\}$ wird $(X,\mathcal{T})$ zu einem hausdorffschen topologischen Raum.

2.5 (Fréchet-Metrik). Sei X ein Vektorraum. Eine Abbildung $d\colon X \to \mathbb{R}$ heißt *Fréchet-Metrik*, falls für alle $x, y \in X$ gilt:

(F1) $d(x) \geq 0 \quad \text{und} \quad d(x) = 0 \iff x = 0$

(F2) $d(-x) = d(x)$

(F3) $d(x+y) \leq d(x) + d(y)$

Dann ist $(x, y) \mapsto d(x - y)$ eine Metrik auf X.

Beispiel: Fréchet-Metriken auf $\mathbb{R}$:

$$x \mapsto |x|^\alpha \quad \text{mit } 0 < \alpha \leq 1$$
$$x \mapsto \frac{|x|}{1 + |x|}$$

2.6 (Norm). X sei ein $\mathbb{K}$-Vektorraum (mit $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$).
Eine Abbildung $\|\cdot\| \colon X \to \mathbb{R}$ heißt *Norm*, falls folgende Bedingungen für alle $x, y \in X, \alpha \in \mathbb{K}$ erfüllt sind:

(N1) $\|x\| \geq 0 \quad \text{und} \quad \|x\| = 0 \iff x = 0$

(N2) $\|\alpha\, x\| = |\alpha|\, \|x\|$

(N3) $\|x + y\| \leq \|x\| + \|y\|$

Dann ist $x \mapsto \|x\|$ eine Fréchet-Metrik. Wir nennen X *Banachraum*, falls X mit einer gegebenen Norm vollständig ist.

X ist eine *Banachalgebra*, falls X eine Algebra ist (d. h. es gibt ein Produkt auf X, das dem Assoziativgesetz und Distributivgestz genügt) und $\|x \cdot y\| \leq \|x\| \cdot \|y\|$ für alle $x, y \in X$ gilt.

2.7 (Skalarprodukt). Sei X ein $\mathbb{K}$-Vektorraum.

(a) $\langle \cdot\,, \cdot \rangle \colon X \times X \to \mathbb{K}$ heißt *Hermitische Form* ($\mathbb{K} = \mathbb{R}$ symmetrische Bilinearform, $\mathbb{K} = \mathbb{C}$ symmetrische Sesquilinearform), falls für alle $x, x_1, x_2, y \in X, \alpha \in \mathbb{K}$ gilt:

(S1) $\langle x, y \rangle = \overline{\langle y, x \rangle}$

(S2) $\langle \alpha\, x, y \rangle = \alpha \, \langle x, y \rangle$

(S3) $\langle x_1 + x_2, y \rangle = \langle x_1, y \rangle + \langle x_2, y \rangle$

(Es folgt: für alle $x \in X$ gilt $\langle x, x \rangle \in \mathbb{R}$.)

(b) $\langle \cdot\,, \cdot \rangle$ heißt *positiv semidefinit*, falls

(S4') $\langle x, x \rangle \geq 0$

und *positiv definit*, falls

(S4) $\langle x, x \rangle \geq 0 \quad \text{und} \quad \langle x, x \rangle = 0 \iff x = 0$

gilt.

(c) $\langle \cdot\,, \cdot \rangle$ heißt *Skalarprodukt*, falls (S1)–(S4) erfüllt sind. Dann ist $\|x\| := \sqrt{\langle x, x \rangle}$ eine Norm auf X und wir nennen X dann einen *Prä-Hilbertraum*. Falls X zusätzlich vollständig ist, so heißt X *Hilbertraum*

Beispiele:

i) $\mathbb{R}^n$ mit $\langle x, y\rangle = \sum_{i=1}^n x_i\, y_i$, $\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$

ii) $X = C^0(K, \mathbb{R})$ für $K \subset \mathbb{R}^n$ kompakt.

$$\langle f, g\rangle \coloneqq \int_K f(x)\, g(x)\, \mathrm{d}x$$

Dann ist $\big(C^0(K), \langle\cdot,\cdot\rangle\big)$ ein Prä-Hilbertraum (aber kein Hilbertraum!)

Satz 2.8. Sei $\langle\,\cdot\,,\,\cdot\,\rangle$ ein Skalarprodukt auf einem Vektorraum X. Dann gelten:

(1) Cauchy-Schwarz-Ungleichung (CSU): $\forall\, x, y \in X\colon \quad |\langle x, y\rangle| \leq \|x\| \cdot \|y\|$.
Gleichheit gilt nur, falls y ein Vielfaches von x ist.

(2) Dreiecksungleichung: $\forall\, x, y \in X\colon \quad \|x + y\| \leq \|x\| + \|y\|$

(3) Parallelogrammidentität: $\forall\, x, y \in X\colon \quad \|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$

Bemerkung: Im Fall $\mathbb{K} = \mathbb{R}$ folgt aus der CSU für $x, y \in X \setminus \{0\}$:

$$\left\langle \frac{x}{\|x\|}, \frac{y}{\|y\|} \right\rangle \in [-1, 1] \tag{$*$}$$

D. h. es gibt genau ein $\theta \in [0, \pi]$, s. d.

$$\left\langle \frac{x}{\|x\|}, \frac{y}{\|y\|} \right\rangle = \cos\theta.$$

Wir interpretieren θ als den Winkel zwischen x und y.

Beweis von Satz 2.8.

(3)

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

Ersetze y durch $-y$ und addiere beide Gleichungen.

(1) Ersetze in $(*)$ y durch $-\frac{\langle x,y\rangle}{\|y\|^2}\, y$ (o. E. $y \neq 0$). Dann ergibt sich:

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

Es folgt die CSU. In der ersten Zeile gilt bei $\leq$ die Gleichheit genau dann, wenn x ein Vielfaches von y ist.

(2)

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\underbrace{\operatorname{Re}\langle x,y\rangle}_{\leq |\langle x,y\rangle| \leq \|x\|\,\|y\|} \leq (\|x\|+\|y\|)^2$$

■

2.9 (Vergleich von Topologien)**.** Seien $\mathcal{T}_1, \mathcal{T}_2$ zwei Topologien auf einer Menge X. Wir sagen $\mathcal{T}_2$ ist *stärker* (oder *feiner*) als $\mathcal{T}_1$ und $\mathcal{T}_1$ ist *schwächer* (oder *gröber*) als $\mathcal{T}_2$, falls $\mathcal{T}_1 \subset \mathcal{T}_2$ gilt.

Sind d_1, d_2 zwei Metriken auf X und $\mathcal{T}_1, \mathcal{T}_2$ die induzierten Topologien (siehe 2.4), so heißt die Metrik d_1 *stärker (bzw. schwächer)* als d_2, falls $\mathcal{T}_1$ stärker (bzw. schwächer) als $\mathcal{T}_2$ ist. Die Metriken heißen äquivalent, falls $\mathcal{T}_1 = \mathcal{T}_2$. Entsprechend heißt eine Norm stärker bzw. schwächer als eine zweite, wenn dies für die induzierten Metriken gilt. Analog für Äquivalenz von Normen.

2.10 (Vergleich von Normen)**.** Seien $\|\cdot\|_1$ und $\|\cdot\|_2$ zwei Normen auf einem $\mathbb{K}$-Vektorraum X. Dann gilt:

(1) $\|\cdot\|_2$ ist stärker als $\|\cdot\|_1$ genau dann, wenn es ein $c \in \mathbb{R}_{>0}$ gibt mit:

$$\forall\, x \in X: \quad \|x\|_1 \leq c\,\|x\|_2$$

(2) Die beiden Normen sind genau dann äquivalent, wenn es $c, C \in \mathbb{R}_{>0}$ gibt mit:

$$\forall\, x \in X: \quad c\,\|x\|_2 \leq \|x\|_1 \leq C\,\|x\|_2$$

Beweis. (1) Es sei $B^i_r(x) = \{x' \in X \mid \|x-x'\|_i < r\}$ und $\mathcal{T}_i$ sei die von $\|\cdot\|_i$ induzierte Topologie.
Sei $\mathcal{T}_1 \subset \mathcal{T}_2$. Da $B^1_1(0) \in \mathcal{T}_1$ gilt, ist $B^1_1(0)$ offen bezüglich $\mathcal{T}_1$ und bezüglich $\mathcal{T}_2$. Es liegt 0 im Inneren (bezüglich $\|\cdot\|_2$) von $B^1_1(0)$. Somit gilt $B^2_\varepsilon(0) \subset B^1_1(0)$ für ein $\varepsilon \in \mathbb{R}_{>0}$. Daher gilt für $x \in X \setminus \{0\}$:

$$\left\| \frac{\varepsilon\, x}{2\|x\|_2} \right\|_2 = \frac{\varepsilon}{2} < \varepsilon$$

$$\implies \quad \left\| \frac{\varepsilon\, x}{2\,\|x\|_2} \right\|_1 < 1 \qquad \implies \quad \|x\|_1 < \frac{2}{\varepsilon}\,\|x\|_2$$

Gilt umgekehrt die Ungleichung in (1) so ist für alle $x \in X$ und $r \in \mathbb{R}_{>0}$

$$B^2_r(x) \subset B^1_{cr}(x)$$

Sei nun $A \in \mathcal{T}_1$. Dann ist $A = A^\circ$ bezüglich $\mathcal{T}_1$. D. h. zu $x \in A$ existiert ein $\varepsilon \in \mathbb{R}_{>0}$, so dass $B^1_\varepsilon(x) \subset A$. Also gilt:

$$B^2_{\varepsilon/c}(x) \subset A$$

Dies zeigt $A \in \mathcal{T}_2$.

(2) Wende den ersten Teil zweimal an.

■

Satz 2.11. Auf einem endlich-dimensionalen Vektorraum sind alle Normen äquivalent. Endlich-dimensionale Vektorräume sind Banachräume. Endlich-dimensionale Unterräume normierter Räume sind abgeschlossen.

Beweis. Sei X ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und $\|\cdot\|$ eine Norm. Sei $e_1, \ldots, e_n$ eine Basis von $(X, \|\cdot\|)$. Jedem $x \in X$ mit $x = \sum_{i=1}^n \alpha_i\, e_i$ ordnen wir den Vektor $\alpha = (\alpha_1, \ldots, \alpha_n)^{\mathsf{t}} \in \mathbb{K}^n$ zu.

Die Abbildungen

$$\begin{aligned} \mathbb{K}^n &\to X \to \mathbb{R} \\ \alpha &\mapsto x \;\; \mapsto \|x\| \end{aligned}$$

sind stetig.

Daher nimmt $\|x\|$ auf der kompakten Menge

$$S := \{\alpha \mid \|\alpha\|_2 = 1\}$$

ein Maximum M und ein Minimum m an. (Dabei gilt $m > 0$, da $\|x\| > 0$ für alle $x \in S$.) Damit gilt für x mit $\|\alpha(x)\|_2 = 1$

$$m \leq \|x\| \leq M.$$

Für allgemeine $x \neq 0$ gilt

$$\left\| \alpha\left(\frac{x}{\|\alpha(x)\|_2}\right) \right\|_2 = 1 \quad \text{und somit} \quad m \leq \left\| \frac{x}{\|\alpha(x)\|_2} \right\| \leq M$$

Dies zeigt die Äquivalenz einer beliebigen Norm zur Norm $x \mapsto \|\alpha(x)\|_2$. Damit sind zwei beliebige Normen äquivalent.

Die Vollständigkeit von X folgt aus der Vollständigkeit von $(\mathbb{K}^n, \|\cdot\|_2)$. Die Tatsache, dass endlich-dimensionale Räume abgeschlossen sind, folgt mit Aufgabe 1 von Übungsblatt 2.

■

2.12 (Folgenräume). Wir bezeichnen mit $\mathbb{K}^{\mathbb{N}}$ die Menge aller Folgen über $\mathbb{K}$, d. h.

$$\mathbb{K}^{\mathbb{N}} := \left\{ x = (x_n)_{n\in\mathbb{N}} \;\middle|\; x_n \in \mathbb{K} \text{ für alle } n \in \mathbb{N} \right\}$$

Es gilt:

1) $\mathbb{K}^{\mathbb{N}}$ ist ein metrischer Raum mit der Fréchet-Metrik

$$\varrho(x) := \sum_{n\in\mathbb{N}} 2^{-n} \frac{|x_n|}{1 + |x_n|} \qquad \text{für } x = (x_n)_{n\in\mathbb{N}}$$

2) Ist $(x^k)_{k\in\mathbb{N}} = ((x_i)^k_{i\in\mathbb{N}})_{k\in\mathbb{N}}$ eine Folge in $\mathbb{K}^\mathbb{N}$ und ist $x = (x_i)_{i\in\mathbb{N}} \in \mathbb{K}^\mathbb{N}$, so gilt:

$$\varrho(x^k - x) \to 0 \text{ für } k \to \infty \qquad \Longleftrightarrow \qquad \forall\, i \in \mathbb{N}\colon\ x_i^k \to x_i \text{ für } k \to \infty.$$

(Vergleiche Aufgabe 2 von Blatt 1.)

3) $\mathbb{K}^\mathbb{N}$ ist mit dieser Metrik vollständig.

4) Definiere für $x = (x_i)_{i\in\mathbb{N}} \in \mathbb{K}^\mathbb{N}$:

$$\begin{aligned} \|x\|_{\ell^p} &:= \Big(\sum_{i\in\mathbb{N}} |x_i|^p\Big)^{\frac{1}{p}} \ \in [0,\infty] \quad \text{für } 1 \le p < \infty \\ \|x\|_{\ell^\infty} &:= \sup_{i\in\mathbb{N}} |x_i| \ \in [0,\infty] \end{aligned}$$

Wir betrachten für $1 \le p \le \infty$ die Mengen

$$\ell^p(\mathbb{K}) := \big\{x \in \mathbb{K}^\mathbb{N} \ \big|\ \|x\|_{\ell^p} < \infty\big\}$$

Diese Räume sind normierte Vektorräume und auch vollständig (also Banachräume). (Beweis, siehe später.) Wenn der Körper $\mathbb{K}$ aus dem Kontext klar ist, lassen wir diesen in der Notation auch weg.

5) Interessante Unterräume von ℓ^∞ sind:

$$\begin{aligned} c &:= \big\{x \in \ell^\infty \ \big|\ \lim_{i\to\infty} x_i \text{ existiert}\big\} \qquad \text{und} \\ c_0 &:= \big\{x \in \ell^\infty \ \big|\ \lim_{i\to\infty} x_i = 0\big\}. \end{aligned}$$

Beide Räume versehen wir mit der $\|\cdot\|_\infty$-Norm. Es gilt: $c_0 \subset c \subset \ell^\infty$

6) Der Raum ℓ^2 besitzt das Skalarprodukt

$$(x,y) := \sum_{i=1}^{\infty} x_i\, y_i \quad \text{für } x, y \in \ell^2$$

Lemma 2.13 (Youngsche Ungleichung)**.** Es seien $p, p' \in (1,\infty)$, so dass $\frac{1}{p} + \frac{1}{p'} = 1$ gilt. Dann gilt für alle $a, b \in \mathbb{R}_{>0}$:

$$ab \le \frac{1}{p}\, a^p + \frac{1}{p'}\, b^{p'}.$$

Beweis.

$$\begin{aligned} \log(ab) = \log(a) + \log(b) &= \frac{1}{p}\log(a^p) + \frac{1}{p'}\log(b^{p'}) \\ &\le \log\Big(\frac{1}{p}\, a^p + \frac{1}{p'}\, b^{p'}\Big) \end{aligned}$$

Die letzte Ungleichheit folgt daraus, dass der Logarithmus eine konkave Funktion ist. Da außerdem exp monoton ist, folgt hieraus die Behauptung.

∎

Satz 2.14 (Höldersche Ungleichung auf ℓ^p). Es sei $1 \le p, p' \le \infty$ mit $\frac{1}{p} + \frac{1}{p'} = 1$. Für $x \in \ell^p$ und $y \in \ell^{p'}$ ist $xy \in \ell^1$ (dabei sei für $x = (x_n)_{n\in\mathbb{N}}$ und $y = (y_n)_{n\in\mathbb{N}}$ das Produkt definiert als: $xy \coloneqq (x_n\, y_n)_{n\in\mathbb{N}}$) und es gilt:

$$\|xy\|_{\ell^1} \le \|x\|_{\ell^p} \cdot \|y\|_{\ell^{p'}}.$$

Beweis. Falls $p = \infty$ setzte $p' = 1$ (und umgekehrt). In diesem Fall ist der Beweis einfach. Sei nun $1 < p < \infty$ und $\|x\|_{\ell^p} > 0, \|y\|_{\ell^{p'}} > 0$. Die Youngsche Ungleichung (Lemma 2.13) liefert:

$$\frac{|x_k|\;|y_k|}{\|x\|_{\ell^p}\,\|y\|_{\ell^{p'}}} \le \frac{1}{p}\,\frac{|x_k|^p}{\|x\|_{\ell^p}^p} + \frac{1}{p'}\,\frac{|y_k|^{p'}}{\|y\|_{\ell^{p'}}^{p'}}$$

Die Reihe über die Terme der rechten Seite ist eine konvergente Reihe und damit folgt aus dem Majorantenkriterium, dass $xy \in \ell^1$ erfüllt sein muss.

■

Satz 2.15. Der Raum ℓ^p ist für $1 \le p \le \infty$ ein Banachraum.

Beweis. Die Vollständigkeit von ℓ^1 ist eine Übungsaufgabe. Für ℓ^p folgt dies ähnlich (vergleiche mit dem späteren Beweis über $L^p(\mu)$.) Die Normeigenschaften abgesehn von der $\triangle$-Ungleichung ergeben sich einfach. Wir zeigen die $\triangle$-Ungleichung für $p \in (1, \infty)$. Es seien also $p, p' \in (1, \infty)$ mit $\frac{1}{p} + \frac{1}{p'} = 1$. Dann gilt:

$$\begin{aligned}
\|x+y\|_{\ell^p}^p &= \sum_{k=1}^{\infty} |x_k + y_k|^p \le \sum_{k=1}^{\infty} |x_k|\,|x_k + y_k|^{p-1} + \sum_{k=1}^{\infty} |y_k|\,|x_k + y_k|^{p-1} \\
&\overset{(\star)}{\le} \Big(\sum_{k=1}^{\infty} |x_k|^p\Big)^{\frac{1}{p}} \Big(\sum_{k=1}^{\infty} |x_k + y_k|^{(p-1)p'}\Big)^{\frac{1}{p'}} + \Big(\sum_{k=1}^{\infty} |y_k|^p\Big)^{\frac{1}{p}} \Big(\sum_{k=1}^{\infty} |x_k + y_k|^{(p-1)p'}\Big)^{\frac{1}{p'}} \\
&= (\|x\|_{\ell^p} + \|y\|_{\ell^p})\,(\|x+y\|_{\ell^p}^{p-1})
\end{aligned}$$

Bei $(\star)$ geht die Höldersche Ungleichung ein. Dies zeigt $\|x+y\|_{\ell^p} \le \|x\|_{\ell^p} + \|y\|_{\ell^p}$.

■

2.16 (Stetige Funktionen auf kompakten Mengen). Ist $K \subset \mathbb{R}^n$ abgeschlossen und beschränkt (also nach Heine-Borel äquivalenterweise kompakt) und Y ein Banachraum über $\mathbb{K}$, so ist $C^0(K, Y)$ ein Unterraum von $B(K, Y)$. (Vergleiche Aufgabe 1 von Blatt 2.)

Satz: Mit $\|f\|_{C^0} \coloneqq \|f\|_{\infty} \coloneqq \sup_{x\in K} |f(x)|$ wird $C^0(K, Y)$ ein Banachraum.

Beweis. Jedes $f \in C^0(K, Y)$ ist beschränkt, denn: Zu $x \in K$ existiert ein $\delta_x \in \mathbb{R}_{>0}$ mit $f\big(B_{\delta_x}(x)\big) \subset B_1\big(f(x)\big)$. Da K kompakt ist, existieren endlich viele Punkte $x_1, \ldots, x_m \in K$ mit

$$K \subset \bigcup_{i=1}^{m} B_{\delta_{x_i}}(x_i).$$

Es folgt:

$$f(K) \subset \bigcup_{i=1}^{m} B_1\big(f(x_i)\big).$$

Die rechte Menge ist beschränkt, also ist auch f beschränkt.

Aufgabe 1 (iv) von Blatt 2 zeigt, dass $B(K,Y)$ ein Banachraum ist. Eine Cauchy-Folge in $C^0(K,Y)$ ist auch eine Cauchy-Folge in $B(K,Y)$. Da $B(K,Y)$ vollständig ist, besitzt jede Cauchy-Folge einen Grenzwert.

Sei jetzt $(f_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $C^0(K,Y)$ mit Grenzwert f in $B(Y,K)$. Für $x,y \in K$ gilt:

$$\|f(y) - f(x)\| \le \underbrace{\|f_i(y) - f_i(x)\|}_{\substack{\to 0 \text{ für } y\to x\\ \text{und jedes } i}} + \underbrace{2\|f - f_i\|_\infty}_{\to 0 \text{ für } i\to\infty}$$

Dies beweist $f \in C^0(K,Y)$.

■

2.17 (Räume differenzierbarer Funktionen)**.** Es sei $\Omega \subset \mathbb{R}^n$ offen und beschränkt und $m \in \mathbb{N}_0$. Dann definieren wir:

$$\begin{aligned} C^m(\overline{\Omega}) \coloneqq \{f\colon \Omega \to \mathbb{R} \mid\ & f \text{ ist } m\text{-mal stetig diffenzierbar auf } \Omega \text{ und für } s \in \mathbb{N}^n \\ & \text{mit } |s| \le m \text{ ist } \partial^s f \text{ auf } \overline{\Omega} \text{ stetig fortsetzbar } \}. \end{aligned}$$

(Dabei ist s ein Multiindex mit $|s| = s_1 + \cdots + s_n$.)

Satz: Der Raum $C^m(\overline{\Omega})$ ist mit der Norm

$$\|f\|_{C^m(\overline{\Omega})} \coloneqq \sum_{|s|\le m} \|\partial^s f\|_{C^0(\overline{\Omega})}$$

ein Banachraum.

Beweis. Wir beweisen die Vollständigkeit von $C^1(\overline{\Omega})$. (Der Fall $m > 1$ folgt induktiv.) Ist $(f_k)_{k\in\mathbb{N}}$ eine Cauchy-Folge in $C^1(\overline{\Omega})$, so sind $(f_k)_{k\in\mathbb{N}}$ und $(\partial_i f_k)_{k\in\mathbb{N}}$ Cauchy-Folgen in $C^0(\overline{\Omega})$ für alle $i \in \{1,\dots,n\}$.

Daher existieren f und g_i in $C^0(\overline{\Omega})$, so dass $f_k \to f$ sowie $\partial_i f_k \to g_i$ gleichmäßig für $k \to \infty$ in $C^0(\overline{\Omega})$. Für $x \in \Omega$ und y nahe x mit $x_t \coloneqq (1-t)x + ty$ folgt aus dem HDI:

$$f_k(x_1) - f_k(x_0) = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} f_k(x_t)\,\mathrm{d}t = \int_0^1 (y-x)\cdot \nabla f_k(x_t)\,\mathrm{d}t$$

Es folgt (mit $g = (g_1,\dots,g_n)$):

$$\begin{aligned} \|f_k(y) - f_k(x) - (y-x)\cdot\nabla f_k(x)\| &= \left\| \int_0^1 \big((y-x)\cdot\nabla f_k(x_t) - (y-x)\cdot\nabla f_k(x)\big)\,\mathrm{d}t \right\| \\ &\overset{\text{CSU}}{\le} \int_0^1 \|\nabla f_k(x_t) - \nabla f_k(x)\|\,\mathrm{d}t\,\|y-x\| \\ &\le \Big(2\|\nabla f_k - g\|_\infty + \sup_{t\in[0,1]} \|g(x_t) - g(x)\|\Big)\,\|y-x\| \end{aligned}$$

Für $k \to \infty$ gilt dann:

$$\|f(y) - f(x) - (y-x)\cdot g(x)\| \leq \underbrace{\sup_{0\leq t\leq 1} \|g(x_t) - g(x)\|}_{\to 0 \text{ für } y \to x \text{ wegen Stetigkeit von } g} \|y-x\|$$

Dies bedeutet f ist in x diff'bar mit $\nabla f(x) = g(x)$.

■

2.18 (Vervollständigung)**.** Sei (X, d) ein metrischer Raum. Wir definieren

$$\tilde{X} := \big\{x = (x_i)_{i\in\mathbb{N}} \;\big|\; x \text{ ist Cauchy-Folge in } X\big\}$$

zusammen mit der Äquivalenzrelation

$$(x_i)_{i\in\mathbb{N}} = (y_i)_{i\in\mathbb{N}} \;\text{ in } \tilde{X} \;:\Longleftrightarrow\; \big(d(x_j, y_j)\big)_{j\in\mathbb{N}} \text{ ist Nullfolge.}$$

Führe Metrik auf $\tilde{X}$ ein: für $(x_i)_{i\in\mathbb{N}}, (y_i)_{i\in\mathbb{N}} \in \tilde{X}$ sei

$$\tilde{d}\big((x_i)_{i\in\mathbb{N}}, (y_i)_{i\in\mathbb{N}}\big) := \lim_{j\to\infty} d(x_j, y_j).$$

Satz:

i) Dann ist $(\tilde{X}, \tilde{d})$ ein vollständiger metrischer Raum.

ii) Durch $J(x) := (x)_{j\in\mathbb{N}}$ ist eine injektive Abbildung $J\colon X \to \tilde{X}$ definiert, welche isometrisch ist, d. h. für alle $x, y \in X$ gilt

$$\tilde{d}\big(J(x), J(y)\big) = d(x, y).$$

iii) Es liegt $J(X)$ dicht in $\tilde{X}$.

Beweis. Für $\tilde{x} = (x_i)_{i\in\mathbb{N}}$ und $\tilde{y} = (y_i)_{i\in\mathbb{N}}$ in $\tilde{X}$ gilt (mithilfe der sog. *Vierecksungleichung*):

$$|d(x_j, y_j) - d(x_i, y_i)| \leq d(x_j, x_i) + d(y_j, y_i) \to 0 \qquad \text{für } i, j \to \infty.$$

Somit existiert $\tilde{d}(\tilde{x}, \tilde{y}) = \lim_{j\to\infty} d(x_j, y_j)$. Für $\tilde{x}^1 = \tilde{x}^2$ und $\tilde{y}^1 = \tilde{y}^2$ in $\tilde{X}$ folgt:

$$|d(x_j^2, y_j^2) - d(x_j^1, y_j^1)| \to 0 \qquad \text{für } i \to \infty.$$

Dies zeigt, dass $\tilde{d}$ wohldefiniert ist. Außerdem gilt:

$$\tilde{d}(\tilde{x}, \tilde{y}) = 0 \qquad \Longleftrightarrow \qquad \tilde{x} = \tilde{y},$$

was direkt aus der Definition der Äquivalenzrelation folgt. Die $\triangle$-Ungleichung und Symmetrie übertragen sich direkt.

Zur Vollständigkeit: Es sei $(x^k)_{k\in\mathbb{N}}$ eine Cauchy-Folge in $\tilde{X}$, mit $x^k = (x^k_j)_{j\in\mathbb{N}}$ für alle $k \in \mathbb{N}$. Zu $k \in \mathbb{N}$ wähle $j_k \in \mathbb{N}$, so dass $d(x^k_i, x^k_j) \le 1/k$ für alle $i, j \ge j_k$ erfüllt ist. Dann gilt:

$$\begin{aligned} d(x^k_{j_k}, x^\ell_{j_k}) &\le d(x^k_{j_k}, x^k_j) + d(x^k_j, x^\ell_j) + d(x^\ell_j, x^\ell_{j_\ell}) \\ &\le \frac{1}{k} + d(x^k_j, x^\ell_j) + \frac{1}{\ell} \qquad \text{für } j \ge j_k, j_\ell \\ &\to \frac{1}{k} + \tilde{d}(x^k, x^\ell) + \frac{1}{\ell} \qquad \text{für } j \to \infty \\ &\to 0 \qquad \text{für } k, \ell \to \infty \end{aligned}$$

Also ist $x^\infty := (x^\ell_{j_\ell})_{\ell\in\mathbb{N}}$ in $\tilde{X}$ und es gilt:

$$\begin{aligned} \tilde{d}(x^\ell, x^\infty) \longleftarrow\ & d(x^\ell_k, x^\infty_k) \qquad \text{für } k \to \infty \\ &\le d(x^\ell_k, x^\ell_{j_\ell}) + d(x^\ell_{j_\ell}, x^k_{j_k}) \\ &\le \frac{1}{\ell} + d(x^\ell_{j_\ell}, x^k_{j_k}) \qquad \text{für } k \ge j_\ell \\ &\to 0 \qquad \text{für } k, \ell \to \infty \end{aligned}$$

Es gilt also $x^\ell \to x^\infty$. Da $(x^k)_{k\in\mathbb{N}}$ eine beliebige Cauchy-Folge in $\tilde{X}$ war, hat also jede Cauchy-Folge einen Grenzwert.

Die Aussagen ii) und iii) sind eine einfache Übung.

■

3 Lineare Operatoren

Definition 3.1.

(a) Seien X, Y zwei $\mathbb{K}$-Vektorräume mit Topologien $\mathcal{T}_X, \mathcal{T}_Y$. Wir definieren

$$L(X,Y) := \{T\colon X \to Y \mid T \text{ ist linear und stetig}\}.$$

Elemente in $L(X,Y)$ heißen *lineare Operatoren von X nach Y*. (Für $T \in L(X,Y)$ und $x \in X$ schreiben wir auch oft Tx statt $T(x)$.)

(b) Der *Dualraum* von X ist

$$X' := L(X, \mathbb{K})$$

und Elemente aus X' nennen wir *lineare Funktionale*.

Beispiele 3.2.

1) Gelte $X = C^2(\overline{\Omega})$ und $Y = C^0(\overline{\Omega})$ für $\Omega \subset \mathbb{R}^n$ offen. Betrachte dann $T\colon X \to Y$ mit

$$(Tu)(x) := -\Delta u(x)$$

für alle $u \in C^2(\overline{\Omega}), x \in \overline{\Omega}$.

2) Sei $\Omega \subset \mathbb{R}^n$ offen und beschränkt und sei $K\colon \overline{\Omega} \times \overline{\Omega} \to \mathbb{R}$ stetig. Sei dann T für alle $u \in C^0(\overline{\Omega}), x \in \overline{\Omega}$ gegeben durch:

$$(Tu)(x) := \int_\Omega K(x,y)\, u(y)\, \mathrm{d}y.$$

Lemma 3.3. Seien X, Y normierte Vektorräume und sei $T\colon X \to Y$ linear. Dann sind die folgenden Aussagen äquivalent:

(1) T ist stetig, also $T \in L(X,Y)$.

(2) T ist stetig in x_0 für ein $x_0 \in X$.

(3) Es gilt für die Operatornorm von T:

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

(4) Es existiert ein $C \in \mathbb{R}_{\geq 0}$, so dass für alle $x \in X$ gilt: $\|Tx\|_Y \leq C\, \|x\|_X$. (Bemerkung: $C = \|T\|_{L(X,Y)}$ ist die kleinste solche Zahl.)

Beweis. (1) $\Longrightarrow$ (2): klar.

(2) $\Longrightarrow$ (3): Es gibt ein $\delta \in \mathbb{R}_{>0}$, so dass

$$T\big(\overline{B_\delta(x_0)}\big) \subset \overline{B_1\big(T(x_0)\big)}$$

erfüllt ist. Für x mit $\|x\|_X \leq 1$ folgt $x_0 + \delta x \in \overline{B_\delta(x_0)}$ und daraus: $T(x_0+\delta x) \in \overline{B_1\big(T(x_0)\big)}$, d. h. es gilt:

$$\|T(x_0 + \delta x) - T(x_0)\| \leq 1.$$

Wegen der Linearität von T gilt $T(x_0 + \delta x) - T(x_0) = \delta T(x)$, weshalb wir $T(x) \leq 1/\delta$ bekommen.

(3) $\Longrightarrow$ (4): Für $x \neq 0$ gilt $\left\|\frac{x}{\|x\|}\right\| = 1$. Daraus folgt:

$$\|Tx\| = \left\| \|x\|\, T\left(\frac{x}{\|x\|}\right)\right\| \leq \|T\|\, \|x\|.$$

(4) $\Longrightarrow$ (1): Für $x, x_0 \in X$ gilt:

$$\|Tx - Tx_0\| = \|T(x - x_0)\| \leq C\, \|x - x_0\|.$$

Also ist T Lipschitz-stetig und somit auch stetig.

■

Lemma 3.4.

(1) X, Y normierte Räume $\Longrightarrow$ $L(X,Y)$ normiert mit der Operatornorm.

(2) Y Banachraum $\Longrightarrow$ $L(X,Y)$ Banachraum

(3) X Banachraum $\Longrightarrow$ $L(X) \coloneqq L(X,X)$ Banachalgebra

(4) $T \in L(X,Y), S \in L(Y,Z)$ $\Longrightarrow$ $ST \in L(X,Z)$ mit $\|ST\| \leq \|S\|\,\|T\|$

Beweis. Zu (1): Wir zeigen nur die $\triangle$-Ungleichung (der Rest ist klar). Es gilt

$$\|(T_1 + T_2)(x)\| \leq \|T_1 x\| + \|T_2 x\| \leq (\|T_1\| + \|T_2\|)\,\|x\|,$$

woraus folgt:

$$\|T_1 + T_2\| \leq \|T_1\| + \|T_2\|.$$

Zu (2): Es sei $(T_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $L(X,Y)$. Für alle $x \in X$ ist dann $(T_n x)_{n\in\mathbb{N}}$ eine Cauchy-Folge in Y. Setze

$$Tx \coloneqq \lim_{n\to\infty} T_n x.$$

Da Grenzwertbilden linear ist, ist auch T linear. Wir behaupten, dass $T \in L(X,Y)$ und $\|T_n - T\| \to 0$ für $n \to 0$ gelten.

Zu $\varepsilon \in \mathbb{R}_{>0}$ wähle $n_0 \in \mathbb{N}$, so dass für alle $n, m \in \mathbb{N}_{\geq n_0}$ gilt:

$$\|T_n - T_m\| < \varepsilon.$$

Sei $x \in X$ mit $\|x\| \leq 1$. Wähle $m_0 = m_0(\varepsilon, x) \geq n_0$ mit

$$\|T_{m_0}x - Tx\| \leq \varepsilon.$$

Für alle $n \in \mathbb{N}_{\geq n_0}$ folgt nun:

$$\|T_n x - Tx\| \leq \|T_n x - T_{m_0}x\| + \|T_{m_0}x - Tx\| \leq \|T_n - T_m\| + \varepsilon \leq 2\varepsilon.$$

Damit folgen nun aber $\|T\| \leq \infty$ sowie $\|T_n - T\| \to 0$ für $n \to \infty$.

Zu (3) und (4): Es gilt:

$$\|STx\| \leq \|S\|\,\|Tx\| \leq \|S\|\,\|T\|\,\|x\|.$$

Also gilt allgemein: $\|ST\| \leq \|S\|\,\|T\|$.

■

Bemerkung 3.5. Es sei $T \in L(X, Y)$ und $(T_n)_{n\in\mathbb{N}}$ eine Folge in $L(X, Y)$ mit $T_k x \to Tx$ für $k \to \infty$ und für alle $x \in X$. Dann folgt i. A. *nicht* $T_k \to T$ in $L(X, Y)$.

Beispiel: $X = c_0$ (Raum der Nullfolgen, siehe 2.12 (5)) mit der Supremumsnorm, $Y = \mathbb{R}$, $T_k x := x_k$. Dann gilt: $\lim_{k\to\infty} T_k x = \lim_{k\to\infty} x_k = 0 =: Tx$. Offensichtlich gilt $\|T_k x\| = 1$ für $x = e_k$. Außerdem gilt $\|T_k x\| = \|x_k\| \leq 1$ für $\|x\| \leq 1$. D. h. $\|T_k\| = 1$, aber $\|T\| = 0$.

Definition 3.6. Für $T \in L(X, Y)$ definieren wir den *Nullraum (Kern) von T* als

$$N(T) := \{x \in X \mid T(x) = 0\} = T^{-1}(\{0\}).$$

Es ist $N(T)$ ein abgeschlossener Unterraum von $L(X, Y)$.

Weiter sei

$$R(T) := \{Tx \in Y \mid x \in X\} = T(X)$$

der *Bildraum* (engl.: „range“) von T. Es ist $R(T)$ ein linearer Unterraum von Y, i. A. aber nicht abgeschlossen.

Beispiel: $X = C^0([0, 1])$,

$$T\colon X \to X, \quad (Tf)(x) := \int_0^x f(\xi)\,\mathrm{d}\xi$$
$$R(T) = \{g \in C^1([0, 1]) \mid g(0) = 0\}$$

Es gilt $T \in L(X, Y)$ aber $R(T)$ ist nicht abgeschlossen in X, denn:

$$\overline{R(T)} = \{g \in C^0([0, 1]) \mid g(0) = 0\}$$

(denn stetige Funktionen können durch C^1-Funktionen in der C^0-Norm approximiert werden, siehe später).

Satz 3.7 (Neumannsche Reihe). Sei X ein Banachraum und sei $A \in L(X)$ mit $\|A\| < 1$. Es bezeichne Id den Identitätsoperator. Dann liegt $(\mathrm{Id} - A)^{-1}$ in $L(X)$ und es gilt:

$$(\mathrm{Id} - A)^{-1} = \sum_{n=0}^{\infty} A^n$$

Beweis. Sei für alle $n \in \mathbb{N}$:

$$B_n := \sum_{k=0}^{n} A^k \qquad \in L(X).$$

Mit $\|A^k\| \leq \|A\|^k$ folgt:

$$\begin{aligned}\|B_n x - B_m x\| &= \left\| \sum_{k=m+1}^{n} A^k x \right\| \qquad \text{für } n > m \\ &\leq \sum_{k=m+1}^{n} \|A\|^k \|x\| \to 0 \qquad \text{für } n, m \to \infty \text{ für alle } x \text{ mit gleichmäßig } \|x\| \leq 1\end{aligned}$$

Also existiert $B \in L(X)$ mit $B = \lim_{n\to\infty} B_n$. Noch zu zeigen: $B(\mathrm{Id} - A) = \mathrm{Id} = (\mathrm{Id} - A)B$. Es gilt:

$$\sum_{k=0}^{n} A^k (\mathrm{Id} - A) = \sum_{k=0}^{n} (A^k - A^{k+1}) = \mathrm{Id} - A^{n+1} \to \mathrm{Id} \qquad \text{für } n \to \infty.$$

Also:

$$B(\mathrm{Id} - A) = \lim_{n\to\infty} B_n (\mathrm{Id} - A) = \lim_{n\to\infty} (\mathrm{Id} - A^{n+1}) = \mathrm{Id}$$

■

3.8 (Invertierbarere Operatoren). Seien X, Y Banachräume. Wir sagen $T \in L(X, Y)$ ist invertierbar, falls T bijektiv ist und $T^{-1} \in L(Y, X)$ gilt.

Satz: :

i) Die Teilmenge $\{T \in L(X, Y) \mid T \text{ invertierbar}\}$ ist offen in $L(X, Y)$.

ii) Es gilt genauer für $T, S \in L(X, Y)$ mit invertierbarem T:

$$\|S\| < \left\|T^{-1}\right\|^{-1} \qquad \Longrightarrow \qquad T - S \text{ invertierbar}$$

Beweis. Es gilt:

$$T - S = T\,(\mathrm{Id}_X - \underbrace{T^{-1} S}_{\in L(X)}).$$

Also folgt mit $\|S\| < \left\|T^{-1}\right\|^{-1}$:

$$\left\|T^{-1} S\right\| \leq \left\|T^{-1}\right\| \|S\| < 1.$$

Mithilfe der Neumannschen Reihe (Satz 3.7) erhalten wir:

$$(\operatorname{Id} - T^{-1}S) \text{ ist invertierbar.}$$

Also ist auch $T - S$ invertierbar

$\blacksquare$

4 Der Satz von Hahn-Banach und seine Konsequenzen

Problem: Setzte ein Funktional stetig von einem Unterraum auf den gesamten Raum fort.

Bisher wissen wir nicht, ob auf jedem normierten Vektorraum ein (nicht-triviales) stetiges lineares Funktional existiert. Da wir im Folgenden grundlegend das *Zorn'sche Lemma* verwenden, wiederholen kurz dir Voraussetzungen dafür.

Definition 4.1.

i) Sei M eine Menge. Eine Teilmenge $H \subset M \times M$ definiert eine *Halbordnung* (wir sagen $a \leq b$, falls $(a, b) \in H$ erfüllt ist), wenn für alle $a, b \in M$ gilt:

   a) $a \leq a$

   b) $a \leq b \wedge b \leq a \implies a = b$

   c) $a \leq b \wedge b \leq c \implies a \leq c$

ii) Eine Teilmenge $K \subset M$ heißt *Kette* (oder *total geordnete Teilmenge*), falls für alle $a, b \in K$ entweder $a \leq b$ oder $b \leq a$ gilt.

iii) Eine *obere Schranke* einer Teilmenge $K \subset M$ ist ein Element $s \in M$ mit $a \leq s$ für alle $a \in K$. (Achtung: s muss *nicht* in K liegen!)

iv) Wir sagen M ist *induktiv geordnet*, falls jede Kette in M eine obere Schranke besitzt.

v) Ein $m \in K$ heißt *maximales Element von K*, wenn für alle $a \in K$ aus $a \geq m$ schon $a = m$ folgt.

4.2 (Zorn'sches Lemma)**.** Jede induktiv geordnete Menge besitzt (mindestens) ein maximales Element.

Bemerkung: Das Zorn'sche Lemma ist äquivalent zum Auswahlaxiom.

Satz 4.3 (Satz von Hahn-Banach)**.** Sei X ein $\mathbb{R}$-Vektorraum und $Y \subset X$ ein Unterraum. Weiter gelte:

(1) $p\colon X \to \mathbb{R}$ ist sublinear, d. h. für alle $x, y \in X$ und $\alpha \in \mathbb{R}_{\geq 0}$ gilt:

$$p(x + y) \leq p(x) + p(y) \quad \text{und} \quad p(\alpha x) = \alpha p(x).$$

(2) $f\colon Y \to \mathbb{R}$ ist linear.

(3) $f \leq p$ auf Y.

Dann existiert eine lineare Abbildung $f\colon X \to \mathbb{R}$ mit $f \leq p$ auf X.

Beweis. Nutze das Zorn'sche Lemma (4.2). Es sei

$$M := \big\{(Z, g) \,\big|\, Y \subset Z \subset X,\ Z \text{ ist Unterraum}, \\ g\colon Z \to \mathbb{R} \text{ ist linear},\ g = f \text{ auf } Y,\ g \leq p \text{ auf } Z\,\big\}.$$

Wir definieren eine Halbordnung auf dieser Menge folgendermaßen:

$$(Z_1, g_1) \leq (Z_2, g_2) \quad :\Longleftrightarrow \quad Z_1 \subset Z_2 \wedge g_2|_{Z_1} = g_1.$$

Es ist zunächst zu zeigen, dass es überhaupt ein F gibt, so dass $(X, F) \in M$ gilt. Hierzu brauchen wir folgende Konstruktion:

Sei $(Z, g) \in M, z_0 \in X \setminus Z$. Definiere dann $Z_0 := Z \oplus \operatorname{span}\{z_0\}$. Das Ziel ist es nun, g auf Z_0 fortzusetzen. Ansatz:

$$g_0(z + \alpha z_0) = g(z) + \alpha c$$

für $z \in Z, \alpha \in \mathbb{R}$. Gesucht ist nun ein geeinetes $c \in \mathbb{R}$.

Es muss für alle $z \in Z$ gelten:

$$g(z) + \alpha c \leq p(z + \alpha z_0).$$

Für $\alpha = 0$ ist dies klar. Für $\alpha > 0$ haben wir:

$$c \leq \frac{p(z + \alpha z_0) - g(z)}{\alpha} = p\left(\frac{z}{\alpha} + z_0\right) - g\left(\frac{z}{\alpha}\right)$$

und für $\alpha < 0$:

$$c \geq \frac{p(z + \alpha z_0) - g(z)}{\alpha} = -p\left(-\frac{z}{\alpha} - z_0\right) + g\left(-\frac{z}{\alpha}\right).$$

Gesucht ist nun ein c, so dass

$$\sup_{z' \in Z} \big(g(z') - p(z' - z_0)\big) \leq c \leq \inf_{z \in Z} \left(p(z + z_0) - g(z)\right) \tag{$\star$}$$

erfüllt ist.

Es gilt für alle $z', z \in Z$:

$$g(z + z') \leq p(z + z') = p(z + z_0 + z' - z_0) \leq p(z + z_0) + p(z' - z_0).$$

Daraus folgt für alle $z', z \in Z$:

$$g(z') - p(z' - z_0) \leq p(z + z_0) - g(z),$$

was wiederum bedeutet, dass wir für c einfach den Wert des Supremums in ($\star$) nehmen können. Somit existiert also ein $c \in \mathbb{R}$, so dass $(Z_0, g_0) \in M$ gilt. Sei nun $N \subset M$ eine Kette. Definiere dann

$$Z_0 := \bigcup_{(Z,g)\in N} Z$$

und

$$g_0 \colon Z_0 \to \mathbb{R}, \quad z_0 \mapsto g(z_0) \text{ falls } z_0 \in Z \text{ mit } (Z,g) \in N.$$

Da N eine Kette ist, ist g_0 tatsächlich wohldefiniert. Es folgt also $(Z_0, g_0) \in M$ und für alle $(Z, g) \in N$ gilt $(Z, g) \leq (Z_0, g_0)$.

Das Zorn'sche Lemma liefert nun: Es existiert ein maximales Element $(Z, g) \in M$. Dann muss schon $Z = X$ gelten, denn: Falls $z_0 \in X \setminus Z$ existiert, konstruiere eine Fortsetzung von g auf $Z \oplus \operatorname{span}\{x_0\}$ wie oben. Dies liefert einen Widerspruch zur Maximalität von (Z, g). Damit ist der Satz gezeigt.

■

Wir wollen nun den Satz von Hahn-Banach auf den komplexen Fall verallgemeinern. Die Frage ist, wie wir „$f \leq p$“ auf $\mathbb{C}$ umgehen. Dazu betrachten wir die Realteilfunktion $\operatorname{Re} f$ von f.

Lemma 4.4. Sei X ein $\mathbb{C}$-Vektorraum.

(a) Sei $\ell \colon X \to \mathbb{R}$ ein $\mathbb{R}$-lineares Funktional, d. h.

$$\ell(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 \ell(x_1) + \lambda_2 \ell(x_2)$$

für alle $\lambda_1, \lambda_2 \in \mathbb{R},\ x_1, x_2 \in X$. Setzen wir

$$\tilde{\ell}(x) := \ell(x) - i\,\ell(ix),$$

so ist $\tilde{\ell} \colon X \to \mathbb{C}$ eine $\mathbb{C}$-lineare Abbildung mit $\ell = \operatorname{Re} \tilde{\ell}$.

(b) Ist $h \colon X \to \mathbb{C}$ eine $\mathbb{C}$-lineare Abbildung, $\ell = \operatorname{Re} h$ und $\tilde{\ell}$ wie in (a), so ist ℓ eine $\mathbb{R}$-lineare Abbildung mit $\tilde{\ell} = h$.

(c) Ist $p \colon X \to \mathbb{R}$ eine Halbnorm (es gelten die Normaxiome bis auf $p(x) = 0 \implies x = 0$) und ist $\ell \colon X \to \mathbb{C}$ eine $\mathbb{C}$-lineare Abbildung, so gilt:

$$\Big(\forall x \in X\colon\ |\ell(x)| \leq p(x)\Big) \iff \Big(\forall x \in X\colon\ |\operatorname{Re} \ell(x)| \leq p(x)\Big).$$

(d) Ist X ein normierter Vektorraum und ist $\ell \colon X \to \mathbb{C}$ eine $\mathbb{C}$-lineare Abbildung und stetig, so ist $\|\ell\| = \|\operatorname{Re} \ell\|$.

Bemerkung: $\ell \mapsto \operatorname{Re} \ell$ ist also eine bijektive $\mathbb{R}$-lineare Abbildung zwischen den $\mathbb{C}$-linearen und den $\mathbb{R}$-linearen, $\mathbb{R}$-wertigen Abbildungen. Im normierten Fall ist die Abbildung sogar eine Isometrie.

Beweis.

(a) Da $x \mapsto ix$ eine $\mathbb{R}$-lineare Abbildung ist, folgt: $\tilde{\ell}$ ist $\mathbb{R}$-linear. Die Gleichheit $\operatorname{Re}\tilde{\ell} = \ell$ gilt nach Konstruktion. Außerdem gilt:

$$\tilde{\ell}(ix) = \ell(ix) - i\,\ell(i^2x) = \ell(ix) - i\,\ell(-x) = i\,\big(\ell(x) - i\ell(ix)\big) = i\,\tilde{\ell}(x).$$

(b) Natürlich ist $\ell = \operatorname{Re} h$ eine $\mathbb{R}$-lineare Abbildung. Es gilt für alle $z \in \mathbb{C}$:

$$\begin{aligned} h(x) &= \operatorname{Re} h(x) + i\ \operatorname{Im} h(x) = \operatorname{Re} h(x) - i\ \operatorname{Re}\big(ih(x)\big) \\ &= \operatorname{Re} h(x) - i\ \operatorname{Re} h(ix) = \ell(x) - i\,\ell(ix) = \tilde{\ell}(x) \end{aligned}$$

(c) Wegen $|\operatorname{Re} z| \leq |z|$ für alle $z \in \mathbb{C}$ gilt die Hinrichtung. Die Rückrichtung ergibt sich wie folgt: Schreibe $\ell(x) = \lambda\,|\ell(x)|$ für ein geeignetes $\lambda \in \mathbb{C}$ mit $|\lambda| = 1$. Dann gilt für alle $x \in X$:

$$|\ell(x)| = \lambda^{-1}\ell(x) = \ell(\lambda^{-1}x) = |\operatorname{Re}\ell(\lambda^{-1}x)| \leq p(\lambda^{-1}x) = p(x).$$

(d) folgt sofort aus (c).

∎

Satz 4.5. Sei X ein $\mathbb{C}$-Vektorraum und sei $U \subset X$ ein Unterraum. Weiter sei $p\colon X \to \mathbb{R}$ sublinear und $\ell\colon U \to \mathbb{C}$ linear mit $\operatorname{Re}\ell(x) \leq p(x)$ für alle $x \in U$. Dann existiert eine lineare Fortsetzung $L\colon X \to \mathbb{C}$ mit $L|_U = \ell$ und $\operatorname{Re} L(x) \leq p(x)$ für alle $x \in X$.

Beweis. Wende den Satz von Hahn-Banach (4.3) auf das $\mathbb{R}$-lineare Funktional $\operatorname{Re}\ell\colon U \to \mathbb{R}$ an und erhalte eine $\mathbb{R}$-lineare Abbildung $F\colon X \to \mathbb{R}$ mit $F|_U = \operatorname{Re}\ell$ und $F(x) \leq p(x)$ für alle $x \in X$. Nach Lemma 4.4 ist $F = \operatorname{Re} L$ für ein $\mathbb{C}$-lineares Funktional $L\colon X \to \mathbb{C}$. Dann ist L eine geeignete Fortsetzung.

∎

Satz 4.6. Sei X ein normierter Vektorraum und sei $U \subset X$ ein Unterraum. Zu jedem stetigen linearen Funktional $u'\colon U \to \mathbb{K}$ existiert ein lineares Funktional $x'\colon X \to \mathbb{K}$ mit $x'|_U = u'$ und $\|x'\| = \|u'\|$.

Beweis. Sei X zunächst ein $\mathbb{R}$-Vektorraum. Definiere für alle $x \in X$

$$p(x) \coloneqq \big\|u'\big\|\,\|x\|,$$

womit $p\colon X \to \mathbb{R}$ sublinear ist. Der Satz von Hahn-Banach (Satz 4.3) liefert: Es existiert eine lineare Abbildung $x'\colon X \to \mathbb{R}$ mit $x'|_U = u'$ und $x'(x) \leq p(x)$ für alle $x \in X$. Da auch $x'(-x) \leq p(-x) = p(x)$ gilt, folgt

$$\big|x'(x)\big| \leq \big\|u'\big\|\,\|x\| \quad \text{also} \quad \big\|x'\big\| \leq \big\|u'\big\|.$$

Umgekehrt gilt:

$$\|u'\| = \sup_{\substack{u \in U \\ \|u\| \le 1}} |u'(u)| = \sup_{\substack{u \in U \\ \|u\| \le 1}} |x'(u)| \le \sup_{\substack{x \in X \\ \|x\| \le 1}} |x'(x)| = \|x'\|.$$

Sei X nun ein $\mathbb{C}$-Vektorraum. Wir erhalten wie in Satz 4.5 ein lineares Funktional $x' \colon X \to \mathbb{C}$ mit $x'|_U = u'$ und $\|\operatorname{Re} x'\| = \|u'\|$. Lemma 4.4 (d) liefert dann wie gewünscht $\|\operatorname{Re} x'\| = \|x'\|$. ■

Bemerkung 4.7.

i) Die Fortsetzungen im Satz von Hahn-Banach (Satz 4.3) und seinen Folgerungen sind im Allgemeinen *nicht* eindeutig.

ii) Für Operatoren (lineare Abbildungen von X nach Y) ist die Aussage in Satz 4.6 im Allgemeinen falsch.

    Beispiel: Es gibt keinen stetigen linearen Operator $T \colon \ell^\infty \to c_0$, der die Identität $\mathrm{Id} \colon c_0 \to c_0$ fortsetzt.

iii) Es gibt eine eindeutige stetige Fortsetzung, falls der Unterraum U dicht in X liegt.

Definition 4.8. Eine *affine Hyperebene* in einem $\mathbb{K}$-Vektorraum X ist eine Teilmenge $H \subset X$ der Form

$$H = \{x \in X \mid f(x) = \alpha\}$$

für eine (nicht-trivale) lineare Abbildung $f \colon X \to \mathbb{K}$ und $\alpha \in \mathbb{K}$. Wir schreiben auch kurz: $H = \{f = \alpha\}$.

Satz 4.9. Sei X ein $\mathbb{R}$-Vektorraum, sei $f \colon X \to \mathbb{R}$ linear und sei $\alpha \in \mathbb{R}$. Die Hyperebene $H = \{f = \alpha\}$ ist genau dann abgeschlossen, wenn f stetig ist.

Beweis. Es ist klar, dass H abgeschlossen ist, wenn f stetig ist, denn es gilt $f^{-1}(\{\alpha\}) = H$ und $\{\alpha\}$ ist abgeschlossen in $\mathbb{R}$. Für die Rückrichtung sei H abgeschlossen in X. Dann ist H^c offen und nicht leer. Jetzt sei $x_0 \in H^\mathsf{c}$ mit $f(x_0) \neq \alpha$, o. E. $f(x_0) < \alpha$. Sei $r \in \mathbb{R}_{>0}$, so dass $B_r(x_0) = \{x \in X \mid \|x - x_0\| < r\} \subset H^\mathsf{c}$. (Abbildung 4.2)

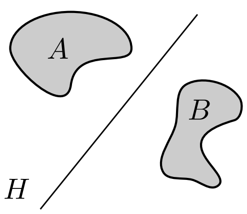


Abbildung 4.1: Zwei Teilmengen A und B eines Vektorraums, getrennt durch eine Hyperebene H (hier eine Gerade im $\mathbb{R}^2$)

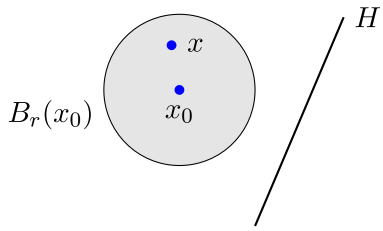


Abbildung 4.2: Hyperebene H und Ball $B_r(x_0)$ um x_0 mit $x \in B_r(x_0)$

Wir behaupten nun, dass dann schon für alle $x \in B_r(x_0)$ die Ungleichung

$$f(x) < \alpha \tag{$*$}$$

gilt. Angenommen dies gilt nicht und es existiert ein $x_1 \in B_r(x_0)$, so dass $f(x_1) > \alpha$ gilt. Das Segment

$$[x_0, x_1] \coloneqq \{x_t \coloneqq (1-t)\, x_0 + t x_1 \mid t \in [0,1]\}$$

ist in $B_r(x_0)$ enthalten (da Bälle in normierten Räumen konvex sind). Somit folgt für alle $t \in [0,1]$:

$$f(x_t) \neq \alpha.$$

Andererseits gilt offenbar

$$f(x_t) = \alpha \quad \text{für} \quad t = \frac{\alpha - f(x_0)}{f(x_1) - f(x_0)}.$$

Dies ist ein Widerspruch, also muss doch schon $(*)$ gelten. Wir erhalten, dass für alle $z \in B_1(0)$

$$f(\underbrace{x_0 + rz}_{\in B_r(x_0)}) < \alpha$$

gilt, woraus sofort

$$f(z) < \frac{1}{r}\big(\alpha - f(x_0)\big)$$

folgt. Nutze diese Ungleichung für z und $-z$ aus $B_1(0)$, um Folgendes für alle $z \in B_1(0)$ zu erhalten:

$$|f(z)| < \frac{1}{r}\big(\alpha - f(x_0)\big).$$

Insgesamt folgt:

$$\|f\| = \sup_{z \in B_1(0)} |f(z)| \leq \frac{1}{r}\big(\alpha - f(x_0)\big).$$

∎

Definition 4.10. Sei X ein $\mathbb{R}$-Vektorraum und seien $A, B \subset X$ zwei Teilmengen von X. Die Hyperebene $H = \{f = \alpha\}$ *trennt die Mengen A und B*, falls für alle $a \in A$ und alle $b \in B$ die Ungleichungen

$$f(a) \leq \alpha \qquad \text{und} \qquad f(b) \geq \alpha$$

gelten.

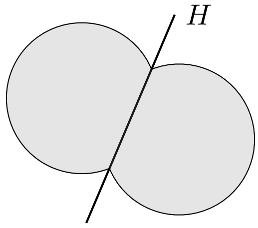


Abbildung 4.3: Zwei *nicht* strikt durch H getrennte Mengen; die Mengen aus Abbildung 4.1 sind hingegen strikt durch H getrennt

Die Hyperebene H *trennt* A *und* B *strikt*, falls es ein $\varepsilon \in \mathbb{R}_{>0}$ gibt, so dass für alle $a \in A$ und alle $b \in B$ die Ungleichungen

$$f(a) \leq \alpha - \varepsilon \qquad \text{und} \qquad f(b) \geq \alpha + \varepsilon$$

gelten.

Bemerkung 4.11.

i) Geometrisch sagt solch eine Trennung aus, dass A auf der einen Seite von H liegt und B auf der anderen Seite. (Siehe Abbildung 4.1 und Abbildung 4.3.)

ii) Ist X ein $\mathbb{C}$-Vektorraum, so sagen wir, dass A und B durch eine *reelle Hyperebene* getrennt werden, falls $f\colon X \to \mathbb{C}$ linear und $\alpha \in \mathbb{R}$ existieren, so dass für alle $a \in A$ und alle $b \in B$ die Ungleichungen

$$\operatorname{Re} f(x) \leq \alpha \qquad \text{und} \qquad \operatorname{Re} f(b) \geq \alpha$$

gelten.

iii) Wir nennen $A \subset X$ konvex, falls für alle $x, y \in A$ auch

$$[x, y] = \{(1-t)\, x + ty \mid t \in [0,1]\} \subset A$$

gilt.

Definition: Sei X ein $\mathbb{K}$-Vektorraum und $K \subset X$. Dann ist das *Minkowski-Funktional zu* K definiert durch

$$p(x) := \inf \left\{ \alpha \in \mathbb{R}_{>0} \;\middle|\; \frac{1}{\alpha}\, x \in K \right\}.$$

Für $K = B_1(0)$ gilt gerade $p(x) = \|x\|$.

Lemma 4.12. Es sei K konvex, offen und $0 \in K$. Dann gilt:

i) Das Minkowski-Funktional p zu K ist sublinear.

ii) Es existiert ein $M \in \mathbb{R}_{>0}$, so dass für alle $x \in X$ gilt:

$$0 \le p(x) \le M \, \|x\|.$$

iii) Zwischen K und p besteht folgender Zusammenhang:

$$K = \{x \in X \mid p(x) < 1\}.$$

Beweis.

i) Seien $\lambda \in \mathbb{R}_{>0}$ und $x \in X$. Dann gilt $p(\lambda x) = \lambda \, p(x)$, denn:

$$\begin{aligned} p(\lambda x) &= \inf \left\{ \alpha \in \mathbb{R}_{>0} \;\middle|\; \frac{1}{\alpha} \lambda x \in K \right\} = \inf \left\{ \alpha' \lambda \;\middle|\; \frac{1}{\alpha' \lambda} \lambda x \in K \right\} \\ &= \lambda \inf \left\{ \alpha' \in \mathbb{R}_{>0} \;\middle|\; \frac{1}{\alpha'} x \in K \right\} = \lambda \, p(x) \end{aligned}$$

Die $\triangle$-Ungleichung zeigen wir später.

ii) Es sei $r \in \mathbb{R}_{>0}$ derart, dass $B_r(0) \subset K$ gilt. Es gilt dann für alle $x \in X$:

$$p(x) \le \frac{1}{r} \, \|x\|,$$

denn:

$$\begin{aligned} p(x) &= \inf \left\{ \alpha \in \mathbb{R}_{>0} \;\middle|\; \frac{1}{\alpha} x \in K \right\} \\ &\le \inf \left\{ \alpha \in \mathbb{R}_{>0} \;\middle|\; \frac{1}{\alpha} x \in B_r(0) \right\} \\ &= \frac{1}{r} \, \|x\| \end{aligned}$$

iii) Es sei $x \in K$. Da K offen ist, folgt $(1 + \varepsilon) \, x \in K$ für $\varepsilon \in \mathbb{R}_{>0}$ klein genug. Dann gilt:

$$p(x) \le \frac{1}{1 + \varepsilon} < 1.$$

Falls $p(x) < 1$ gilt, muss ein $\alpha \in (0, 1)$ geben, so dass $x/\alpha \in K$ erfüllt ist. Damit gilt:

$$x = \alpha \left(\frac{x}{\alpha} \right) + (1 - \alpha) \cdot 0 \in K,$$

denn K ist nach Voraussetzung konvex.

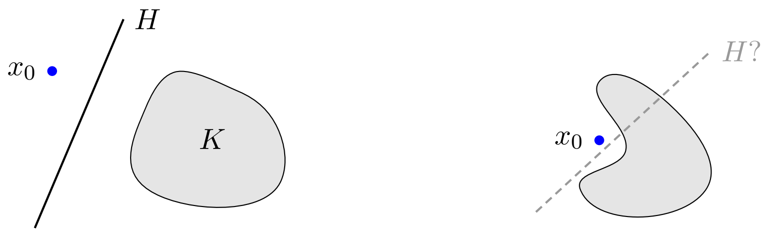


Abbildung 4.4: Links eine konvexe Menge K, getrennt von $\{x_0\}$ durch H; rechts eine nicht konvexe Menge, so dass diese und x_0 nicht durch eine Hyperebene H getrennt werden können

i) Es bleibt die $\triangle$-Ungleichung zu zeigen. Seien $x, y \in X$ und $\varepsilon \in \mathbb{R}_{>0}$. Aus dem bisher Gezeigten folgt:

$$\frac{x}{p(x)+\varepsilon}, \frac{y}{p(y)+\varepsilon} \in K.$$

Damit gilt also für alle $t \in [0,1]$:

$$\frac{tx}{p(x)+\varepsilon} + \frac{(1-t)\,y}{p(y)+\varepsilon} \in K.$$

Wähle nun $t = \frac{p(x)+\varepsilon}{p(x)+p(y)+2\varepsilon}$, dann erhalten wir

$$\frac{x+y}{p(x)+p(y)+2\varepsilon} \in K \quad \text{und mit (iii) folgt} \quad p\left(\frac{x+y}{p(x)+p(y)+2\varepsilon}\right) < 1.$$

Es folgt:

$$p(x+y) < p(x)+p(y)+2\varepsilon.$$

Da $\varepsilon \in \mathbb{R}_{>0}$ beliebig war, folgt die Behauptung.

■

Lemma 4.13. Sei X ein $\mathbb{K}$-Vektorraum und sei $K \subset X$ nicht-leer, offen und konvex. Sei weiter $x_0 \in K^{\mathsf{c}}$. Dann existiert ein $x' \in X'$, so dass für alle $x \in K$ gilt:

$$\operatorname{Re} x'(x) < \operatorname{Re} x'(x_0).$$

Insbesondere trennt im Fall $\mathbb{K} = \mathbb{C}$ die reelle Hyperebene $\{\operatorname{Re} x' = \operatorname{Re} x'(x_0)\}$ somit $\{x_0\}$ und K.

Beweis. Sei zunächst $\mathbb{K} = \mathbb{R}$. Ohne Einschränkung können wir $0 \in K$ annehmen. Sei p das Minkowski-Funktional zu K. Sei weiter $U := \operatorname{span}\{x_0\}$ und $g\colon U \to \mathbb{R}$ sei gegeben durch $g(tx_0) := t$ für alle $t \in \mathbb{R}$. Dann gilt für alle $x \in U$:

$$g(x) \leq p(x)$$

und für x_0 haben wir $g(x_0) = 1 \leq p(x_0)$, da x_0 nicht in K liegt. (Achtung: tx_0 mit $t < 0$ ist kein Problem, da $g(tx_0) < 0$.)

Wende nun Hahn-Banach (Satz 4.3) an, womit wir ein $x'\colon X \to \mathbb{R}$ erhalten, mit $x'(x) \leq p(x)$ für alle $x \in X$ und außerdem $x'|_U = g$. Insbesondere gilt also $x'(x_0) = 1$. Außerdem ist x' stetig (vgl. Lemma 4.12 (ii)). Mit Lemma 4.12 (iii) erhalten wir: für alle $x \in K$ gilt

$$x'(x) < 1.$$

Der komplexe Fall (also $\mathbb{K} = \mathbb{C}$) folgt aus dem Obigen und Lemma 4.4.

∎

Satz 4.14 (Satz von Hahn-Banach (erste geometrische Formulierung))**.** Sei X ein normierter $\mathbb{K}$-Vektorraum und seien $A, B \subset X$ nicht-leer, konvex und disjunkt. Außerdem sei A offen. Dann existiert $x' \in X'$ mit $\operatorname{Re} x'(a) < \operatorname{Re} x'(b)$ für alle $a \in A$ und $b \in B$.

Bemerkung: Ist X ein $\mathbb{R}$-Vektorraum, so trennt die abgeschlossene Hyperebene $\{x' = \alpha\}$ mit

$$\alpha \in \big[\sup_{a\in A} x'(a), \inf_{b \in B} x'(b)\big]$$

die Mengen A und B.

Beweis. Es sei $C \coloneqq A - B \coloneqq \{a - b \mid a \in A,\, b \in B\}$. Dann ist C konvex (leichte Rechnung) und offen, denn:

$$C = \bigcup_{b\in B} \underbrace{(A - \{b\})}_{\text{offen}}.$$

Da A und B disjunkt sind, liegt 0 nicht in C. Aus Lemma 4.13 folgt die Existenz eines $x' \in X'$, welches für alle $x \in C$ die Ungleichung

$$\operatorname{Re} x'(x) < 0 = \operatorname{Re} x'(0)$$

erfüllt. Das heißt, es gilt für alle $a \in A$ und $b \in b$

$$\operatorname{Re} x'(a - b) < 0 \quad \text{oder äquivalent} \quad \operatorname{Re} x'(a) < \operatorname{Re} x'(b).$$

∎

Satz 4.15 (Satz von Hahn-Banach (zweite geometrische Formulierung))**.** Sei X ein normierter $\mathbb{K}$-Vektorraum und seien $A, B \subset X$ nicht-leer, konvex und disjunkt. Weiter sei A abgeschlossen und B kompakt. Dann exisistiert ein $x' \in X'$ sowie ein $\alpha \in \mathbb{R}$ und ein $\varepsilon \in \mathbb{R}_{>0}$, so dass für alle $a \in A$ und $b \in B$ die Ungleichungen

$$\operatorname{Re} x'(a) + \varepsilon \leq \alpha \leq \operatorname{Re} x'(b) - \varepsilon$$

gelten.

Beweis. Es sei $C \coloneqq A - B$ wie bei Satz 4.14. Damit ist C konvex und abgeschlossen (siehe unten) und es gilt $0 \notin C$. Damit existiert ein $r \in \mathbb{R}_{>0}$, so dass $B_r(0) \cap C = \emptyset$ gilt. Satz 4.14 liefert: Es existiert ein $x' \in X'$ mit $x' \not\equiv 0$, so dass für alle $a \in A$, $b \in B$ und $z \in B_1(0)$ gilt:

$$\operatorname{Re} x'(a - b) < \operatorname{Re} x'(rz).$$

Also gilt für alle $a \in A$ und $b \in B$:

$$\operatorname{Re} x'(a-b) \leq -r\,\|x'\|.$$

Für $\varepsilon r \|x'\|/2 > 0$ ergibt sich, dass für alle $a \in A$ und alle $b \in B$ gilt:

$$\operatorname{Re} x'(a) + \varepsilon \leq \operatorname{Re} x'(b) - \varepsilon.$$

Wähle $\alpha \in \mathbb{R}_{>0}$, so dass

$$\sup_{a\in A} \big(x'(a) + \varepsilon\big) \leq \alpha \leq \sup_{b \in B} \big(x'(b) - \varepsilon\big)$$

erfüllt ist. Noch zu zeigen: C ist abgeschlossen. Sei $(c_n)_{n\in\mathbb{N}} = (a_n - b_n)_{n\in\mathbb{N}}$ eine Folge in C mit Grenzwert $c \in X$. Da B kompakt ist, existiert eine Teilfolge $(b_{n_k})_{k\in\mathbb{N}}$ von $(b_n)_{n\in\mathbb{N}}$, mit $b_{n_k} \to b \in B$ für $k \to \infty$. Damit ergibt sich:

$$a_{n_k} = c_{n_k} + b_{n_k} \xrightarrow[n\to\infty]{} c + b.$$

Da A abgeschlossen ist, folgt $c + b \in A$ und damit $c = (c+b) - b \in A - B = C$.

■

Im Allgemeinen lassen sich konvexe Mengen mit $A \cap B = \emptyset$ nicht trennen. Es gibt Beispiele mit A, B zusätzlich abgeschlossen, in denen Trennungen nicht möglich ist. (Siehe Übungen.)

Korollar 4.16. Sei X ein normierter Vektorraum und $U \subset X$ ein Unterraum mit $\overline{U} \neq X$. Dann existiert ein $x' \in X'$ mit $x' \neq 0$ und $x'|_U = 0$.

Beweis. Es sei $x_0 \in X$ mit $x_0 \notin \overline{U}$. Wende Satz 4.15 auf $A = \overline{U}$ und $B = \{x_0\}$ an. Wir erhalten somit ein $x' \in X'$ und ein $\alpha \in \mathbb{R}$ mit $\operatorname{Re} x'(x) < \alpha < \operatorname{Re} x'(x_0)$ für alle $x \in \overline{U}$. Es folgt für alle $\lambda \in \mathbb{R},\ x \in \overline{U}$:

$$\operatorname{Re} x'(\lambda x) < \alpha.$$

Also muss schon $\operatorname{Re} x'(x) = 0$ für alle $x \in \overline{U}$ gelten. Wegen $\operatorname{Re} x'(x_0) > \operatorname{Re} x'(x) = 0$ für alle $x \in U$ ist außerdem $x' \neq 0$.

■

Bemerkung: Korollar 4.16 wird genutzt, um zu zeigen, dass ein Unterraum U dicht in einem umgebenden Raum X liegt. Kann man zeigen, dass für alle $x' \in X'$ aus $x'|_U = 0$ schon $x' = 0$ folgt, so ergibt sich $\overline{U} = X$.

Definition 4.17. Sei $(X, \|\cdot\|)$ ein normierter $\mathbb{K}$-Vektorraum und X' der Dualraum zu X (Definition 3.1 (b)). Dann ist $X'' := (X')'$ der *Bidualraum von X*.

Wir können auf kanonische Weise eine Abbildung $J_X \colon X \to X''$ wie folgt definieren:

$$x \mapsto \begin{pmatrix} X' \to \mathbb{K} \\ x' \mapsto x'(x) \end{pmatrix}.$$

Dann ist J_X linear und stetig, denn es gilt für alle $x' \in X'$ und alle $x \in X$ die Ungleichung $|x'(x)| \le \|x'\| \cdot \|x\|$ und damit für alle $x \in X$:

$$\|J_X(x)\| \le \|x\|. \tag{$*$}$$

Sei

$$\bar{B}_1^{X'} := \overline{B_1^{X'}(0)} = \big\{x' \in X' \;\big|\; \|x'\| \le 1\big\}.$$

Dann gilt sogar

$$\|x\| = \sup_{x' \in \bar{B}_1^{X'}} |x'(x)| \quad \text{für alle } x \in X.$$

Sei $x_0 \in X \setminus \{0\}$. Setze dann das Funktional

$$u' \colon \operatorname{span}\{x_0\} \to \mathbb{K}, \qquad x \mapsto \lambda \|x_0\| \quad \text{falls } \lambda \in \mathbb{K} \text{ mit } x = \lambda x_0$$

normgleich auf X fort. Es gilt dann $\|x'\| = \|u'\| = 1$ und $x'(x_0) = \|x_0\|$. Damit ist in $(*)$ sogar Gleichheit gezeigt. Insgesamt folgt:

Satz 4.18. Die Abbildung J_X ist eine (im Allgemeinen nicht surjektive) lineare Isometrie, d. h. für alle $x \in X$ gilt $\|J_X(x)\|_{X''} = \|x\|_X$. (Insbesondere ist J_X als Isometrie stets injektiv.)

Definition 4.19. Ein Banachraum X ist *reflexiv*, wenn J_X surjektiv (also bijektiv) ist.

Bemerkung: Da J_X injektiv ist, kann X mit einem Unterraum von X'' identifiziert werden.

Definition 4.20. Sei X ein normierter Raum, $M \subset X$ ein Unterraum und $N \subset X'$ ein Unterraum des Dualraums. Wir definieren dann den *Annihilator von M* als

$$\begin{aligned} M^\perp &:= \big\{x' \in X' \;\big|\; \forall\, x \in M \colon\; x'(x) = 0\big\} \\ &= \big\{x' \in X' \;\big|\; x'|_M = 0\big\} \end{aligned}$$

und den *Annihilator von N* als

$$N^\perp := \big\{x \in X \;\big|\; \forall\, x' \in N \colon\; x'(x) = 0\big\}.$$

Bemerkung 4.21.

(i) Es ist $N^\perp$ eine Teilmenge von X und *nicht* von X''.

(ii) Es sind $M^\perp$ und $N^\perp$ abgeschlossene Unterräume.

Satz 4.22. Sei X ein normierter Raum und $M \subset X$ ein Unterraum. Dann gilt

$$\big(M^\perp\big)^\perp = \overline{M}.$$

Sei außerdem $N \subset X'$ ein Unterraum. Dann gilt

$$\big(N^\perp\big)^\perp \supset \overline{N}.$$

(Im Allgemeinen ist diese Inklusion echt.)

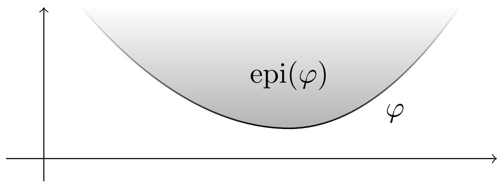


Abbildung 4.5: Epigraph einer Funktion φ

Definition 4.23.

i) Es sei E eine Menge und $\varphi\colon E \to (-\infty, \infty] = \mathbb{R} \cup \{\infty\}$ eine Abbildung. Wir definieren dann
$$D(\varphi) \coloneqq \{x \in E \mid \varphi(x) < \infty\} = \varphi^{-1}(\mathbb{R}).$$

ii) Der *Epigraph von* φ (Abbildung 4.5) ist die Menge
$$\operatorname{epi}(\varphi) \coloneqq \{(x, \lambda) \in E \times \mathbb{R} \mid \varphi(x) \leq \lambda\}.$$

Definition 4.24. Sei $(E, \mathcal{T})$ ein topologischer Raum. Eine Abbildung $\varphi\colon E \to (-\infty, \infty]$ ist *unterhalbstetig*, wenn für alle $\lambda \in \mathbb{R}$ die Menge
$$\{\varphi \leq \lambda\} \coloneqq \{x \in E \mid \varphi(x) \leq \lambda\} = \varphi^{-1}(\mathbb{R}_{\leq \lambda}) \subset E$$
abgeschlossen ist.

Lemma 4.25. Sei $(E, \mathcal{T})$ ein topologischer Raum und $\varphi\colon E \to (-\infty, \infty]$ eine Abbildung. Dann gelten folgende Aussagen:

(i) Es ist φ genau dann unterhalbstetig, wenn $\operatorname{epi}(\varphi)$ abgeschlossen in $E \times \mathbb{R}$ (mit der Produkttopologie) ist.

(ii) Es ist φ genau dann unterhalbstetig, wenn für alle $x \in E$ und alle $\varepsilon \in \mathbb{R}_{>0}$ eine Umgebung V von x existiert, so dass für alle $y \in V$ gilt: $\varphi(y) \geq \varphi(x) \cdot \varepsilon$.

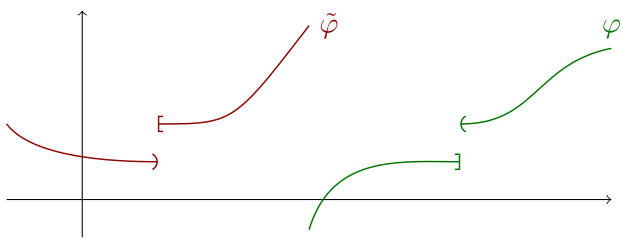


Abbildung 4.6: Die Funktion $\tilde{\varphi}$ ist *nicht* unterhalbstetig, φ schon

(iii) Ist φ unterhalbstetig, so gilt für jede Folge $(x_n)_{n\in\mathbb{N}}$ in E mit $\lim_{n\to\infty} x_n = x \in E$:

$$\liminf_{n\to\infty} \varphi(x_n) \geq \varphi(x).$$

Falls E ein metrischer Raum ist, so gilt auch die Umkehrung.

(iv) Sind φ und $\tilde{\varphi}\colon E \to (-\infty, \infty]$ unterhalbstetig, so auch $\varphi + \tilde{\varphi}$.

(v) Ist $(\varphi_i)_{i\in I}$ eine Familie unterhalbstetiger Abbildungen $E \to (-\infty, \infty]$ mit

$$\varphi(x) = \sup_{i\in I} \varphi_i(x)$$

für alle $x \in E$, so ist auch φ unterhalbstetig.

(vi) Ist $E \neq \emptyset$ folgenkompakt und φ unterhalbstetig, so nimmt die Funktion φ ihr Minimum an, d. h. es existiert ein $x_0 \in E$ mit $\varphi(x_0) = \inf_{x\in E} \varphi(x)$.

Beweis. Siehe Übungen für Teile der Aussagen. Wir beweisen hier nur (vi):

Sei also E folgenkompakt und nicht leer und sei φ unterhalbstetig. Sei dann $(x_n)_{n\in\mathbb{N}}$ eine Folge in E, für welche $\big(\varphi(x_n)\big)_{n\in\mathbb{N}}$ gegen $\inf_{x\in E} \varphi(x)$ konvergiert. Weil E folgenkompakt ist, existiert dann eine konvergente Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ und wir definieren $x_0 := \lim_{k\to\infty} x_{n_k}$. Aus (iii) folgt dann:

$$\varphi(x_0) \leq \inf_{x\in E} \varphi(x) \leq \varphi(x_0).$$

(Dies zeigt auch, dass $\inf_{x\in E} \varphi(x) > -\infty$ gelten muss.)

■

Definition 4.26. Sei X ein Vektorraum. Eine Funktion $\varphi\colon X \to (-\infty, \infty]$ ist *konvex*, wenn φ für alle $x, y \in X$ und alle $t \in [0, 1]$ die Ungleichung

$$\varphi\big(tx + (1-t)\,y\big) \;\leq\; t\,\varphi(x) + (1-t)\,\varphi(y)$$

erfüllt. (Abbildung 4.7)

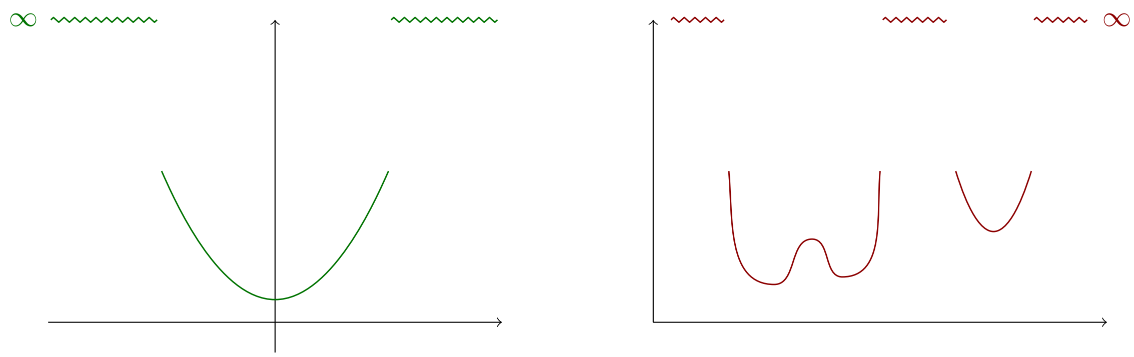


Abbildung 4.7: Konvexe Funktion links und *nicht* konvexe Funktion rechts

Lemma 4.27. Sei X ein Vektorraum.

(i) Es ist $\varphi\colon X \to (-\infty, \infty]$ genau dann konvex, wenn $\operatorname{epi}(\varphi)$ eine konvexe Teilmenge von $X \times \mathbb{R}$ ist.

(ii) Ist $\varphi\colon X \to (-\infty, \infty]$ konvex, so ist die Menge $\{\varphi \leq \lambda\}$ für alle $\lambda \in \mathbb{R}$ konvex. (Die Umkehrung gilt im Allgemeinen nicht.)

(iii) Sind $\varphi_1, \varphi_2\colon X \to (-\infty, \infty]$ konvex, so auch $\varphi_1 + \varphi_2$.

(iv) Ist $(\varphi_i)_{i \in I}$ eine Familie konvexer Abbildungen $X \to (-\infty, \infty]$, so ist auch

$$\sup_{i \in I} \varphi_i \coloneqq \big(x \mapsto \sup_{i \in I} \varphi_i(x)\big)$$

konvex.

Ab jetzt betrachten wir vornehmlich normierte $\mathbb{R}$-Vektorräume.

Definition 4.28. Sei X ein normierter $\mathbb{R}$-Vektorraum und sei $\varphi\colon X \to (-\infty, \infty]$ eine Funktion mit $D(\varphi) \neq \emptyset$ (d. h. φ ist nicht konstant ∞). Dann ist die *Legendre-Transformation* (oder *konjugierte Funktion*) von φ die Abbildung

$$\begin{aligned} \varphi^*\colon X' &\to (-\infty, \infty] \\ f &\mapsto \sup_{x \in X} \big(f(x) - \varphi(x)\big). \end{aligned}$$

Bemerkung 4.29.

(i) Sei $n \in \mathbb{N}$ und $\varphi\colon \mathbb{R}^n \to (-\infty, \infty]$ eine Abbildung. Ist $f \in (\mathbb{R}^n)'$, so gibt es genau einen Vektor $y_f \in \mathbb{R}^n$ mit $f(x) = x \cdot y \coloneqq \langle x, y \rangle_{\text{eukl}}$ für alle $x \in \mathbb{R}^n$. Wir identifizieren dann $(\mathbb{R}^n)'$ mit $\mathbb{R}^n$ vermöge

$$\begin{aligned} (\mathbb{R}^n)' &\longleftrightarrow \mathbb{R}^n \\ f &\longmapsto y_f \\ (x \mapsto x \cdot y) &\longleftarrow y\,, \end{aligned}$$

und somit gilt für alle $y \in \mathbb{R}^n$:

$$\varphi^*(y) = \sup_{x \in \mathbb{R}^n} \big(x \cdot y - \varphi(x)\big).$$

(ii) Es ist φ^* stets konvex und unterhalbstetig. Dies folgt daraus, dass wir das Supremum betrachten und dass $f \mapsto f(x) - \varphi(x)$ konvex und stetig ist (da affin linear).

(iii) Es gilt für alle $x \in X$ und alle $f \in X'$ die Ungleichung

$$f(x) \leq \varphi(x) + \varphi^*(f),$$

was direkt aus der Definiton von φ^* folgt.

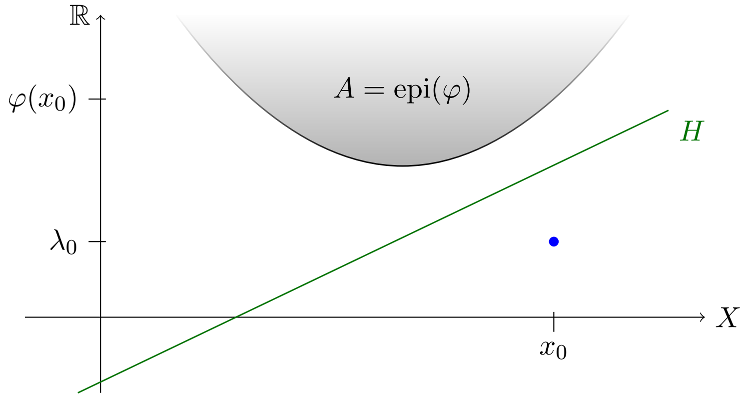


Abbildung 4.8: Skizze zum Beweis von Theorem 4.30

(iv) Die Youngsche Ungleichung (Lemma 2.13) ist ein Spezialfall von (iii). Seien $p, p' \in (1, \infty)$ mit $\frac{1}{p} + \frac{1}{p'} = 1$ und setze $\varphi(x) := \frac{1}{p} |x|^p$. Dann gilt für alle $y \in \mathbb{R}_{\geq 0}$:

$$\varphi^*(y) = \frac{1}{p'} |y|^{p'}.$$

(Siehe Übungen.)

Theorem 4.30. Sei X ein normierter $\mathbb{R}$-Vektorraum und $\varphi\colon X \to (-\infty, \infty]$ konvex und unterhalbstetig mit $D(\varphi) \neq \emptyset$. Dann gilt $D(\varphi^*) \neq \emptyset$ und φ ist von unten durch eine affin lineare Funktion beschränkt.

Beweis. Sei $x_0 \in D(\varphi)$ und sei $\lambda_0 \in \mathbb{R}$ mit $\lambda_0 < \varphi(x_0)$. Wende nun die zweite geometrische Form des Satzes von Hahn-Banach (4.15) auf den Raum $X \times \mathbb{R}$, die abgeschlossene Menge $A := \mathrm{epi}(\varphi)$ und die kompakte Menge $B := \{(x_0, \lambda_0)\}$ an. (Abbildung 4.8) Wir erhalten somit ein stetiges lineares Funktional $\Phi\colon X \times \mathbb{R} \to \mathbb{R}$ und ein $\alpha \in \mathbb{R}$, so dass die abgeschlossene Hyperebene $H = \{\Phi = \alpha\} \subset X \times \mathbb{R}$ die Mengen A und B trennt. Die Abbildung

$$\begin{aligned} f\colon X &\to \mathbb{R} \\ x &\mapsto \Phi\big((x, 0)\big) \end{aligned}$$

ist stetig und es gilt $f \in X'$. Mit $k := \Phi\big((0, 1)\big)$ gilt für alle $(x, \lambda) \in X \times \mathbb{R}$

$$\Phi\big((x, \lambda)\big) = f(x) + k\lambda.$$

Es gilt weiter $\Phi|_A > \alpha$ und $\Phi|_B < \alpha$. Dann gilt also $f(x_0) + k\lambda_0 < \alpha$ und

$$f(x) + k\lambda > \alpha$$

für alle $(x, \lambda) \in \mathrm{epi}(\varphi)$. Somit erhalten wir für alle $x \in D(\varphi)$:

$$f(x) + k\varphi(x) > \alpha \tag{$\star$}$$

und für den Punkt (x_0, λ_0):

$$f(x_0) + k\varphi(x_0) > \alpha > f(x_0) + k\lambda_0.$$

Dies zeigt $k > 0$ (da $\varphi(x_0) > \lambda_0$ nach Wahl von λ_0). Aus $(\star)$ folgt, dass für alle $x \in D(\varphi)$ gilt:

$$-\frac{1}{k} f(x) - \varphi(x) < -\frac{\alpha}{k}.$$

Daraus folgt $\varphi^*\big(-\frac{1}{k}f\big) < \infty$ (nach Definition von φ^*) und damit

$$\varphi(x) > -\frac{1}{k} f(x) + \frac{\alpha}{k},$$

aber gerade das wollten wir zeigen.

$\blacksquare$

Wir können auch die Funktion φ^{**} betrachten. Dies wäre eigentlich eine Abbildung von X'' nach $\mathbb{R}$. Wir schränken diese aber auf X ein (unter der Einbettung von X nach X'' vermöge der Isometrie J_X, siehe Definition 4.17 ff.).

Definition 4.31. Es sei $\varphi\colon X \to (-\infty, \infty]$ mit $D(\varphi) \neq \emptyset$. Wir definieren $\varphi^{**}\colon X \to \mathbb{R}$ für alle $x \in X$ durch

$$\varphi^{**}(x) \coloneqq \sup_{f \in X'} \big(f(x) - \varphi^*(f)\big).$$

Theorem 4.32 (Fenchel-Moreau)**.** Sei $\varphi\colon X \to (-\infty, \infty]$ konvex, unterhalbstetig und $D(\varphi) \neq \emptyset$. Dann gilt $\varphi^{**} = \varphi$.

Beweis. Schritt 1: Wir setzen $\varphi \geq 0$ voraus. Da für alle $x \in X$ und alle $f \in X'$

$$f(x) - \varphi^*(f) \leq \varphi(x)$$

gilt, erhalten wir zunächst $\varphi^{**} \leq \varphi$. Angenommen es existiert ein $x_0 \in X$ mit

$$\varphi^{**}(x_0) < \varphi(x_0)$$

(wobei $\varphi(x_0) = \infty$ möglich ist). Nutze wieder den Satz von Hahn-Banach in der zweiten geometrischen Formulierung (4.15) mit $A = \operatorname{epi}(\varphi)$ abgeschlossen und $B = \{(x_0, \varphi^{**}(x_0)\}$ kompakt. (Vgl. Beweis von Theorem 4.30.) Wir erhalten somit ein $f \in X'$ und $k, \alpha \in \mathbb{R}$ mit $f(x_0) + k\varphi^{**}(x_0) < \alpha$ und

$$f(x) + k\lambda > \alpha. \tag{$\diamond$}$$

für alle $(x, \lambda) \in \operatorname{epi}(\varphi)$. Wähle $x \in D(\varphi)$ und betrachte $\lambda \to \infty$ in der letzten Ungleichung. Es folgt $k \geq 0$. Jetzt sei $\varepsilon \in \mathbb{R}_{>0}$. Da $\varphi \geq 0$ gilt, folgt aus $(\diamond)$, dass für alle $x \in D(\varphi)$ gilt:

$$f(x) + (k + \varepsilon)\, \varphi(x) \geq \alpha.$$

Somit erhalten wir für alle $x \in D(\varphi)$:

$$-\frac{1}{k+\varepsilon}\, f(x) - \varphi(x) \leq -\frac{\alpha}{k+\varepsilon}.$$

Dies zeigt:

$$\varphi^*\left(-\frac{f}{k+\varepsilon}\right) \leq -\frac{\alpha}{k+\varepsilon}.$$

Die Definition von φ^{**} liefert für $\varphi^{**}(x_0)$:

$$\varphi^{**}(x_0) \geq -\frac{f}{k+\varepsilon}(x_0) - \varphi^*\left(-\frac{f}{k+\varepsilon}\right) \geq -\frac{f}{k+\varepsilon}(x_0) + \frac{\alpha}{k+\varepsilon}.$$

Damit folgt

$$f(x_0) + (k+\varepsilon)\,\varphi^{**}(x_0) \geq \alpha,$$

was aber für $\varepsilon \to 0$ einen Widerspruch zu $f(x_0) + k\varphi^{**}(x_0) < \alpha$ liefert.

Schritt 2 (allgemeiner Fall): Theorem 4.30 sichert uns $D(\varphi^*) \neq \emptyset$. Wähle dann $f_0 \in D(\varphi^*)$ und setze für alle $x \in X$

$$\bar{\varphi}(x) \coloneqq \varphi(x) - f_0(x) + \varphi^*(f_0).$$

Es gilt (wie einfache Rechnungen zeigen), dass $\bar{\varphi}$ konvex und unterhalbstetig ist, und wir haben $\bar{\varphi} \geq 0$ (denn $\varphi^*(f_0) \geq f_0(x) - \varphi(x)$ für $x \in X$). Dann gilt nach Schritt 1: $(\bar{\varphi})^{**} = \bar{\varphi}$. Wir berechnen

$$\begin{aligned}(\bar{\varphi})^*(f) &= \sup_{x\in X}\big(f(x) - \bar{\varphi}(x)\big) = \sup_{x\in X}\big(f(x) - \varphi(x) + f_0(x) - \varphi^*(f_0)\big)\\ &= \varphi^*(f+f_0) - \varphi^*(f_0)\end{aligned}$$

und weiter

$$\begin{aligned}(\bar{\varphi})^{**} &= \sup_{f\in X'}\big(f(x) - (\bar{\varphi})^*(f)\big) = \sup_{f\in X'}\big(f(x) - \varphi^*(f+f_0) + \varphi^*(f_0)\big)\\ &= \sup_{f\in X'}\big((f+f_0)(x) - \varphi^*(f+f_0) - f_0(x) + \varphi^*(f_0)\big)\\ &= \varphi^{**}(x) - f_0(x) + \varphi^*(f_0).\end{aligned}$$

Da $(\bar{\varphi})^{**} = \bar{\varphi}$ gilt, folgt $\varphi^{**} = \varphi$.

■

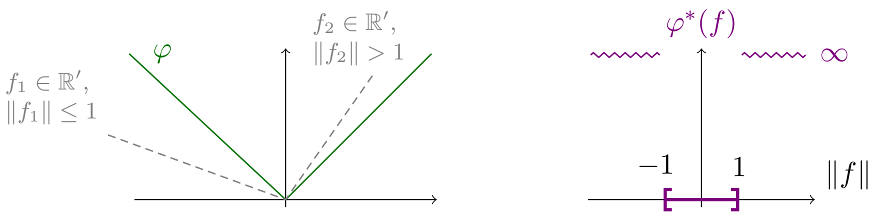


Abbildung 4.9: Beispiel 4.33 (i) für $\varphi(x) = |x|$ auf $\mathbb{R}$ mit zugehörigem φ^*

Beispiele 4.33.

(i) Sei $(X, \|\cdot\|)$ ein normierter $\mathbb{R}$-Vektorraum und $\varphi(x) \coloneqq \|x\|$ für alle $x \in X$. Dann gilt:

$$\varphi^*(f) = \sup_{x\in X}\big(f(x) - \varphi(x)\big) = \big(f(x) - \|x\|\big) = \begin{cases} 0, & \text{falls } \|f\| \le 1 \\ \infty, & \text{falls } \|f\| > 1. \end{cases}$$

Dies erhalten wir wie folgt. Es gilt:

$$\|f\| = \sup_{x\in X\setminus\{0\}} \frac{f(x)}{\|x\|}.$$

Für $\|f\| > 1$ existiert ein $x \in X$ mit $f(x)/\|x\| > 1$ und damit $f(x) - \|x\| > 0$. Ersetze nun x durch αx mit $\alpha \in \mathbb{R}_{>0}$ und betrachte $\alpha \to \infty$. Es folgt:

$$\sup_{x\in X}\big(f(x) - \|x\|\big) = \infty.$$

Der andere Fall ergibt sich ähnlich. (Abbildung 4.9) Es folgt mit Theorem 4.32:

$$\|x\| = \varphi(x) = \varphi^{**}(x) = \sup_{f\in X'}\big(f(x) - \varphi^*(f)\big) = \sup_{\substack{f\in X',\\ \|f\|\le 1}} f(x).$$

(ii) Sei X ein normierter Raum und $K \subset X$. Wir definieren die sogenannte *Indikatorfunktion von K* für alle $x \in X$ durch

$$I_K(x) \coloneqq \begin{cases} 0, & \text{falls } x \in K \\ \infty & \text{falls } x \notin K. \end{cases}$$

(Achtung: dies ist *nicht* die charakteristische Funktion von K.) Einfache Überlegungen liefern: I_K ist genau dann konvex, wenn K konvex ist und I_K ist genau dann unterhalbstetig, wenn K abgeschlossen ist. Die *Trägerfunktion zu K* ist dann definiert durch die konjugierte Funktion $(I_K)^*$ von I_K. Man kann nun folgende Aussagen zeigen:

– Falls $K = M \subset X$ ein Unterraum ist, so gilt:

$$(I_M)^* = I_{M^\perp}, \quad (I_M)^{**} = I_{(M^\perp)^\perp}.$$

– Falls M zusätzlich abgeschlossen ist, so gilt

$$(I_M)^{**} = I_M \quad \text{und somit} \quad \big(M^\perp\big)^\perp = M.$$

Für $a, b \in \mathbb{R}$ und $\emptyset \neq K = [a, b] \subset \mathbb{R}$ erhalten wir $(I_K)^*$ als Funktion von $\mathbb{R}$ nach $\mathbb{R}$:

$$(I_K)^*(y) = \sup_{x\in\mathbb{R}}\big(x\cdot y - I_K(x)\big) = \sup_{x\in[a,b]} x\cdot y = \begin{cases} by, & \text{falls } y \ge 0 \\ ay, & \text{falls } y < 0. \end{cases}$$

(Siehe Abbildung 4.10.)

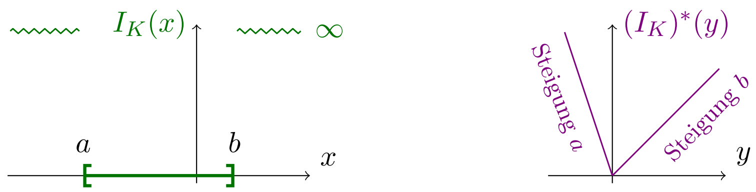


Abbildung 4.10: Beispiel 4.33 (ii) für $K = [a, b]$ mit $a = -3$ und $b = 1$

(iii) Sei $g\colon \mathbb{R} \to \mathbb{R}$ stetig differenzierbar mit

$$\lim_{x\to\pm\infty} \frac{g(x)}{|x|} = \infty.$$

Dann gilt: In

$$\sup_{x\in\mathbb{R}} \big(x \cdot y - g(x)\big)$$

wird das Supremum für ein endliches $x \in \mathbb{R}$ angenommen (da $x \cdot y - g(x) \to -\infty$ für $x \to \pm\infty$). Berechne x maximal als Lösung von $y - g'(x) = 0$. Falls g' streng monoton ist, gibt es höchstens eine solche Lösung. Für $y \in \mathbb{R}$ gilt dann

$$g^*(y) = (g')^{-1}(y)\, y - g\big((g')^{-1}(y)\big).$$

Falls $g \in C^2$ gilt, so können wir folgende Rechnung machen:

$$\begin{aligned}(g^*)'(y) &= (g')^{-1}(y) + \big((g')^{-1}(y)\big)'\, y - g'\big((g')^{-1}(y)\big)\, \big((g')^{-1}(y)\big)' \\ &= (g')^{-1}(y).\end{aligned}$$

Das heißt, dass wir unter geeigneten Voraussetzungen an g die Formel

$$(g^*)'(y) = (g')^{-1}(y)$$

erhalten. In der Theorie erhalten wir dann g^*, indem wir g ableiten, die Umkehrfunktion von g' bestimmen und zu dieser eine Stammfunktion finden.

5 Bairescher Kategoriensatz und seine Konsequenzen

Satz 5.1 (Bairescher Kategoriensatz). Sei (X, d) ein nicht-leerer vollständiger metrischer Raum. Sei $(A_k)_{k\in\mathbb{N}}$ eine Folge von in X abgeschlossenen Mengen und gelte

$$X = \bigcup_{k\in\mathbb{N}} A_k.$$

Dann gibt es ein $k_0 \in \mathbb{N}$, so dass das Innere von A_{k_0} nicht leer ist, d. h. so dass $A_{k_0}^\circ \neq \emptyset$ gilt.

Beweis. Angenommen für alle $k \in \mathbb{N}$ gilt $A_k^\circ = \emptyset$. Dann ist für alle offenen, nicht-leeren Teilmengen $U \subset X$ und alle $k \in \mathbb{N}$ die Menge $U \setminus A_k$ offen und nicht leer; insbesondere existiert in dieser Situation ein $x \in X$ und ein $\varepsilon \in \mathbb{R}_{>0}$, o. E. $\varepsilon \leq 1/k$, mit

$$\overline{B_\varepsilon(x)} \subset (U \setminus A_k).$$

Wir wählen für den ersten Schritt $U = X$ und konstruieren dann auf obige Weise induktiv Folgen $(x_k)_{k\in\mathbb{N}}$ in X und $(\varepsilon_k)_{k\in\mathbb{N}}$ in $\mathbb{R}_{>0}$, so dass für alle $k \in \mathbb{N}$

$$\varepsilon_k \leq 1/k \quad \text{und} \quad \overline{B_{\varepsilon_k}(x_k)} \subset \big(B_{\varepsilon_{k-1}}(x_{k-1}) \setminus A_k\big)$$

gilt. Dann ist $(x_k)_{k\in\mathbb{N}}$ eine Cauchy-Folge in X, denn: Zu jedem $\varepsilon \in \mathbb{R}_{>0}$ gibt es nach Konstruktion ein $k \in \mathbb{N}$ mit $\varepsilon_k \leq \varepsilon$ und für alle $\ell \in \mathbb{N}_{\geq k}$ gilt dann

$$x_\ell \in B_{\varepsilon_k}(x_k).$$

Da X vollständig ist, existiert ein $x \in X$ mit $x = \lim_{n\to\infty} x_n$. Weil für alle $k \in \mathbb{N}$ fast alle Folgenglieder von $(x_k)_{k\in\mathbb{N}}$ im abgeschlossenen ε_k-Ball um x_k liegen, muss dies auch für den Grenzwert x gelten, d h. für alle $k \in \mathbb{N}$ gilt:

$$x \in \overline{B_{\varepsilon_k}(x_k)} \subset X \setminus A_k.$$

Es folgt der Widerspruch

$$x \in \bigcap_{k\in\mathbb{N}} (X \setminus A_k) = X \setminus \bigcup_{k\in\mathbb{N}} A_k = \emptyset.$$

∎

Bemerkung:
Die Vollständigkeit von X ist hier entscheidend. Als Gegenbeispiel betrachte man $X = \mathbb{Q}$.

Satz 5.2 (Prinzip der gleichmäßigen Beschränktheit). Sei (X, d) ein vollständiger metrischer Raum, Y ein normierter Raum und $\mathcal{F} \subset C^0(X, Y)$. Es gelte für alle $x \in X$:

$$\sup_{f \in \mathcal{F}} \|f(x)\|_Y < \infty.$$

Dann existieren ein $x_0 \in X$, ein $\varepsilon_0 \in \mathbb{R}_{>0}$ und ein $C \in \mathbb{R}_{>0}$, so dass gilt:

$$\forall\, x \in \overline{B_\varepsilon(x_0)} \ \ \forall\, f \in \mathcal{F}\colon \quad \|f(x)\|_Y \leq C.$$

Beweis. Die Menge

$$A_k := \bigcap_{f \in \mathcal{F}} \big\{x \in X \bigm| \|f(x)\|_Y \leq k\big\}$$

ist für alle $k \in \mathbb{N}$ abgeschlossen. Außerdem gibt es nach Voraussetzung für alle $x \in X$ ein $k \in \mathbb{N}$, so dass für alle $f \in \mathcal{F}$ gilt: $\|f(x)\| \leq k$. Also gilt

$$X = \bigcup_{k \in \mathbb{N}} A_k.$$

Der Bairescher Kategoriensatz (5.1) liefert: es existieren $k_0 \in \mathbb{N}$, $x_0 \in X$, $\varepsilon_0 \in \mathbb{R}_{>0}$ mit

$$\overline{B_{\varepsilon_0}(x_0)} \subset A_{k_0}.$$

■

Satz 5.3 (Banach-Steinhaus). Sei X ein Banachraum, Y ein normierter Raum und $\mathcal{T} \subset L(X, Y)$. Für alle $x \in X$ gelte: $\sup_{T \in \mathcal{T}} \|Tx\|_Y < \infty$. Dann folgt schon

$$\sup_{T \in \mathcal{T}} \|T\| < \infty.$$

Beweis. Das Prinzip der gleichmäßigen Beschränktheit (Satz 5.2) liefert: es gibt $x_0 \in X$, $\varepsilon_0 \in \mathbb{R}_{>0}$, $C \in \mathbb{R}_{>0}$, so dass für alle $x \in X$ gilt:

$$\|x - x_0\| \leq \varepsilon_0 \implies \forall\, T \in \mathcal{T}\colon\ \|Tx\| \leq C.$$

Damit folgt, dass für alle $x \in \overline{B_{\varepsilon_0}(x_0)}$ und alle $T \in \mathcal{T}$ gilt:

$$\left\| T\left(\frac{x - x_0}{\varepsilon_0}\right)\right\| \leq \frac{C + \sup_{\tilde{T} \in \mathcal{T}} \|\tilde{T} x_0\|_Y}{\varepsilon_0} =: C_0.$$

Es folgt $\|T\| \leq C_0$, denn $\frac{x - x_0}{\varepsilon_0}$ nimmt alle Vektoren der Norm kleiner-gleich eins an.

■

Bemerkung 5.4.

(i) Der Satz von Banach-Steinhaus (5.3) liefert das erstaunliche Resultat, dass eine punktweise beschränkte Familie von stetigen linearen Operatoren schon beschränkt in der Operatornorm ist.

(ii) Punktweise Grenzwerte von stetigen Funktionen sind im Allgemeinen nicht stetig. Aus dem Satz von Banach-Steinhaus (5.3) folgt aber, dass dies für lineare Abbildungen doch gilt.

(iii) Ist $(T_n)_{n\in\mathbb{N}}$ eine Folge von linearen Operatoren mit punktweisem Grenzwert T, so gilt im Allgemeinen nicht $\|T_n - T\| \to 0$ für $n \to \infty$. Es gilt aber eine etwas schwächere Aussage, siehe Korollar 5.5.

Korollar 5.5. Seien X und Y Banachräume. Sei $(T_n)_{n\in\mathbb{N}}$ eine Folge in $L(X,Y)$, so dass für alle $x \in X$ die Folge $(T_n x)_{n\in\mathbb{N}}$ in Y konvergiert. Setzen wir

$$Tx \coloneqq \lim_{n\to\infty} T_n x$$

für alle $x \in X$, so gilt:

(a) $\sup_{n\in\mathbb{N}} \|T_n\| < \infty$

(b) $T \in L(X,Y)$

(c) $\|T\| \leq \liminf_{n\to\infty} \|T_n\|$

Beweis. Zunächst ist aufgrund der Linearität des Grenzwerts klar, dass auch T linear ist. Aussage (a) folgt unmittelbar aus dem dem Satz von Banach-Steinhaus (5.3). Es existiert somit ein $C \in \mathbb{R}_{>0}$, welches $\|T_n\| \leq C$ für alle $n \in \mathbb{N}$ erfüllt. Somit gilt für alle $x \in X$ und alle $n \in \mathbb{N}$:

$$\begin{aligned} \|T_n x\| &\leq \|T_n\| \cdot \|x\| \\ &\leq C \|x\| \end{aligned}$$

Aus der unteren Ungleichung erhalten wir im Grenzwert $n \to \infty$ für alle $x \in X$ die Ungleichung $\|Tx\| \leq C\|x\|$, woraus folgt, dass T stetig ist. Damit ist also (b) gezeigt. Aus der oberen Ungleichung erhalten wir

$$\liminf_{n\to\infty} \|T_n x\| \leq \liminf_{n\to\infty} \|T_n\| \cdot \|x\|,$$

und weil $(T_n x)_{n\in\mathbb{N}}$ für alle $x \in X$ konvergiert, gilt für alle $x \in X$ mit $\|x\| \leq 1$ (mithilfe der Stetigkeit der Norm):

$$\|Tx\| = \lim_{n\to\infty} \|T_n x\| = \liminf_{n\to\infty} \|T_n x\| \leq \liminf_{n\to\infty} \|T_n\| \cdot \|x\| \leq \liminf_{n\to\infty} \|T_n\|.$$

Daraus folgt (c).

∎

Korollar 5.6. Sei Z ein Banachraum über $\mathbb{K}$ und $B \subset Z$ eine Teilmenge von Z. Für alle $f \in Z'$ sei $f(B) \subset \mathbb{K}$ beschränkt. Dann ist B beschränkt in Z.

Beweis. Nutze Banach-Steinhaus (Satz 5.3) mit (in den dortigen Bezeichnern) $X = Z'$, $Y = \mathbb{K}$ und

$$\mathcal{T} = \{J_{Z'}(b) \mid b \in B\} \subset Z''$$

(mit $J_{Z'}$ wie nach Definition 4.17, d. h. für $b \in B$ und $f \in Z'$ gilt $J_{Z'}(b)(f) = f(b)$). Nach Voraussetzung gilt für alle $f \in Z'$:

$$\sup_{b \in B} |J_{Z'}(b)(f)| = \sup_{b \in B} |f(b)| < \infty.$$

Nach dem Satz von Banach-Steinhaus gilt also:

$$\sup_{b \in B} \|J_{Z'}(b)\| = \sup_{T \in \mathcal{T}} \|T\| < \infty.$$

Nach Satz 4.18 ist J_X aber eine Isometrie, also folgt

$$\sup_{b \in B} \|b\| = \sup_{b \in B} \|J_{Z'}(b)\| < \infty,$$

aber dies bedeutet gerade, dass B in Z beschränkt ist.

■

Bemerkung: Im endlich-dimensionalen besagt dieses Korollar: Eine Menge ist beschränkt, falls die Projektion auf alle Komponenten beschränkt ist. (Das Korollar ist also eine Verallgemeinerung dieser Tatsache.)

Korollar 5.7. Sei Z ein Banachraum über $\mathbb{K}$ und sei $A \subset Z'$ eine Teilmenge des Dualraums. Sei außerdem für alle $x \in Z$ die Menge

$$A(x) := \{f(x) \mid f \in A\}$$

beschränkt in $\mathbb{K}$. Dann ist A beschränkt in Z'.

Beweis. Folgt unmittelbar aus Banach-Steinhaus (Satz 5.3) mit (in den dortigen Bezeichnern) $X = Z$, $Y = \mathbb{K}$ und $\mathcal{T} = A$.

■

Definition 5.8. Seien X, Y topologische Räume und $f\colon X \to Y$ eine Abbildung. Dann ist f *offen*, wenn Bilder offener Mengen offen sind, d. h. wenn für alle in X offenen Teilmengen $U \subset X$ auch $f(U)$ offen in Y ist.

Bemerkung: Sind X, Y normierte Räume und ist ist f linear, so ist f genau dann offen, wenn es ein $\delta \in \mathbb{R}_{>0}$ gibt mit

$$B_\delta(0) \subset f\big(B_1(0)\big).$$

Satz 5.9 (Satz von der offenen Abbildung)**.** Seien X und Y Banachräume und sei $T \in L(X, Y)$. Dann ist T genau dann surjektiv, wenn T offen ist.

Beweis. „$\Leftarrow$“: Es existiert also ein $\delta \in \mathbb{R}_{>0}$ mit

$$B_\delta(0) \subset T\big(B_1(0)\big).$$

Durch Skalierung erhalten wir für alle $k \in \mathbb{N}$:

$$B_{k\delta}(0) \subset T\big(B_k(0)\big).$$

Daraus folgt, dass T surjektiv ist.

„$\Rightarrow$“: Weil T surjektiv ist, gilt

$$Y = \bigcup_{k\in\mathbb{N}} \underbrace{\overline{T\big(B_k(0)\big)}}_{=:A_k}.$$

Wir wenden den Bairescher Kategoriensatz (5.1) auf die Folge $(A_k)_{k\in\mathbb{N}}$ an und erhalten somit $\varepsilon_0 \in \mathbb{R}_{>0},\ y_0 \in Y,\ k_0 \in \mathbb{N}$ mit

$$\overline{B_{\varepsilon_0}(y_0)} \subset \overline{T\big(B_{k_0}(0)\big)}.$$

Sei $y \in \overline{B_{\varepsilon_0}(0)}$, dann gilt $y_0 + y \in \overline{B_{\varepsilon_0}(y_0)}$. Wähle dann eine Folge $(x_n)_{n\in\mathbb{N}}$ in $B_{k_0}(0)$ mit

$$Tx_n \to y_0 + y \qquad \text{für } n \to \infty.$$

Sei außerdem $x_0 \in X$ mit $Tx_0 = y_0$. Dann erhalten wir

$$T(x_n - x_0) = Tx_n - y_0 \to y \qquad \text{für } n \to \infty.$$

Daraus folgt:

$$T\Bigg(\underbrace{\frac{x_n - x_0}{k_0 + \|x_0\|}}_{\in B_1(0)}\Bigg) \to \underbrace{\frac{y}{k_0 + \|x_0\|}}_{\in B_\delta(0)} \qquad \text{für } n \to \infty$$

mit $\delta := \frac{\varepsilon_0}{k_0 + \|x_0\|}$. Das bedeutet aber:

$$B_\delta(0) \subset \overline{T\big(B_1(0)\big)}.$$

Um die Behauptung zu zeigen, benötigen wir diese Inklusion aber ohne den Abschluss auf der rechten Seite. Sei dazu nun $y \in B_\delta(0)$. Dann gibt es ein $x \in B_1(0)$ mit $\|y - Tx\| < \delta/2$. Daraus folgt:

$$2(-Tx + y) \in B_\delta(0).$$

Indem wir $y_1 := y$ setzen, erhalten wir so induktiv Folgen $(y_k)_{k\in\mathbb{N}}$ in $B_\delta(0)$ und $(x_k)_{k\in\mathbb{N}}$ in $B_1(0)$ mit folgender Eigenschaft für alle $k \in \mathbb{N}$:

$$y_{k+1} = 2(y_k - Tx_k).$$

Es folgt für alle $k \in \mathbb{N}$

$$2^{-k} y_{k+1} = 2^{-k+1} y_k - \underbrace{T(2^{-k+1} x_k)}_{=:a_k}$$

oder durch umstellen

$$a_k = 2^{-(k-1)} y_k - 2^{-k} y_{k+1}.$$

Sei $m \in \mathbb{N}$. Dann ist $\sum_{k=1}^{m} a_n$ offenbar eine Teleskopsumme, also gilt:

$$\sum_{k=1}^{m} a_k = y_1 - 2^{-m} y_{m+1}.$$

Da $(y_k)_{k \in \mathbb{N}}$ beschränkt ist, gilt $2^{-m} y_{m+1} \to 0$ für $m \to \infty$. Also erhalten wir:

$$\lim_{m \to \infty} T\Big(\sum_{k=1}^{m} 2^{-(k-1)} x_k\Big) = \lim_{m \to \infty} \sum_{k=1}^{m} a_k = y_1.$$

Außerdem konvergiert die Reihe $\sum_{k=1}^{\infty} 2^{-(k-1)} x_k$ in X, denn es gilt

$$\sum_{k=1}^{\infty} \|2^{-(k-1)} x_k\| \leq \sum_{k=1}^{\infty} 2^{-(k-1)} = \sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k = 2,$$

womit man leicht zeigt, dass $\big(\sum_{k=1}^{m} 2^{-(k-1)} x_k\big)_{m \in \mathbb{N}}$ eine Cauchy-Folge in X ist (und damit auch konvergent, aufgrund der Vollständigkeit von X). Sei also $\tilde{x} := \sum_{k=1}^{\infty} 2^{-(k-1)} x_k$. Dann gilt wegen der Stetigkeit von T auch $T(\tilde{x}) = y_1 = y$ und aus der obigen Betrachtung folgt außerdem $\|\tilde{x}\| \leq 2 < 3$. Also gilt $B_\delta(0) \subset T\big(B_3(0)\big)$ und durch Skalierung erhalten wie gewünscht:

$$B_{\delta/3}(0) \subset T\big(B_1(0)\big).$$

■

Satz 5.10 (Satz von der inversen Abbildung)**.** Seien X und Y Banachräume und sei $T \in L(X, Y)$ bijektiv. Dann gilt: $T^{-1} \in L(X, Y)$.

Beweis. Weil T surjektiv ist, liefert der Satz von der offenen Abbildung (5.9), dass T offen ist. Daraus folgt unmittelbar, dass T^{-1} stetig ist.

■

Korollar 5.11. Sei X ein Vektorraum mit zwei Normen $\|\cdot\|_{①}$ und $\|\cdot\|_{②}$. Außerdem gebe es ein $C_1 \in \mathbb{R}_{>0}$, so dass für alle $x \in X$ die Ungleichung $\|x\|_{②} \leq C_1 \|x\|_{①}$ gilt. Sei weiter $(X, \|\cdot\|_{①})$ ein Banachraum. Dann gilt: $(X, \|\cdot\|_{②})$ ist genau dann ein Banachraum, wenn es ein $C_2 \in \mathbb{R}_{>0}$ gibt mit $\|x\|_{①} \leq C_2 \|x\|_{②}$ für alle $x \in X$.

Beweis. „$\Leftarrow$“ ist klar.
„$\Rightarrow$“: Die Identität

$$\mathrm{id}\colon (X, \|\cdot\|_{①}) \to (X, \|\cdot\|_{②})$$

ist stetig und bijektiv. Der Satz von der inversen Abbildung (5.10) liefert:

$$\mathrm{id}^{-1}\colon (X, \|\cdot\|_{②}) \to (X, \|\cdot\|_{①})$$

ist stetig. Daraus erhalten wir eine Konstante C_2 wie gefordert.

■

Satz 5.12 (Satz vom abgeschlossenen Graphen)**.** Seien X und Y Banachräume, $D(T)$ ein Unterraum von X und sei $T\colon D(T) \to Y$ eine lineare Abbildung. Sei

$$\operatorname{graph}(T) \coloneqq \big\{(x, Tx) \in X \times Y \;\big|\; x \in D(T)\big\}$$

der *Graph von T in $X \times Y$*. Dabei wird $X \times Y$ zu einem Banachraum bezüglich folgender Norm:

$$\|\cdot\|\colon X \times Y \to \mathbb{R}_{\geq 0}, \quad (x, y) \mapsto \|x\|_X + \|y\|_Y.$$

Dann gilt: Ist $D(T)$ abgeschlossen, so ist $\operatorname{graph}(T)$ genau dann abgeschlossen in $X \times Y$, wenn $T \in L\big(D(T), Y\big)$ gilt.

Beweis. „$\Leftarrow$": klar, denn: Sei $(x_n, Tx_n)_{n\in\mathbb{N}}$ eine Folge in $\operatorname{graph}(T)$, die in $X \times Y$ konvergiert. Sei $x \in X$ der Grenzwert von $(x_n)_{n\in\mathbb{N}}$, so gilt wegen der Stetigkeit von T auch $Tx_n \to Tx$ für $n \to \infty$, aber $(x, Tx) \in \operatorname{graph}(T)$ gilt nach Definition von $\operatorname{graph}(T)$.

„$\Rightarrow$": Da $\operatorname{graph}(T)$ abgeschlossen ist, muss dieser Raum (mit der Einschränkung der Norm von $X \times Y$) ein Banachraum sein. Seien P_X und P_Y die Projektionen $X \times Y \to X$ bzw. $X \times Y \to Y$. Dann sind P_X, P_Y stetig und linear und P_X ist bijektiv von $\operatorname{graph}(T)$ auf $D(T)$. Der Satz von der inversen Abbildung (5.10) liefert:

$$P_X^{-1} \in L\big(D(T), \operatorname{graph}(T)\big).$$

Daraus folgt: $T = P_Y P_X^{-1} \in L\big(D(T), Y\big)$.

■

5.13 (Projektoren)**.** Sei Z ein Vektorraum und $A \subset Z$. Sei weiter X ein normierter Raum und $Y \subset X$ ein Unterraum.

(1) Eine Abbildung $P\colon Z \to Z$ ist eine *Projektion auf* A, falls folgende Bedingungen erfüllt sind:

$$P(Z) \subset A \quad \text{und} \quad P|_A = \mathrm{Id}_A\,.$$

Äquivalent kann man fordern:

$$P(Z) = A \quad \text{und} \quad P^2 = P \circ P = P.$$

Es folgt:

$$P\,(\mathrm{Id} - P) = (\mathrm{Id} - P)\,P = 0.$$

Beispiel: orthogonale Projektionen im euklidschen Raum sind Projektionen.

(2) Sei $P\colon X \to Y$ eine lineare Projektion auf Y. Dann gilt $Y = R(P)$ und $\mathrm{Id} = (\mathrm{Id} - P) + P$ und $(\mathrm{Id} - P)$ ist eine Projektion auf $N(P)$. (Zur Definition des Bildraums $R(\,\cdot\,)$ und des Nullraums $N(\,\cdot\,)$, siehe Definition 3.6.)

Es gilt $X = N(P) \oplus R(P) = R(\mathrm{Id} - P) \oplus N(\mathrm{Id} - P)$, $N(P) = R(\mathrm{Id} - P)$ und $R(P) = N(\mathrm{Id} - P)$, denn:

Für $x \in X$ gilt: $x = (x - Px) + Px$ mit $(x - Px) \in N(P)$ und $Px \in R(P)$. Ist $x \in N(P) \cap R(P)$, so gilt $Px = 0$ und $x = Px$, also $x = 0$. Außerdem gelten folgende Äquivalenzen:

$$x \in N(\mathrm{Id} - P) \iff x - Px = 0 \iff x = Px \iff x \in R(P).$$

Also gilt $N(\mathrm{Id} - P) = R(P)$.

(3) Eine Abbildung $P\colon X \to X$ ist ein *Projektor auf* Y, falls P eine stetige lineare Projektion auf Y ist. Es sei

$$\mathrm{Pr}(X) \coloneqq \{P\colon X \to X \mid P \text{ ist Projektor}\}.$$

Falls P ein Projektor ist, so ist klarerweise auch $\mathrm{Id} - P$ ein Projektor und $N(P)$ sowie $R(P) = N(\mathrm{Id} - P)$ sind abgeschlossen.

Satz 5.14 (Satz vom abgeschlossenen Komplement)**.** Sei X ein Banachraum und seien Y und Z Unterräume von X. Außerdem gelte $X = Z \oplus Y$ und Y sei abgeschlossen. Dann gilt: Der Unterraum Z ist genau dann abgeschlossen, wenn es einen stetigen Projektor P auf Y mit $N(P) = Z$ gibt.

Beweis. „$\Leftarrow$" ist klar, da P stetig ist.

„$\Rightarrow$": Sei $\tilde{X} \coloneqq Z \times Y$. Definiere

$$T\colon \tilde{X} \to X, \quad (z, y) \mapsto z + y.$$

Dann ist T linear, bijektiv und stetig (wie man mithilfe der $\triangle$-Ungleichung einsieht). Weil Y und Z abgeschlossen sind, ist auch $\tilde{X}$ ein abgeschlossener Unterraum von $X \times X$ und damit ein Banachraum. Der Satz von der inversen Abbildung (5.10) liefert $T^{-1} \in L(X, \tilde{X})$. Sei

$$P\colon X \to Y, \quad z + y \mapsto y \quad (\text{mit } z \in Z \text{ und } y \in Y).$$

Dann ist P eine lineare Projektion auf Y und es gilt $N(P) = Z$. Außerdem ist P stetig, denn es gilt $P = P_Y T^{-1}$, wobei P_Y die Projektion $\tilde{X} \to Y$ auf die zweite Komponente ist. Damit ist P der gesuchte Projektor.

■

Definition 5.15. Seien X und Y normierte Vektorräume und $T \in L(X, Y)$. Dann ist der *adjungierte Operator* T' die Abbildung

$$\begin{aligned} T'\colon Y' &\to X' \\ y' &\mapsto \begin{pmatrix} X \mapsto \mathbb{K} \\ x \mapsto y'(Tx) \end{pmatrix}. \end{aligned}$$

Bemerkung: Es gilt

$$|(T'y')(x)| = |y'(Tx)| \leq \|y'\| \cdot \|Tx\| \leq \|y'\| \cdot \|T\| \cdot \|x\|$$

für alle $y' \in Y'$ und alle $x \in X$. Dies zeigt, dass T' wohldefiniert ist ($T'y' \in X'$ für alle $y' \in Y'$) und dass $\|T'y'\| \leq \|y'\| \cdot \|T\|$ gilt. Daraus folgt $T' \in L(Y', X')$.

Beispiel 5.16 (Shift-Operator). Wir betrachten auf ℓ^2 über $\mathbb{R}$ den Shift-Operator

$$T(x_1, x_2, x_3, \dots) \coloneqq (x_2, x_3, \dots).$$

Wir möchten nun T' bestimmen. Später zeigen wir: $(\ell^2)'$ kann mit ℓ^2 identifiziert werden. Jedes $x \in \ell^2$ definiert einen linearen Operator auf ℓ^2 durch $x'(y) = (x, y)_{\ell^2} \coloneqq \sum_{i=1}^\infty x_i y_i$ für alle $y \in \ell^2$. Dies liefert schon $(\ell^2)'$. Wir fordern für $y' \in (\ell^2)'$, dargestellt durch $(y_n)_{n\in\mathbb{N}}$:

$$y'(Tx) = \sum_{n=1}^\infty x_{n+1} y_n = \sum_{n=2}^\infty x_n \tilde{y}_n \stackrel{!}{=} (T'y')(x),$$

wobei $\tilde{y}_n = y_{n-1}$ für alle $n \in \mathbb{N}_{\geq 2}$. Dies zeigt, dass

$$T' \colon Y' \to X', \quad (y_1, y_2, \dots) \mapsto (0, y_1, y_2, \dots)$$

gilt. Dabei haben wir $TT' = \mathrm{Id}$, aber $T'T \neq \mathrm{Id}$.

Satz 5.17. Seien X und Y normierte Räume.

(i) Die Abbildung

$$\begin{aligned} ' \colon L(X,Y) &\to L(Y', X') \\ y &\mapsto y' \end{aligned}$$

ist eine lineare Isometrie.

(ii) Sei Z ein weiterer normierter Raum. Für alle $T \in L(X,Y)$ und $S \in L(Y,Z)$ gilt

$$(ST)' = T'S'.$$

Beweis.

(i) Die Linearität rechnet man leicht nach. Sei $T \in L(X,Y)$. Dann gilt für alle $y' \in Y'$

$$\|T'y'\| \leq \|y'\| \cdot \|T\|, \quad \text{woraus} \quad \|T'\| \leq \|T\| \quad \text{folgt.}$$

Dass wir tatsächlich auch Gleichheit haben, sehen wir wie folgt ein (– um die Notation übersichtlicher zu halten, bezeichne $\bar{B}_1^X \coloneqq \overline{B_1^X(0)}$ die abgeschlossene Einheitskugel in X und $\bar{B}_1^{Y'} \coloneqq \overline{B_1^{Y'}(0)}$ diejenige in Y'):

$$\begin{aligned} \|T\| &= \sup_{x\in\bar{B}_1^X} \|Tx\| = \sup_{x\in\bar{B}_1^X} \sup_{y'\in\bar{B}_1^{Y'}} |y'(Tx)| \\ &= \sup_{y'\in\bar{B}_1^{Y'}} \sup_{x\in\bar{B}_1^X} |y'(Tx)| = \sup_{y'\in\bar{B}_1^{Y'}} \|T'y'\| = \|T'\|. \end{aligned}$$

(Die zweite Gleichheit gilt dabei nach Satz 4.18.)

(ii) Seien $T \in L(X,Y), S \in L(Y,Z)$ und $z' \in Z'$ sowie $x \in X$. Dann gilt:

$$\big((ST)'z'\big)(x) = z'(STx) = (S'z')(Tx) = (T'S'z')(x).$$

Es folgt die Behauptung.

■

Mit Hilfe der adjungierten Abbildung können wir die Lösbarkeit von linearen Gleichungen untersuchen. Dazu benötigen wir zunächst einige Hilfsaussagen.

Satz 5.18. Seien X, Y Banachräume und sei $T \in L(X,Y)$. Dann gilt

$$\overline{R(T)} = \big(N(T')\big)^{\perp}.$$

Beweis. „$\subset$": Sei $x \in X$ und $y := Tx \in R(T)$. Gilt $y' \in N(T')$, so folgt

$$y'(y) = y'(Tx) = \underbrace{(T'y')}_{=0}(x) = 0.$$

Es folgt $R(T) \subset (N(T'))^{\perp}$. Da $N(T')^{\perp}$ abgeschlossen ist (Bemerkung 4.21 (ii)), folgt

$$\overline{R(T)} \subset \big(N(T')\big)^{\perp}.$$

„$\supset$": Setze $U := \overline{R(T)}$. Es sei $y \notin \overline{U} = U$. Zeigen wir nun $y \notin (N(T'))^{\perp}$, so sind wir fertig. Aus Korollar 4.16 erhalten wir ein $y' \in Y'$ mit $Y'|_U = 0$ und $y'(y) \neq 0$. Insbesondere gilt $y'(Tx) = 0$ für alle $x \in X$. Also haben wir sicher $y' \in N(T')$ und wegen $y'(y) \neq 0$ gilt auch $y' \notin (N(T'))^{\perp}$.

■

Korollar 5.19. Seien X, Y Banachräume und sei $T \in L(X,Y)$. Sei weiter $R(T)$ abgeschlossen und $y \in Y$. Dann ist $Tx = y$ genau dann lösbar, wenn $y'(y) = 0$ für alle $y' \in N(T')$.

Beweis. Die Hinrichtung ist klar. Zur Rückrichtung: Aus $y'(y) = 0$ für alle $y' \in N(T')$ folgt sofort $y \in (N(T'))^{\perp}$ und damit:

$$y \in \big(N(T')\big)^{\perp} = \overline{R(T)} = R(T).$$

■

Bemerkung:

(i) Korollar 5.19 liefert die Existenz einer Lösung durch Bedingungen an den Nullraum von T'.

(ii) Falls T' injektiv ist, brauchen wir im Fall, dass $R(T)$ abgeschlossen ist, *keine* Zusatzbedingungen mehr, denn: Mit $N(T') = \{0\}$ und $R(T)$ abgeschlossen sind die Voraussetzungen aus Korollar 5.19 sicher erfüllt.

Im Allgemeinen ist es nicht einfach, die Abgeschlossenheit von $R(T)$ zu überprüfen. Wir wollen nun also Kriterien finden, die dies vereinfachen.

Satz 5.20. Sei X ein normierter Raum und Y ein abgeschlossener Unterraum. Dann wird durch

$$X/U \to \mathbb{R}_{\geq 0}, \quad \hat{x} \mapsto \inf_{y \in \hat{x}} \|y\|$$

eine Norm auf dem Quotientenvektorraum definiert. Ist X vollständig, so auch X/U.

Der Beweis ist eine einfache Übungsaufgabe.

Lemma 5.21. Seien X, Y Banachräume und sei $T \in L(X,Y)$. Sei weiter $R(T)$ abgeschlossen. Dann existiert ein $K \in \mathbb{R}_{\geq 0}$ mit folgender Eigenschaft:

$$\forall\, y \in R(T)\ \exists\, x \in X\colon\ Tx = y \wedge \|x\| \leq K\|y\|.$$

Beweis. Sei $\hat{T}$ die lineare und stetige Bijektion, die folgendes Diagramm kommutativ macht:

$$\begin{tikzcd}
X \arrow[r, "T"] \arrow[d] & R(T) \\
X/N(T) \arrow[ur, "\hat{T}"'] &
\end{tikzcd}$$

(dabei sei die vertikale Abbildung die kanonische Projektion). Da $R(T)$ abgeschlossener Teilraum des Banachraums Y ist, ist auch $R(T)$ ein Banachraum. Auch $X/N(T)$ ist ein Banachraum, da $N(T)$ abgeschlossen ist. Der Satz von der inversen Abbildung (5.10) liefert, dass $\hat{T}^{-1}$ stetig ist. Also existiert ein $\hat{K} \in \mathbb{R}_{\geq 0}$ mit $\|\hat{T}^{-1}y\| \leq \hat{K}\|y\|$. Mit der Definition der Norm auf $X/N(T)$ folgt: Es existiert ein $x \in \hat{T}^{-1}y$ mit $\|x\| \leq (\hat{K}+1)\|y\|$ und da $\hat{T}^{-1}y$ gerade alle $x \in X$ mit $Tx = y$ enthält, folgt die Behauptung.

■

Lemma 5.22. Seien X, Y Banachräume über $\mathbb{K}$ und sei $T \in L(X,Y)$. Weiter sei $K \in \mathbb{R}_{>0}$, so dass für alle $y' \in Y'$ gilt:

$$K\|y'\| \leq \|T'y'\|.$$

Dann ist T offen und insbesondere surjektiv.

Beweis. Es bezeichne $U_\varepsilon \coloneqq B_\varepsilon^X(0)$ den offenen ε-Ball in X und $V_\varepsilon \coloneqq B_\varepsilon^Y(0)$ denjenigen in Y. Bei der Betrachtung offener Abbildungen hatten wir schon gesehen, dass es genügt $V_K \subset T(U_1)$ zu zeigen. Wie im Beweis des Satzes von der offenen Abbildung (5.9) genügt es sogar, $V_K \subset \overline{T(U_1)} \eqqcolon D$ zu zeigen. Sei $y_0 \in V_K$, d. h. $\|y_0\| < K$. Angenommen es gilt $y_0 \notin D$. Dann existiert nach Hahn-Banach in der zweiten geometrischen Formulierung (Satz 4.15) ein $y' \in Y'$ und ein $\alpha \in \mathbb{R}$ mit

$$\operatorname{Re} y'(y) \leq \alpha < \operatorname{Re} y'(y_0) \leq |y'(y_0)|$$

für alle $y \in D$. Wegen $0 \in D$ und $y'(0) = 0$ gilt $0 \leq \alpha$. Wegen der echten Ungleichheit in $\alpha < \operatorname{Re} y'(y_0)$ können wir o. E. $\alpha > 0$ annehmen und indem wir y' durch y'/α ersetzen, können wir annehmen, dass $\alpha = 1$ gilt. Weil T linear ist, liegt für alle $y \in D$ und $\lambda \in \mathbb{K}$ mit $|\lambda| \leq 1$ auch λy in D. Ist also $y \in D$ mit $y'(y) \neq 0$, so gilt auch $y\,|y'(y)|/y'(y) \in D$, und damit folgt:

$$|y'(y)| = y'(y)\frac{|y'(y)|}{y'(y)} = y'\left(y\,\frac{|y'(y)|}{y'(y)}\right) \leq 1.$$

Also gilt für alle $y \in D$ sogar

$$|y'(y)| \leq 1 < |y'(y_0)|.$$

Für alle $x \in U_1$ (und damit $Tx \in D$) gilt somit

$$|y'(Tx)| = |(T'y')(x)| \leq 1,$$

woraus wir $\|T'y'\| \leq 1$ erhalten.

Es folgt
$$1 < |y'(y_0)| \le \|y'\|\,\|y_0\| \le K\|y'\| \le \|T'y'\| \le 1,$$
ein Widerspruch. Also muss doch $y_0 \in D$ gelten und damit sind wir fertig. ■

Bemerkung: Aus Satz 5.18 und Korollar 5.19 erhalten wir, dass T' genau dann injektiv ist, wenn T dichtes Bild hat. Außerdem folgt aus der Injektivität von T' und abgeschlossenem $R(T)$, dass T surjektiv ist. Obiges Lemma 5.22 verschärft diese Aussage noch.

Satz 5.23 (Satz vom abgeschlossenen Bild)**.** Seien X, Y Banachräume über $\mathbb{K}$ und sei $T \in L(X,Y)$. Dann sind folgende Aussagen äquivalent:

(i) $R(T)$ ist abgeschlossen

(ii) $R(T) = \big(N(T')\big)^{\perp}$

(iii) $R(T')$ ist abgeschlossen

(iv) $R(T') = \big(N(T)\big)^{\perp}$

Beweis. „(i)⇔(ii)“ folgt aus Satz 5.18.

„(i)⇒(iv)“: Es gilt stets $R(T') \subset (N(T))^{\perp}$ (da $T'y'(x) = y'(Tx) = 0$ für alle $x \in N(T)$). Sei $x' \in (N(T))^{\perp}$. Betrachte dann die Abbildung
$$z' \colon R(T) \to \mathbb{K}, \qquad y \mapsto x'(x) \quad \text{falls } Tx = y.$$
Es ist z' wohldefiniert, denn: für $x_1, x_2 \in X$ mit $Tx_1 = y = Tx_2$ gilt $T(x_1 - x_2) = 0$, also $x_1 - x_2 \in N(T)$, also $x'(x_1 - x_2) = 0$, also $x'(x_1) = x'(x_2)$. Die Abbildung z' ist auch stetig, denn: Aus Lemma 5.21 erhalten wir ein $K \in \mathbb{R}_{\ge 0}$, so dass für alle $y \in R(T)$ ein $x \in X$ existiert mit $Tx = y$ und $\|x\| \le K\|y\|$. Für solche x gilt:
$$|z'(y)| = |x'(x)| \le \|x'\|\,\|x\| \le K\|x'\|\,\|y\|.$$
Daraus folgt die Stetigkeit von z'. Wir setzen nun z' mittels Satz 4.6 zu $y' \in Y'$ fort. Dann gilt $x' = T'y'$, denn für alle $x \in X$ gilt
$$x'(x) = z'(Tx) = y'(Tx) = (T'y')(x).$$
Dies zeigt $(N(T))^{\perp} \subset R(T')$.

„(iv)⇒(iii)“: Klar, da $\big((N(T)\big)^{\perp}$ stets abgeschlossen ist (Bemerkung 4.21).

„(iii)⇒(i)“: Definiere $Z \coloneqq \overline{R(T)}$ und $S \in L(X,Z)$ durch $Sx \coloneqq Tx$ für alle $x \in X$. Für $y' \in Y'$ und $x \in X$ gilt dann
$$(T'y')(x) = y'(Tx) = y'|_Z(Sx) = \big(S'(y'|_Z)\big)(x).$$

D. h. $T'y' = S'(y'|_Z)$. Dies zeigt $R(T') \subset R(S')$. Ist $S'(z') \in R(S')$, so gilt für jede gemäß Satz 4.6 gewählte Fortsetzung $y' \in Y'$ von $z' \in Z'$ (nach demselben Argument wie oben) $S'z' = T'y'$. Dies zeigt $R(T') = R(S')$. Nach Voraussetzung ist somit $R(S')$ abgeschlossen.

Nach Satz 5.18 gilt $\overline{R(S)} = (N(S'))^{\perp}$ und da S dichtes Bild hat, folgt, dass S' injektiv ist. Also ist S' eine stetige Bijektion zwischen den Banachräumen Z' und $R(S')$. Aus dem Satz von der inversen Abbildung (5.10) folgt:

$$\exists\, C \in \mathbb{R}_{>0}\ \forall\, z' \in Z': \quad C\|z'\| \leq \|S'z'\|.$$

Lemma 5.22 liefert $R(S) = Z$. Damit gilt aber auch $R(T) = Z$ und weil Z nach Definition abgeschlossen ist, hat T somit abgeschlossenes Bild.

■

Bemerkung 5.24. Oft ist es einfacher zu zeigen, dass $R(T')$ abgeschlossen ist, als $R(T)$ zu bestimmen. Dann erhält man mit dem Satz vom abgeschlossenen Bild (5.23) direkt: $R(T)$ abgeschlossen, $R(T) = (N(T'))^{\perp}$. Man bekommt also Informationen über $R(T)$, indem man T' betrachtet.

Beispiel: Ist T' injektiv, dann folgt $N(T) = \{0\}$ und daraus $R(T) = (N(T'))^{\perp} = Y$.

6 Hilberträume

Siehe 2.7 (c). Eine wichtige Eigenschaft von Hilberträumen ist, dass stets orthogonale Projektionen auf abgeschlossene, konvexe Teilmengen existieren.

Satz 6.1 (Projektionssatz)**.** Sei H ein Hilbertraum, $K \subset H$ eine nicht-leere, abgeschlossene und konvexe Teilmenge. Dann existiert für jedes $f \in H$ ein eindeutiges $u \in K$, so dass gilt:

$$\|f - u\| = \min_{v \in K} \|f - v\| = \operatorname{dist}(f, K).$$

Das Element $u \in H$ ist eindeutig charakterisiert durch folgende Eigenschaft:

$$u \in K \quad \text{und für alle } v \in K \text{ gilt} \quad \operatorname{Re}\langle f - u, v - u\rangle \leq 0.$$

Das Element u nennen wir dann *Projektion von f auf K* und wir schreiben dafür $u = \mathrm{P}_K(f)$.

Beweis. Wähle eine Minimalfolge $(v_n)_{n\in\mathbb{N}}$ in K für das Infimum $\operatorname{dist}(f, K)$, also mit $d_n := \|f - v_n\| \to \inf_{v\in K} \|f - v\| =: d$. Wir behaupten, dass dann $(v_n)_{n\in\mathbb{N}}$ schon eine Cauchyfolge sein muss. Dazu wenden wir die Parallelogramm-Identität (Satz 2.8 (3)) auf $f - v_n$ und $f - v_m$ an. Wir erhalten:

$$\|2f - (v_n + v_m)\|^2 + \|v_n - v_m\|^2 = 2\big(d_n^2 + d_m^2\big).$$

Da K konvex ist, gilt $\frac{v_n+v_m}{2} \in K$, und es folgt:

$$\left\| f - \frac{v_n + v_m}{2} \right\| \geq d.$$

Nach Multiplizieren der vorherigen Ungleichung mit 1/4 folgt somit:

$$\left\| \frac{v_n - v_m}{2} \right\|^2 \leq \frac{1}{2}\big(d_n^2 + d_m^2\big) - d^2 \xrightarrow[n,m\to\infty]{} 0.$$

Also ist $(v_n)_{n\in\mathbb{N}}$ tatsächlich eine Cauchy-Folge. Weil K eine abgeschlossene Teilmenge eines vollständigen Raums ist, ist K insbesondere vollständig. Also konvergiert $(v_n)_{n\in\mathbb{N}}$ in K und es gibt ein $v \in K$ mit $v = \lim_{n\to\infty} v_n$. Für dieses gilt dann

$$\|f - v\| = \lim_{n\to\infty} \|f - v_n\| = d = \operatorname{dist}(f, K).$$

Wir zeigen nun, dass die beiden Charakterisierungen aus der Behauptung äquivalent sind. Sei hierzu für $u \in K$ die erste Bedingung erfüllt. Wähle $w \in K$. Dann folgt für $t \in [0, 1]$

$$v = (1 - t)u + tw \in K.$$

Damit erhalten wir für alle $t \in [0,1]$:

$$\begin{aligned} & \|f-u\| \le \|f-(1-t)u+tw\| = \|(f-u)-t(w-u)\| \\ \Longrightarrow \quad & \|f-u\|^2 \le \|f-u\|^2 - 2\operatorname{Re}\langle f-u, w-u\rangle + t^2\|w-u\|^2 \\ \Longrightarrow \quad & 2\operatorname{Re}\langle f-u, w-u\rangle \le t^2\|w-u\|^2 \end{aligned}$$

Für $t \to 0$ erhalten wir die zweite Charakterisierung. Sei umgekehrt letztere gegeben für $u \in K$. Dann gilt

$$\|u-f\|^2 - \|v-f\|^2 = \|u-f\|^2 - \|v-u+(u-f)\|^2 = 2\operatorname{Re}\langle f-u, v-u\rangle - \|u-v\|^2 \le 0,$$

woraus die erste Charakterisierung folgt.

Zur Eindeutigkeit: Angenommen $u_1, u_2 \in K$ erfüllen beide die zweite Charakterisierung. Dann gilt für alle $v \in K$

$$\operatorname{Re}\langle f-v_1, v-u_1\rangle \le 0 \quad \text{und} \quad \operatorname{Re}\langle f-v_2, v-u_2\rangle \le 0.$$

Wähle in der ersten Ungleichung $v = u_2$, in der zweiten $v = u_1$ und addiere beide Gleichungen. Dann erhalten wir

$$\|u_1-u_2\|^2 = \operatorname{Re}\langle u_1-u_2, v_1-v_2\rangle \le 0,$$

woraus $u_1 = u_2$ folgt.

∎

Bemerkung 6.2. Das Problem im obigen Satz ist ein Minimierungsproblem. Auch in anderen bereits bekannten Problemen wird ein Minimum durch Ungleichungen beschrieben. Betrachte zum Beispiel $F \in C^1([0,1])$ und $F(u) = \min_{v\in[0,1]} F(v)$. Dann gilt $F'(u) = 0$, falls $u \in (0,1)$, $F'(u) \ge 0$, falls $u = 0$ und $F'(u) \le 0$, falls $u = 1$. Oder zusammengefasst:

$$u \in [0,1] \quad \text{und für alle } v \in [0,1] \text{ gilt} \quad F'(u)(v-u) \ge 0.$$

Satz 6.3. Sei H ein Hilbertraum, $K \subset H$ eine nicht-leere, abgeschlossene und konvexe Teilmenge. Dann nimmt der Abstand durch die Anwendung von P_K nicht zu, d. h. P_K ist Lipschitz zur Konstante 1, d. h. für alle $f_1, f_2 \in H$ gilt

$$\|\mathrm{P}_K f_1 - \mathrm{P}_K f_2\| \le \|f_1 - f_2\|.$$

Beweis. Sei $u_1 := \mathrm{P}_K f_1$ und $u_2 := \mathrm{P}_K f_2$. Dann gilt für alle $v \in K$:

$$\operatorname{Re}\langle f_1-u_1, v-u_1\rangle \le 0 \quad \text{und} \quad \operatorname{Re}\langle f_2-u_2, v-u_2\rangle \le 0.$$

Wähle einmal $v = u_2$ und einmal $v = u_1$ und erhalte durch Addition:

$$\operatorname{Re}\langle f_2-u_2-f_1+u_1, u_1-u_2\rangle \le 0.$$

Daraus folgt:

$$\|u_1-u_2\|^2 \le \operatorname{Re}\langle f_1-f_2, u_1-u_2\rangle \overset{\text{CSU}}{\le} \|f_1-f_2\|\,\|u_1-u_2\|.$$

Nach Dividieren durch $\|u_1-u_2\|$ folgt die Behauptung.

∎

Korollar 6.4. Sei H ein Hilbertraum und $M \subset H$ ein abgeschlossener Unteraum. Für $f \in H$ ist $\mathrm{P}_M f =: u$ charakterisiert durch: $u \in M$ und $\langle f-u, v\rangle = 0$ für alle $v \in M$. Insbesondere ist P_M ein linearer stetiger Operator.

Beweis. Es gilt $\mathrm{Re}\langle f-u, v-u\rangle \leq 0$ für alle $v \in M$. Sei $w \in M$. Dann gilt auch $v = u + \langle f-u, w\rangle\, w \in M$ und damit

$$|\langle f-u, w\rangle|^2 = \mathrm{Re}\big(\overline{\langle f-u, w\rangle}\langle f-u, w\rangle\big) \leq 0,$$

woraus wie gewünscht

$$\langle f-u, w\rangle = 0$$

folgt. Gilt andererseits $\langle f-u, w\rangle = 0$ für alle $w \in M$, so folgt für alle $v \in M$ schon

$$\mathrm{Re}\langle f-v, v-u\rangle \leq 0,$$

da $v - u \in M$. Die Linearität ist nach der obigen Charakterisierung klar und die Stetigkeit folgt aus Satz 6.3.

■

6.5 (Dualraum eines Hilbertraums)**.** In einem Hilbertraum H können wir durch

$$H \ni u \mapsto \langle u, f\rangle \in \mathbb{K}$$

für jedes $f \in H$ ein lineares Funktional definieren. Der folgende Satz zeigt, dass wir dadurch sogar bereits alle linearen Funktionale erhalten.

Satz: (Riesz'scher Darstellungssatz)
Sei H ein Hilbertraum über $\mathbb{K}$. Dann ist

$$\begin{aligned} J\colon H &\to H' \\ x &\mapsto \begin{pmatrix} H \to \mathbb{K} \\ \quad y \mapsto \langle y, x\rangle \end{pmatrix} \end{aligned}$$

ein isometrischer, konjugiert linearer Isomorphismus. (Dabei bedeutet konjugiert linear, dass $J(\alpha x + y) = \bar{\alpha} J(x) + J(y)$ für alle $x, y \in H$, $\alpha \in \mathbb{K}$ gilt.)

Beweis. Seien $x, y \in H$. Aus der Cauchy-Schwarz-Ungleichung (Satz 2.8 (1)) folgt

$$|J(x)(y)| \leq \|x\|\,\|y\|.$$

Daraus folgt $J(x) \in X'$ mit $\|J(x)\| \leq \|x\|$. Wegen $|J(x)(x)| = \|x\|^2$ gilt $\|J(x)\| \geq \|x\|$. Also ist J eine Isometrie und damit insbesondere injektiv. Der wesentliche Schritt ist nun, die Surjektivität von J zu zeigen. Sei dazu $x_0' \in X' \backslash \{0\}$. Wähle P als orthogonale Projektion auf $N(x_0')$ (was ein abgeschlossener Unterraum ist). Wähle $e \in X$ mit $x_0'(e) = 1$ und definiere $x_0 \coloneqq e - Pe$. Dann gilt $x_0'(x_0) = 1 \neq 0$. Aus dem Projektionssatz (6.1) folgt:

$$\forall\, y \in N(x_0')\colon \quad \langle y, x_0\rangle = 0. \tag{$\star$}$$

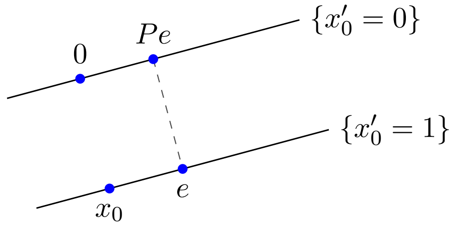


Abbildung 6.1: Situation im Beweis des Riesz'schen Darstellungssatzes

Sei wieder $x \in X$. Dann gilt

$$x = \underbrace{(x - x_0'(x)\,x_0)}_{\in N(x_0')} + x_0'(x)\,x_0$$

und damit wegen $(\star)$:

$$\langle x, x_0 \rangle = \big\langle x_0'(x)\,x_0, x_0 \big\rangle = x_0'(x)\,\|x_0\|^2.$$

Es folgt

$$x_0'(x) = \left\langle x, \frac{x_0}{\|x_0\|^2} \right\rangle = J\left(\frac{x_0}{\|x_0\|^2}\right)(x)$$

und daraus $x' = J\big(x_0/\|x_0\|^2\big)$. Also ist J surjektiv.

■

6.6 (Soll man H mit H' identifizieren?)**.** Der Rieszsche Darstellungssatz erlaubt es uns, H mit H' zu identifizieren. Wir werden dies oft tun, aber nicht immer. Wir wollen eine Situation betrachten, in der man vorsichtig mit einer solchen Identifikation sein sollte:

Sei H ein Hilbertraum mit Skalarprodukt $\langle \cdot\,, \cdot \rangle$ und assoziierter Norm $\|\cdot\|$. Sei nun $V \subset H$ ein linearer Unterraum, der dicht in H liegt. Wir nehmen an, dass V ein Banachraum ist mit Norm $\|\cdot\|_V$. Weiter sei die Inklusion $V \hookrightarrow H$ stetig, d. h. es existiert ein $c \in \mathbb{R}_{>0}$, so dass für alle $v \in V$ gilt: $\|v\|_V \leq c\|v\|$.

Ein Beispiel für eine solche Situation ist gegeben durch

$$\begin{aligned} H &= L^2([0,1]) \\ &= \Big\{v\colon [0,1] \to \mathbb{R} \text{ Lebesgue-messbar} \;\Big|\; \textstyle\int_0^1 (v(x))^2 \,\mathrm{d}x < \infty\Big\} \Big/ \{v \mid v = 0 \text{ fast überall}\} \end{aligned}$$

mit Skalarprodukt $\langle u, v \rangle = \int_0^1 u(x)\,v(x)\,\mathrm{d}x$ für $u, v \in H$ und $V = C([0,1])$. Dann existiert eine kanonische Abbildung

$$T\colon H' \to V', \quad x' \mapsto \big(v \mapsto x'(v)\big).$$

Wir sehen einfach ein, dass T stetig und injektiv ist. Nun identifizieren wir H' mit H, geschrieben $H \simeq H'$, und erhalten: $V \subset H \simeq H' \subset V'$ $(\diamond)$, wobei dies alles stetige Injektionen sind. Problematisch wird diese Situation, falls V ein Hilbertraum mit eigenem Skalarprodukt $\langle \cdot\,, \cdot \rangle_V$ und $\|\cdot\|_V$ die zugehörige Norm ist. Wir könnten V mit V' identifizieren. Aber was bedeutet dann $(\diamond)$? Wir können also nicht gleichzeitig H mit H' und V mit V' identifizieren. Typischerweise verwendet man dann nur $H \simeq H'$.

Wir betrachten noch ein weiteres Beispiel:

$$H = \ell^2(\mathbb{R}) \quad \text{mit Skalarprodukt} \quad (u, v) \mapsto \langle u, v\rangle := \sum_{n=1}^{\infty} u_n v_n.$$

Weiter sei

$$V = \Big\{ u \in H \;\Big|\; \sum_{n=1}^{\infty} n^2 u_n^2 < \infty \Big\},$$

ausgestattet mit dem Skalarprodukt

$$\langle\!\langle \,\cdot\,, \,\cdot\, \rangle\!\rangle \colon V \times V \to \mathbb{R}, \quad (u, v) \mapsto \langle\!\langle u, v \rangle\!\rangle := \sum_{n=1}^{\infty} n^2 u_n v_n.$$

Es gilt $V \subset H$ und die Inklusion $V \hookrightarrow H$ ist stetig. Wir identifizieren H' mit H und V' mit dem Raum

$$V' = \Big\{ f \in \mathbb{R}^{\mathbb{N}} \;\Big|\; \sum_{n=1}^{\infty} \frac{1}{n^2} f_n^2 < \infty \Big\},$$

indem wir für $f \in V'$, $v \in V$ die Anwendung von f auf v durch $f(v) = \sum_{n=1}^{\infty} f_n v_n$ erklären (man rechnet leicht mit der Hölderschen Ungleichung nach, dass dies wohldefiniert ist). Dieser Raum ist offenbar echt größer als H. Die Isometrie $J \colon V \to V'$ aus dem Riesz'schen Darstellungssatz (6.5) ist dann gegeben durch

$$u \mapsto \big(n^2 u_n\big)_{n\in\mathbb{N}}.$$

Bemerkung 6.7. Sei H ein Hilbertraum über $\mathbb{K}$. Dann ist H reflexiv (Definition 4.19). Sei J der konjugiert lineare Isomorphismus aus dem Riesz'schen Darstellungssatz (6.5). Es ist zu zeigen, dass J_H (aus Satz 4.18) surjektiv ist. Sei also $x'' \in H''$. Dann ist

$$x' \colon H \to \mathbb{K}, \quad y \mapsto \overline{x''(Jy)}$$

ein Element in H'. Setze $x := J^{-1}x'$. Sei $y' \in H'$ und $y \in H$ mit $Jy = y'$. Dann gilt:

$$\begin{aligned} x''(y') &= x''(Jy) = \overline{x'(y)} = \overline{(Jx)(y)} = \overline{\langle y, x\rangle} \\ &= \langle x, y\rangle = (Jy)(x) = y'(x) = (J_H x)(y') \end{aligned}$$

Also gilt $x'' = J_H x$ und damit ist J_H surjektiv.

Bemerkung 6.8. Sei $H \simeq H'$ ein Hilbertraum und $M \subset H$ ein Unterraum. Dann kann man den Annihilator $M^\perp$ identifizieren mit

$$M^\perp = \big\{ u \in H \;\big|\; \forall v \in M \colon\; \langle v, u\rangle = 0 \big\}.$$

Außerdem gilt $M \cap M^\perp = \{0\}$. Wenn M abgeschlossen ist, so gilt $M + M^\perp = H$.

Definition 6.9. Sei H ein Hilbertraum. Wir nennen eine Sesquilinearform a auf H stetig, wenn es ein $C \in \mathbb{R}_{>0}$ gibt, so dass gilt:

$$\forall\, u, v \in H\colon \quad |a(u,v)| \leq C\, \|u\|\, \|v\|.$$

Wir nennen a *koerziv*, wenn ein $\alpha \in \mathbb{R}_{>0}$ existiert, so dass gilt:

$$\forall\, u \in H\colon \quad \operatorname{Re} a(u,u) \geq \alpha \|u\|^2.$$

Für den Beweis des folgenden Theorems benötigen wir den Banach'schen Fixpunktsatz:
Satz: Sei (X,d) ein vollständiger metrischer Raum und $S\colon X \to X$ eine (strikte) Kontraktion. Dann besitzt S genau einen Fixpunkt.

Theorem 6.10 (Stampacchia)**.** Sei H ein Hilbertraum über $\mathbb{K}$ und $a\colon H \times H \to \mathbb{K}$ eine stetige, koerzive Sesquilinearform. Weiter sei $K \subset H$ nicht leer, abgeschlossen und konvex. Dann gibt es für alle $\varphi \in H'$ genau ein $u \in K$, so dass gilt:

$$\forall\, v \in K\colon \quad \operatorname{Re} a(v-u, u) \geq \operatorname{Re} \varphi(v-u). \tag{$\star$}$$

Ist a außerdem symmetrisch, so ist das Element $u \in H$ eindeutig charakterisiert durch folgende Eigenschaft:

$$u \in K \quad \text{und} \quad \frac{1}{2}\, a(u,u) - \operatorname{Re} \varphi(u) = \min_{v \in K} \big(\tfrac{1}{2}\, a(v,v) - \operatorname{Re} \varphi(v) \big).$$

Beweis. Sei $\varphi \in H'$ gegeben. Riez (6.5) liefert: es gibt genau ein $f \in H$, so dass für alle $v \in H$ schon $\varphi(v) = \langle v, f \rangle$ gilt. Andererseits ist für $u \in H$ die Abbildung

$$v \mapsto a(v,u)$$

ein Element in H'. Also erhalten wir erneut mit Riesz: Es existiert genau ein $Au \in H$, so dass für alle $v \in H$ gilt:

$$a(v,u) = \langle v, Au \rangle.$$

Seien $C, \alpha \in \mathbb{R}_{>0}$ entsprechende Konstanten für a wie in Definition 6.9. Dann gilt:

$$\|Au\| \leq \|a(\,\cdot\,, u)\| \leq C\, \|u\|.$$

Da a und $\langle\,\cdot\,,\,\cdot\,\rangle$ im zweiten Argument konjugiert linear sind, folgt, dass A linear ist. Somit ergibt sich:

$$A \in L(H) \quad \text{und} \quad \|A\| \leq C.$$

Außerdem gilt für alle $u \in H$:

$$\operatorname{Re}\langle u, Au \rangle = \operatorname{Re} a(u,u) \geq \alpha \|u\|^2.$$

Die Eigenschaft $(\star)$ aus der Behauptung ist äquivalent zu:

$$\forall\, v \in K\colon \quad \operatorname{Re}\langle v-u, Au \rangle \geq \operatorname{Re}\langle v-u, f \rangle.$$

Für $\varrho \in \mathbb{R}_{>0}$ ist dies wiederum äquivalent zu:

$$\forall\, v \in K\colon \quad \operatorname{Re}\langle v-u, \varrho f - \varrho A u + u - u\rangle \leq 0.$$

Aus dem Projektionssatz (6.1) folgt, dass dies äquivalent dazu ist, ein $u \in K$ zu finden mit

$$u = \mathrm{P}_K(\varrho f - \varrho A u + u).$$

Definiere für $v \in K$:

$$S(v) := \mathrm{P}_K(\varrho f - \varrho A v + v).$$

Satz 6.3 liefert: für alle $v_1, v_2 \in K$ gilt:

$$\|Sv_1 - Sv_2\| \leq \|(v_1 - v_2) - \varrho(Av_1 - Av_2)\|.$$

Es folgt:

$$\begin{aligned}\|Sv_1 - Sv_2\|^2 &\leq \|v_1 - v_2\|^2 - 2\varrho \underbrace{\operatorname{Re}\langle Av_1 - Av_2, v_1 - v_2\rangle}_{\geq \alpha \|v_1 - v_2\|^2} + \varrho^2 \|Av_1 - Av_2\|^2 \\ &\leq \|v_1 - v_2\|^2 \underbrace{\left(1 - 2\varrho\alpha + \varrho^2 C^2\right)}_{<1 \text{ für } \varrho \in (0, 2\alpha/C^2)}\end{aligned}$$

Für ϱ klein genug ist S also eine Kontraktion und damit liefert der Banach'sche Fixpunktsatz die Existenz eines eindeutigen Fixpunkts. Nach den vorherigen Überlegungen ist dieser dann die eindeutige Lösung von $(\star)$.

Sei nun a symmetrisch. Dann definiert $(u,v) \mapsto a(u,v)$ ein Skalarprodukt auf H. Die zugehörige Norm $u \mapsto \sqrt{a(u,u)}$ ist äquivalent zur Norm $u \mapsto \|u\| = \sqrt{\langle u, u\rangle}$ (was man leicht aus den Voraussetzungen an a folgern kann). Also ist H auch ein Hilbertraum bezüglich des von a gegebenen Skalarprodukts. Der Riesz'sche Darstellungssatz (6.5) liefert: Für alle $\varphi \in H'$ gibt es ein eindeutiges $g \in H$ mit $\varphi = a(\,\cdot\,, g)$. Damit ist $(\star)$ aber äquivalent dazu, ein $u \in K$ zu finden, welches folgende Eigenschaft erfüllt:

$$\forall\, v \in K\colon \quad \operatorname{Re} a(v-u, g-u) \leq 0.$$

Das heißt aber, dass u die Projektion von g auf K bezüglich des durch a gegebenen Skalarprodukts ist. Der Projektionssatz (6.1) liefert dann, dass dies äquivalent ist zu

$$u \in K \quad \text{und} \quad \sqrt{a(g-u, g-u)} = \min_{v \in K} \sqrt{a(g-v, g-v)},$$

was offenbar genau dann der Fall ist, wenn u die Funktion

$$v \mapsto a(g-v, g-v) = a(v,v) - 2\operatorname{Re} a(v,g) + a(g,g) = a(v,v) - 2\operatorname{Re}\varphi(v) + a(g,g)$$

auf K minimiert, oder äquivalent die Funktion

$$v \mapsto \frac{1}{2}\, a(v,v) - \operatorname{Re}\varphi(v).$$

■

Bemerkung:
Ist a eine positiv semidefinite Bilinearform auf einem $\mathbb{R}$-Vektorraum H, so ist die Abbildung $v \mapsto a(v,v)$ konvex. Falls a sogar positiv definit ist, ist diese Abbildung strikt konvex. In der Regel besitzen strikt konvexe Funktionen mit $f(x) \to \infty$ für $\|x\| \to \infty$ ein eindeutiges Minimum.

Satz 6.11 (Lax-Milgram)**.** Sei H ein Hilbertraum über $\mathbb{K}$ und $a\colon H \times H \to \mathbb{K}$ eine stetige, koerzive Sesquilinearform. Dann gibt es für jedes $\varphi \in H'$ genau ein $u \in H$, so dass gilt:

$$\forall\, v \in H\colon \qquad \operatorname{Re} a(v,u) = \operatorname{Re} \varphi(v). \tag{$\star\star$}$$

Ist a außerdem symmetrisch, so ist das Element $u \in H$ eindeutig charakterisiert durch folgende Eigenschaft:

$$\frac{1}{2}\, a(u,u) - \operatorname{Re} \varphi(u) = \min_{v \in H} \big(\tfrac{1}{2}\, a(v,v) - \operatorname{Re} \varphi(v)\big).$$

Beweis. Sei $\varphi \in H'$ und $w \in H$. Wählen wir $K = H$ im Theorem von Stampacchia (6.10), so erhalten wir ein $u \in H$ mit folgender Eigenschaft:

$$\forall\, v \in H\colon \qquad \operatorname{Re} a(v-u,u) \geq \operatorname{Re} \varphi(v-u).$$

Also erhalten wir für $\pm w + u \in H$ zwei Ungleichungen, die zusammen

$$\operatorname{Re} a(w,u) = \operatorname{Re} \varphi(w)$$

implizieren. Es folgt die Behauptung.

■

Bemerkung 6.12.

(i) Der Satz von Lax-Milgram (6.11) kann zur Lösung elliptischer partieller Differentialgleichungen genutzt werden.

(ii) Sei a symmetrisch und sei $F\colon H \to \mathbb{R}$, $v \mapsto \frac{1}{2} a(v,v) - \operatorname{Re} \varphi(v)$. Dann bedeutet ($\star\star$), dass für das Minimum „$F'(u) = 0$“ gilt. Betrachte dazu $\frac{\mathrm{d}}{\mathrm{d}t} F(u+tv)|_{t=0}$.

Es gilt

$$F(u+tv) = \frac{1}{2}\big(a(u,u) + ta(u,v) + ta(v,u) + t^2 a(v,v)\big) - \operatorname{Re} \varphi(u) - t \operatorname{Re} \varphi(v),$$

also erhalten wir:

$$\frac{\mathrm{d}}{\mathrm{d}t} F(u+tv) = \operatorname{Re} a(u,v) + ta(v,v) - \operatorname{Re} \varphi(v).$$

Definition 6.13. Sei H ein Hilbertraum und $(E_n)_{n\in\mathbb{N}}$ eine Folge von abgeschlossenen Unterräumen von H. Dann nennen wir H die *Hilbertsumme von* $(E_n)_{n\in\mathbb{N}}$, falls

(a) $(E_n)_{n\in\mathbb{N}}$ paarweise orthogonal ist, d. h. $\langle u,v\rangle = 0$ für alle $u \in E_n$, $v \in E_m$ mit $n \neq m$, und

(b) der lineare Raum $\operatorname{span}\big(\bigcup_{n=1}^{\infty} E_n\big)$ ist dicht in H.

In diesem Fall schreiben wir

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

Satz 6.14. Sei H ein Hilbertraum und gelte $H = \bigoplus_{n=1}^{\infty} E_n$ für eine Folge abgeschlossener Unterräume $(E_n)_{n\in\mathbb{N}}$. Sei $u \in H$ und für alle $n \in \mathbb{N}$ sei $u_n \coloneqq \mathrm{P}_{E_n}(u)$ sowie $S_n \coloneqq \sum_{k=1}^{n} u_k$. Dann gilt

$$\lim_{n\to\infty} S_n = u$$

(wofür wir auch die Notation $\sum_{k=1}^{\infty} u_k = u$ verwenden) und

$$\sum_{k=1}^{\infty} \|u_k\|^2 = \|u\|^2,$$

die sogenannte *Bessel-Parseval-Identität.*

Für den Beweis benötigen wir zunächst eine Hilfsaussage.

Lemma 6.15. Sei H ein Hilbertraum und $(v_n)_{n\in\mathbb{N}}$ ein Folge paarweise orthogonaler Vektoren in H (d. h. für alle $n, m \in \mathbb{N}$ mit $n \neq m$ gilt $\langle v_n, v_m\rangle = 0$). Gelte außerdem $\sum_{k=1}^{\infty} \|v_k\|^2 < \infty$. Sei $S_n \coloneqq \sum_{k=1}^{n} v_k$. Dann existiert der Grenzwert $S \coloneqq \lim_{n\to\infty} S_n$ und es gilt

$$\|S\|^2 = \sum_{n=1}^{\infty} \|v_k\|^2.$$

Beweis. Seien $m, n \in \mathbb{N}$ mit $m > n > 1$. Dann gilt:

$$\begin{aligned}
\|S_m - S_{n-1}\|^2 &= \Big\|\sum_{k=n}^{m} v_k\Big\|^2 = \Big\langle \sum_{k=n}^{m} v_k,\ \sum_{\ell=n}^{m} v_\ell \Big\rangle \\
&= \sum_{k=n}^{m}\sum_{\ell=n}^{m} \langle v_k, v_\ell\rangle = \sum_{k=n}^{m} \|v_k\|^2
\end{aligned}$$

Dies zeigt, dass $(S_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge ist. Da H vollständig ist, existiert also auch ein Grenzwert $S \coloneqq \lim_{n\to\infty} S_n$. Außerdem folgt aus der obigen Rechnung

$$\|S_n\|^2 = \sum_{k=1}^{n} \|v_k\|^2,$$

woraus für $n \to \infty$ die zweite Behauptung folgt.

■

Beweis von Satz 6.14. Aus Korollar 6.4 folgt:

$$\forall\, n \in \mathbb{N}\ \forall\, v \in E_n\colon \quad \langle u - u_n, v\rangle = 0.$$

Insbesondere erhalten wir daraus $\langle u, u_n\rangle = \|u_n\|^2$ für alle $n \in \mathbb{N}$. Sei $m \in \mathbb{N}$. Dann gilt also:

$$\langle u, S_m\rangle = \sum_{n=1}^{m} \|u_n\|^2 = \|S_m\|^2.$$

(Für die zweite Gleichheit, vergleiche Beweis von Lemma 6.15.) Die Cauchy-Schwarz-Ungleichung (Satz 2.8 (1)) liefert

$$\|S_m\|^2 \leq \|u\| \, \|S_m\|,$$

woraus $\|S_m\| \leq \|u\|$ folgt. Dies zeigt:

$$\sum_{n=1}^{m} \|u_n\|^2 = \|S_m\|^2 \leq \|u\|^2.$$

Damit konvergiert $\sum_{n=1}^{\infty} \|u_n\|^2$ und wir können Lemma 6.15 anwenden. Wir erhalten die Existenz des Grenzwerts $S \coloneqq \lim_{n\to\infty} S_n$. Wir wollen nun S bestimmen (zunächst ohne die Dichtheit von $\operatorname{span}\big(\bigcup_{n=1}^{\infty} E_n\big)$ vorauszusetzen). Sei $F \coloneqq \operatorname{span}\big(\bigcup_{n=1}^{\infty} E_n\big)$. Wir behaupten

$$S = \mathrm{P}_{\overline{F}}\, u. \tag{$**$}$$

Sei $m \in \mathbb{N},\ v \in E_m$ und sei $n \in \mathbb{N}$ mit $m \leq n$. Dann gilt:

$$\begin{aligned} \langle u - S_n, v\rangle &= \Big\langle u - \sum_{k=1}^{n} u_k, v \Big\rangle = \langle u, v\rangle - \sum_{k=1}^{n} \underbrace{\langle u_k, v\rangle}_{=0 \text{ für } k \neq m} \\ &= \langle u, v\rangle - \langle u_m, v\rangle = \langle u - u_m, v\rangle = 0. \end{aligned}$$

Für $n \to \infty$ folgt mit der Stetigkeit des Skalarprodukts $\langle u - S, v\rangle = 0$. Es folgt $\langle u - S, \tilde{v}\rangle = 0$ für alle $\tilde{v} \in F$. Dies impliziert (erneut wegen der Stetigkeit des Skalarprodukts)

$$\forall\, \tilde{v} \in \overline{F}\colon \quad \langle u - S, \tilde{v}\rangle = 0.$$

Wegen $S_n \in F$ für alle $n \in \mathbb{N}$, gilt $S \in \overline{F}$. Mithilfe von (6.4) folgt nun ($**$). Wir benutzen nun zusätzlich, dass F dicht in H liegt, also $\overline{F} = H$. Dann ergibt ($**$) direkt $S = u$. Gehen wir in $\sum_{n=1}^{m} \|u_n\|^2 = \|S_m\|^2$ zum Grenzwert $m \to \infty$ über, so folgt die Bessel-Parseval-Identität.

■

Definition 6.16 (Schauder-Basis)**.** Sei X ein normierter Raum und $(e_k)_{k\in\mathbb{N}}$ eine Folge in X. Dann nennen wir $\{e_k \mid k \in \mathbb{N}\}$ eine Schauder-Basis von X, falls gilt: Für alle $x \in X$ existiert eine eindeutige Folge $(\alpha_k)_{k\in\mathbb{N}}$ in $\mathbb{K}$ mit

$$\sum_{k=1}^{n} \alpha_k\, e_k \to x \qquad \text{für } n \to \infty.$$

Definition: (Orthonormalbasis) Sei H ein Prä-Hilbertraum und $(e_k)_{k\in\mathbb{N}}$ eine Folge in H. Dann nennen wir $(e_k)_{k\in\mathbb{N}}$ eine *Orthonormalbasis* (oder *Hilbertbasis*), falls $(e_k)_{k\in\mathbb{N}}$ eine Schauder-Basis ist und $\langle e_m, e_n\rangle = \delta_{mn}$ für alle $m, n \in \mathbb{N}$ gilt (mit dem Kroneckerdelta δ).

Satz 6.17. Sei H ein Prä-Hilbertraum und $(e_k)_{k\in\mathbb{N}}$ ein Orthonormalsystem, d. h. für alle $m, n \in \mathbb{N}$ gilt $\langle e_m, e_n\rangle = \delta_{mn}$. Dann sind äquivalent:

(1) $\operatorname{span}\{e_k \mid k \in \mathbb{N}\}$ liegt dicht in H

(2) $(e_k)_{k\in\mathbb{N}}$ ist eine Schauder-Basis

(3) $\forall\, x \in H\colon \quad x = \sum_{k=1}^{\infty} \langle x, e_k\rangle\, e_k$

(4) $\forall\, x, y \in H\colon \quad \langle x, y\rangle = \sum_{k=1}^{\infty} \langle x, e_k\rangle \overline{\langle y, e_k\rangle}$ (Parseval-Identität)

(5) $\forall\, x \in H\colon \quad \|x\|^2 = \sum_{k=1}^{\infty} |\langle x, e_k\rangle|^2$ (Vollständigkeitsrelation)

Falls eine dieser Bedinungen gilt, ist $(e_k)_{k\in\mathbb{N}}$ also eine Hilbertbasis.

Beweis. „(1) ⇒ (3)“: Nutze Satz 6.14 (in den dortigen Bezeichnern) mit $E_n = \operatorname{span}\{e_n\}$. Es gilt dann $u_n = \alpha_n e_n$ mit $\alpha_n = \langle e_n, u\rangle$. Das heißt:

$$S_n = \sum_{k=1}^{n} \langle e_k, u\rangle\, e_k$$

(folgt aus der Orthogonalität der Projektion). Also folgt $S_n \to u$ für $n \to \infty$. Aus 6.14 folgt dann die Behauptung.

„(3) ⇒ (2)“: Wir müssen die Eindeutigkeit der Koeffizienten zeigen. Wegen der Linearität reicht es, den Fall $x = 0$ zu betrachten. Gilt

$$0 = \sum_{k=1}^{\infty} \alpha_k e_k,$$

so folgt mit der Stetigkeit des Skalarprodukts:

$$0 = \Big\langle \sum_{k=1}^{\infty} \alpha_k e_k, e_\ell \Big\rangle = \sum_{k=1}^{\infty} \alpha_k \langle e_k, e_\ell\rangle = \alpha_\ell.$$

„(2) ⇒ (1)“ folgt aus der Definition der Schauder-Basis.

„(3) ⇒ (4)“: Aus der Stetigkeit des Skalarprodukts erhalten wir:

$$\langle x, y\rangle = \lim_{n\to\infty} \Big\langle \sum_{k=1}^{n} \langle x, e_k\rangle e_k, \sum_{\ell=1}^{n} \langle y, e_\ell\rangle e_\ell \Big\rangle = \lim_{n\to\infty} \sum_{k=1}^{n} \sum_{\ell=1}^{n} \langle x, e_k\rangle \overline{\langle y, e_\ell\rangle}\, \underbrace{\langle e_k, e_\ell\rangle}_{\delta_{k\ell}}.$$

„(4) ⇒ (5)“ ist klar.

„(5) ⇒ (3)“:

$$\begin{aligned} \Big\| x - \sum_{k=1}^{n} \langle x, e_k\rangle\, e_k \Big\|^2 &\leq \|x\|^2 - \sum_{k=1}^{n} \langle x, e_k\rangle \overline{\langle x, e_k\rangle} - \sum_{k=1}^{n} \langle x, e_k\rangle \langle e_k, x\rangle + \sum_{k=1}^{n} |\langle x, e_k\rangle|^2 \\ &= \|x\|^2 - \sum_{k=1}^{n} |\langle x, e_k\rangle|^2 \quad \to 0 \qquad \text{für } n \to \infty \end{aligned}$$

■

Definition 6.18 (separabel). Ein topologischer Raum X heißt *separabel*, falls X eine abzählbare dichte Teilmenge enthält.

Beispiele:

(a) $\mathbb{R}$ ist separabel, da $\mathbb{Q} \subset \mathbb{R}$ dicht liegt.

(b) $C^0([a,b])$ mit der Supremumsnorm ist separabel, da die Polynome mit rationalen Koeffizienten dicht liegen. (Weierstraßscher Approximationssatz)

(c) $\ell^2(\mathbb{R})$ ist separabel, da die abbrechenden Folgen aus $\ell^2(\mathbb{Q})$ dicht liegt (vgl. Beweis von Lemma 6.19).

Lemma 6.19. Sei X ein unendlich-dimensionaler normierter Raum. Dann sind äquvialent:

(1) X ist separabel

(2) Es gibt eine Folge $(X_n)_{n\in\mathbb{N}}$ endlich-dimensionaler Unterräume von X mit folgenden Eigenschaften: Für alle $n \in \mathbb{N}$ gilt $X_n \subset X_{n+1}$ und $\bigcup_{n\in\mathbb{N}} X_n$ liegt dicht in X.

(3) Es gibt eine Folge $(E_n)_{n\in\mathbb{N}}$ endlich-dimensionaler Unterräumen von X mit folgenden Eigenschaften: Für alle $n, m \in \mathbb{N}$ mit $n \neq m$ gilt $E_n \cap E_m = \{0\}$ und

$$\bigoplus_{k\in\mathbb{N}} E_k := \bigcup_{k\in\mathbb{N}} (E_1 \oplus \cdots \oplus E_k)$$

liegt dicht in X.

(4) Es gibt eine linear unabhängige Menge $\{e_k \mid k \in \mathbb{N}\}$ von Vektoren aus X, so dass $\operatorname{span}\{e_k \mid k \in \mathbb{N}\}$ dicht in X liegt.

Beweis. „(1)$\Rightarrow$(2)“: Sei $\{x_n \mid n \in \mathbb{N}\}$ dicht in X. Definiere $X_n := \operatorname{span}\{x_1, \ldots, x_n\}$ für alle $n \in \mathbb{N}$, dann erfüllt die Folge $(X_n)_{n\in\mathbb{N}}$ die geforderten Bedingungen.

„(2)$\Rightarrow$(3)“: Sei $E_1 := X_1$ und $n \in \mathbb{N}$. Da X_{n+1} endlich-dimensional ist, gibt es einen Teilraum $E_{n+1} \subset X_{n+1}$ mit $X_{n+1} = X_n \oplus E_{n+1}$. Wir erhalten so eine Folge von Unterräumen $(E_n)_{n\in\mathbb{N}}$. Dann gilt $X_n = E_1 \oplus \cdots \oplus E_n$ und nach Voraussetzung liegt aber $\bigcup_{n\in\mathbb{N}} X_n$ dicht in X.

„(3)$\Rightarrow$(4)“: Für alle $n \in \mathbb{N}$ sei $(e_{n,j})_{j\in\{1,\ldots,\dim E_n\}}$ eine Basis von E_n. Setzte

$$X_n := E_1 \oplus \cdots \oplus E_n = \operatorname{span}\{e_{i,j} \mid 1 \leq i \leq n,\ 1 \leq j \leq \dim E_i\}.$$

Dann gilt

$$\operatorname{span}\{e_{i,j} \mid i \in \mathbb{N},\ 1 \leq j \leq \dim E_i\} = \operatorname{span}\Big(\bigcup_{n\in\mathbb{N}} X_n\Big)$$

und nach Voraussetzung liegt die rechte Menge dicht in X. Also ist

$$\{e_{i,j} \mid i \in \mathbb{N},\ 1 \leq j \leq \dim E_i\}$$

die gesuchte linear unabhängige Menge.

„(4) ⇒ (1)": Für $n \in \mathbb{N}$ ist

$$A_n := \Big\{ \sum_{k=1}^{n} \alpha_k\, e_k \;\Big|\; \forall\, k \in \{1, \dots, n\}\colon\; \alpha_k \in \mathbb{K} \cap \mathbb{Q}(i) \Big\}$$

abzählbar (wobei $\mathbb{Q}(i) = \{x + iy \in \mathbb{C} \mid x, y \in \mathbb{Q}\}$) mit

$$\overline{A}_n = \operatorname{span}\{e_k \mid 1 \le k \le n\}.$$

Da abzählbare Vereinigungen abzählbarer Mengen wieder abzählbar sind, liefert die Teilmenge $\bigcup_{n\in\mathbb{N}} A_n$ von X die Behauptung.

■

Satz 6.20. Für jeden unendlich-dimensionalen Hilbertraum H über $\mathbb{K}$ ist äquivalent:

(1) H ist separabel

(2) H besitzt eine Hilbertbasis

Weiter gilt: Ist eine dieser Bedingungen erfüllt, so ist H isometrisch isomorph zu $\ell^2(\mathbb{K})$.

Beweis. „(1) ⇒ (2)": Sei $(e_k)_{k\in\mathbb{N}}$ wie in Lemma 6.19 (4) und für alle $n \in \mathbb{N}$ sei $H_n := \operatorname{span}\{e_1, \dots, e_n\}$. Definiere induktiv für alle $n \in \mathbb{N}$ Elemente $\hat{e}_n$ durch

$$\begin{aligned} \tilde{e}_n &:= e_n - \sum_{k=1}^{n-1} \langle e_n, \tilde{e}_k\rangle\, \hat{e}_k \in H_n \setminus H_{n-1} \\ \hat{e}_n &:= \frac{\tilde{e}_n}{\|\tilde{e}_n\|}. \end{aligned}$$

Dann gilt $\hat{e}_n \in H_n \cap H_{n-1}^{\perp}$ für alle $n \in \mathbb{N}$. Also ist $(\hat{e}_n)_{n\in\mathbb{N}}$ ein Orthonormalsystem. Wegen $\operatorname{span}\{\hat{e}_k \mid 1 \le k \le n\} = H_n$ folgt:

$$\operatorname{span}\{\hat{e}_k \mid k \in \mathbb{N}\} \text{ liegt dich in } H.$$

Nach Satz 6.17 ist damit $\{\hat{e}_k \mid k \in \mathbb{N}\}$ eine Hilbertbasis und es gilt $x = \sum_{k=1}^{\infty} \langle x, \hat{e}_k\rangle\, \hat{e}_k$ für alle $x \in H$. Die Abbildung

$$J\colon H \to \ell^2(\mathbb{K}), \quad x \mapsto \big(\langle x, \hat{e}_k\rangle\big)_{k\in\mathbb{N}}$$

erfüllt dann $\langle Jx, Jy\rangle = \langle x, y\rangle$ für alle $x, y \in H$ wieder nach Satz 6.17. Außerdem ist J linear und bijektiv.

„(2) ⇒ (1)" folgt aus Lemma 6.19, da $\{e_k \mid k \in \mathbb{N}\}$ existiert mit $\operatorname{span}\{e_k \mid k \in \mathbb{N}\}$ dicht in H.

■

Bemerkung: Auf $\ell^2(\mathbb{K})$ ist $(e_n)_{n\in\mathbb{N}} := \big((\delta_{kn})_{k\in\mathbb{N}}\big)_{n\in\mathbb{N}}$ eine Hilbertbasis. (Dabei ist δ das Kroneckerdelta, also hat für $n \in \mathbb{N}$ der Vektor e_n nur in der n-ten Komponente den Eintrag 1, ansonsten 0.)

Es bildet $(e_n)_{n\in\mathbb{N}}$ ein Orthonormalsystem. Außerdem liegt $\operatorname{span}\{e_k \mid k \in \mathbb{N}\}$ dicht in ℓ^2. Sei $x \in \ell^2$ und sei

$$x^{(k)} := (x_1, \dots, x_k, 0, \dots) = \sum_{i=1}^{k} x_i\, e_i.$$

Dann gilt:

$$\|x^{(k)} - x\|^2 = \sum_{i=k+1}^{\infty} x_i^2 \xrightarrow[k\to\infty]{} 0.$$

Bemerkung: Die Tatsache, dass alle separabelen Hilberträume isometrisch isomorph zu ℓ^2 sind, könnte dazu verführen, nur noch ℓ^2 zu betrachten. Viele Operatoren zwischen separablen Hilberträumen haben aber eine komplizierte Struktur, wenn man sie in ℓ^2 ausdrückt. Fazit: Oft ist es besser, doch direkt mit dem entsprechenden Hilbertraum zu arbeiten.

7 Schwache Konvergenz

Zur Erinnerung: Im $\mathbb{R}^n$ kann man aus jeder beschränkten Folge eine konvergente Teilfolge auswählen. Außerdem gilt für Teilmengen $K \subset \mathbb{R}^n$ der *Satz von Heine-Borel*: K ist genau dann kompakt, wenn K beschränkt und abgeschlossen ist. Beides ist in unendlichdimensionalen Banachräumen i. A. nicht mehr gegeben. Wir versuchen, Ersatz dafür zu finden.

Definition 7.1. (i) Sei $(X, \mathcal{T})$ ein topologischer Raum und $A \subset X$. Dann heißt A *kompakt*, falls jede offene Überdeckung von A eine endliche Teilüberdeckung besitzt, d. h. falls A folgende Eigenschaft besitzt: Ist I eine Menge und $(U_i)_{i\in I}$ eine Familie offener Mengen in X mit $A \subset \bigcup_{i\in I} U_i$, so gibt es eine eine endliche Teilmenge $J \subset I$ mit $A \subset \bigcup_{j\in J} U_j$.

(ii) Sei (X, d) ein metrischer Raum und $A \subset X$. Dann heißt A *präkompakt*, falls für alle $\varepsilon \in \mathbb{R}_{>0}$ endlich viele $x_1, \ldots, x_n \in X$ existieren mit $A \subset \bigcup_{i=1}^n B_\varepsilon(x_i)$.

Bemerkung: Man kann Definition 7.1 (ii) dahingehend verschärfen, dass die Mittelpunkte $x_1, \ldots, x_n$ der ε-Kugeln in A liegen müssen. Dies ist äquivalent zur obigen Definition.

Satz 7.2. Sei (X, d) ein metrischer Raum und $A \subset X$ eine Teilmenge. Dann sind die folgenden Aussagen äquivalent:

(1) A ist kompakt

(2) A ist folgenkompakt, d. h. jede Folge in A besitzt eine konvergente Teilfolge mit Grenzwert in A.

(3) A ist präkompakt und $(A, d|_A)$ ist vollständig.

Beweis. „(1) $\Rightarrow$ (2)“: Sei $(x_k)_{k\in\mathbb{N}}$ eine Folge in A. Besitzt A keinen Häufungspunkt so gibt es für alle $y \in A$ ein $r_y \in \mathbb{R}_{>0}$, so dass

$$N_y := \{k \in \mathbb{N} \mid x_k \in B_{r_y}(y) \cap A\}$$

endlich ist. Weil $\big(B_{r_y}(y)\big)_{y\in A}$ eine offene Überdeckung von A bildet und A kompakt ist, gibt es $y_1, \ldots, y_n \in A$ mit $A \subset \bigcup_{i=1}^n B_{r_{y_i}}(y_i)$. Damit wäre $\mathbb{N} \subset \bigcup_{i=1}^n N_{y_i}$ endlich. Widerspruch.

„(2) $\Rightarrow$ (3)“: Jede Cauchy-Folge in A besitzt wegen der Folgenkompaktheit von A eine konvergente Teilfolge. Da aber eine Cauchyfolge mit konvergenter Teilfolge schon selbst konvergiert (gegen den Grenzwert der Teilfolge), ist A also vollständig. Zur Präkompaktheit: Falls es für ein $\varepsilon \in \mathbb{R}_{>0}$ keine endliche ε-Überdeckung von A gibt, wähle induktiv

$$x_{k+1} \in A \setminus \bigcup_{i=1}^{k} B_\varepsilon(x_i).$$

Damit hätte $(x_k)_{k\in\mathbb{N}}$ keinen Häufungspunkt, im Widerspruch zur Folgenkompaktheit von A.

„(3) $\Rightarrow$ (1)“: Sei $(U_i)_{i\in I}$ eine offene Überdeckung von A. Angenommen $(U_i)_{i\in I}$ besitzt keine endliche Teilüberdeckung von A. Wähle zunächst endlich viele $y_1, \dots, y_\ell \in X$, so dass $A \subset \bigcup_{j=1}^{\ell} B_{1/2}(y_j)$ gilt (was wegen der Präkompaktheit von A möglich ist). Dann gibt es ein $B_{1/2}(y_j)$, so dass $B_{1/2}(y_j) \cap A$ nicht von endlich vielen U_i überdeckt wird. Wähle dann $x_1 := y_j$. Falls nun x_n schon definiert ist, so gehen wir analog vor: Seien $y_1^n, \dots, y_{\ell_n}^n \in X$, so dass

$$B_{1/2^n}(x_n) \cap A \;\subset\; A \;\subset\; \bigcup_{j=1}^{\ell_n} B_{1/2^{n+1}}(y_j^n)$$

gilt. Weil nach Konstruktion $B_{1/2^n}(x_n) \cap A$ nicht von endlich vielen U_i überdeckt wird, muss es ein $B_{1/2^{n+1}}(y_j^n)$ geben, das $B_{1/2^n}(x_n) \cap A$ nicht-trivial schneidet und das nicht von endlich vielen U_i überdeckt wird. Setze dann $x_{n+1} := y_j^n$. Somit erhalten wir induktiv eine Folge $(x_n)_{n\in\mathbb{N}}$ in X mit $A_n := B_{1/2^{n+1}}(x_{n+1}) \cap B_{1/2^n}(x_n) \cap A \neq \emptyset$ für alle $n \in \mathbb{N}$. Wähle dann $z_n \in A_n$ für alle $n \in \mathbb{N}$; dies liefert eine Cauchy-Folge in A, was aus

$$d(z_n, z_{n+1}) \leq d(z_n, x_{n+1}) + d(x_{n+1}, z_{n+1}) \leq \frac{1}{2^{n+1}} + \frac{1}{2^{n+1}} = \frac{1}{2^n}$$

für alle $n \in \mathbb{N}$ folgt. Da A vollständig ist, folgt: Es gibt ein $z \in A$ mit $z_n \to z$ für $n \to \infty$. Weil $(U_i)_{i\in I}$ eine Überdeckung von A ist, gibt es also ein $i_0 \in I$ mit $z \in U_{i_0}$. Da U_{i_0} offen ist, gibt es ein $R \in \mathbb{R}_{>0}$, so dass $B_R(z)$ komplett in U_{i_0} enthalten ist. Weil $(z_n)_{n\in\mathbb{N}}$ gegen z konvergiert und nach Wahl von z_n für alle $n \in \mathbb{N}$ auch $d(x_n, z_n) < 1/2^n$ gilt, finden wir ein $N \in \mathbb{N}$ mit

$$B_{1/2^N}(x_N) \subset B_R(z) \subset U_{i_0}.$$

Damit wird aber $B_{1/2^N}(x_N) \cap A$ von U_{i_0} überdeckt, was nach Konstruktion ausgeschlossen war. Also muss es doch eine endliche Teilüberdeckung von $(U_i)_{i\in I}$ geben.

$\blacksquare$

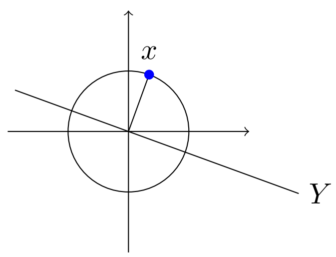


Abbildung 7.1: Maximaler Abstand zwischen Gerade und Einheitssphäre

Satz 7.3 (Fast orthogonales Element). Sei X ein normierter Raum und $Y \subsetneq X$ ein abgeschlossener Unterraum. Weiter sei $\theta \in (0,1)$ oder, falls X ein Hilbertraum ist, $\theta \in (0,1]$. Dann gibt es ein $x_\theta \in X$ mit (s. Abbildung 7.1)

$$\|x_\theta\| = 1 \quad \text{und} \quad \theta \le \operatorname{dist}(x_\theta, Y) \le 1.$$

Beweis. Sei zunächst $\theta \in (0,1)$. Wähle $x \in X \backslash Y$. Da Y abgeschlossen ist, gilt $\operatorname{dist}(x,Y) > 0$. Es gibt daher ein $y_\theta \in Y$ mit

$$\|x - y_\theta\| \le \frac{1}{\theta} \operatorname{dist}(x,Y).$$

Setze nun

$$x_\theta \coloneqq \frac{x - y_\theta}{\|x - y_\theta\|}.$$

Dann gilt für alle $y \in Y$

$$\|x_\theta - y\| = \frac{1}{\|x - y_\theta\|} \big\|x - (y_\theta + \|x - y_\theta\|\, y)\big\|,$$

woraus folgt:

$$\|x_\theta - y\| \ge \frac{\operatorname{dist}(x,Y)}{\frac{1}{\theta}\operatorname{dist}(x,Y)} = \theta.$$

Falls X ein Hilbertraum und $\theta = 1$ ist, so wähle $y_1 = \mathrm{P}_Y(x)$. ■

Satz 7.4 (Heine-Borel). Sei X ein normierter Raum. Dann gilt:

$$\overline{B_1(0)} \text{ kompakt} \quad \Longleftrightarrow \quad \dim X < \infty.$$

Beweis. „$\Leftarrow$“: Sei $e_1, \dots, e_n$ eine Basis von X. Zu $x \in X$ bezeichne $\alpha(x)$ den eindeutigen Vektor in $\mathbb{K}^n$, für den $x = \sum_{i=1}^n \alpha(x)_i\, e_i$ gilt. Da alle Normen äquivalent sind, genügt es, die Kompaktheit von $\overline{B_1(0)}$ in folgender Norm zu zeigen: für $\alpha \in \mathbb{K}^n$ und $x \in X$ sei

$$\|\alpha\|_{\max} \coloneqq \max_i |\alpha_i| \quad \text{und} \quad \|x\|_{\max} \coloneqq \|\alpha(x)\|_{\max}.$$

Nun gilt:

$$\overline{B_1(0)} \quad \subset \bigcup_{\substack{z \in \mathbb{Z}^n, \\ \|z\|_{\max} \le m}} B_{2/m}\Big(\sum_{i=1}^n \frac{z_i}{m}\, e_i\Big).$$

Damit ist $\overline{B_1(0)}$ präkompakt und da in endlich-dimensionalen Räumen abgeschlossene Teilmengen auch vollständig sind, folgt mit Satz 7.2 die Behauptung.

„⇒“: Sei $\varepsilon \in (0,1)$ und
$$\overline{B_1(0)} \subset \bigcup_{i=1}^{n_\varepsilon} B_\varepsilon(x_i)$$
für geeignete $n_\varepsilon \in \mathbb{N},\ x_1, \dots, x_{n_\varepsilon} \in X$. Dann ist
$$Y := \operatorname{span}\{x_1, \dots, x_{n_\varepsilon}\}$$
ein abgeschlossener Unterraum (da endlich-dimensional). Angenommen $Y \subsetneq X$. Dann liefert Satz 7.3 ein x_ε mit $\|x_\varepsilon\| = 1$ und $\varepsilon \leq \operatorname{dist}(x_\varepsilon, Y)$. Dann folgt aber $x \in \overline{B_1(0)}$ und damit muss es ein $i_0 \in \{1, \dots, n_\varepsilon\}$ geben, so dass $x \in B_\varepsilon(x_{i_0})$ gilt. Somit erhalten wir folgenden Widerspruch:
$$\varepsilon \leq \operatorname{dist}(x_\varepsilon, Y) \leq \|x_{i_0} - x_\varepsilon\| < \varepsilon.$$
Also war die Annahme falsch und es muss doch schon $Y = X$ und damit $\dim X \leq n_\varepsilon < \infty$ gelten.

■

Definition 7.5 (Schwache Konvergenz). Sei X ein Banachraum.

1. Eine Folge $(x_n)_{n\in\mathbb{N}}$ in X *konvergiert schwach gegen* $x \in X$, falls gilt:
$$\forall\, x' \in X' : \qquad x'(x_n) \to x'(x) \qquad \text{für } n \to \infty.$$
Notation: $x_n \to x$ schwach in X für $n \to \infty$, oder $x_n \rightharpoonup x$ in X für $n \to \infty$.

2. Eine Folge $(x'_n)_{n\in\mathbb{N}}$ in X' *konvergiert schwach-∗ gegen* $x' \in X'$, falls gilt:
$$\forall\, x \in X : \qquad x'_n(x) \to x'(x) \qquad \text{für } n \to \infty.$$
(Dies entspricht gerade punktweiser Konvergenz von x'.) Notation: $x'_n \to x'$ schwach-∗ in X' für $n \to \infty$, oder $x'_n \overset{*}{\rightharpoonup} x'$ in X' für $n \to \infty$.

3. Eine Teilmenge $M \subset X$ (bzw. $M \subset X'$) heißt *schwach* (bzw. *schwach-∗*) *folgenkompakt*, falls jede Folge in M eine schwach (bzw. schwach-∗) konvergente Teilfolge besitzt, deren Grenzwert in M liegt.

Bemerkungen:

(i) Falls $x_n \to x$ in X für $n \to \infty$ (bezüglich Normkonvergenz) gilt, so sagen wir zur besseren Unterscheidung auch: $(x_n)_{n\in\mathbb{N}}$ konvergiert *stark* gegen x. Die Topologie, die X von der Norm erhält, nennen wir auch *starke Topologie.*

(ii) Die schwache Konvergenz kann als schwach-∗ Konvergenz im Bidualraum aufgefasst werden.
$$\begin{aligned} x_n \rightharpoonup x \quad \text{für } k \to \infty \qquad &\iff \qquad \forall\, x' \in X' : \quad x'(x_n) \to x'(x) \quad \text{für } n \to \infty \\ &\iff \qquad \forall\, x' \in X' : \quad \big(J_X x_n\big)(x') \to \big(J_X x\big)(x') \quad \text{für } n \to \infty \end{aligned}$$

Frage: Welche Konvergenz in X' ist stärker? Schwache Konvergenz oder schwach-$*$ Konvergenz?

$$\begin{aligned} x'_n \rightharpoonup x' \quad &\text{bedeutet} \quad \forall\, x'' \in X'' \colon\; x''(x'_n) \to x''(x') \\ x'_n \overset{*}{\rightharpoonup} x' \quad &\text{bedeutet} \quad \forall\, x \in X \colon\; x'_n(x) \to x'(x) \end{aligned}$$

Es gilt $x'' = J_X x$ für gewisse $x'' \in X''$ (und geeignetes $x \in X$), aber i. A. ist J_X nicht surjektiv (d. h. es gibt „mehr x'' als $J_X x$"). Also verlangt $x'_n \rightharpoonup x'$ mehr als $x'_n \overset{*}{\rightharpoonup} x'$. Damit ist schwach-$*$ Konvergenz im Allgemeinen schwächer als schwache Konvergenz (d. h. es gibt mehr konvergente Folgen bezüglich schwach-$*$ Konvergenz als bezüglich schwacher Konvergenz).

Lemma 7.6. Sei X ein Banachraum. Dann gelten folgende Aussagen:

(1) Der schwache Limes und der schwach-$*$ Limes sind eindeutig bestimmt.

(2) Starke Konvergenz impliziert schwache Konvergenz.

(3) Aus $x'_n \to x'$ schwach-$*$ in X' für $n \to \infty$ folgt:

$$\|x'\| \leq \liminf_{n\to\infty} \|x'_n\|.$$

(4) Aus $x_n \to x$ schwach in X für $n \to \infty$ folgt:

$$\|x\| \leq \liminf_{n\to\infty} \|x_n\|.$$

(Dies zeigt, dass die Norm unterhalbstetig ist bezüglich schwacher Konvergenz.)

(5) Schwach und schwach-$*$ konvergente Folgen sind (bezüglich der Norm) beschränkt.

(6) Seien $(x_n)_{n\in\mathbb{N}}$ bzw. $(x'_n)_{n\in\mathbb{N}}$ Folgen in X bzw. X' und seien $x \in X$ und $x' \in X'$. Sind für $n \to \infty$ die Bedingungen

$$x_n \to x \text{ stark} \quad \text{und} \quad x'_n \overset{*}{\rightharpoonup} x'$$

oder die Bedingungen

$$x_n \rightharpoonup x \quad \text{und} \quad x'_n \to x' \text{ stark}$$

erfüllt, so folgt:

$$x'_n(x_n) \to x'(x) \qquad \text{für } n \to \infty.$$

Beweis. (1) Punktweise Grenzwerte sind eindeutig, also auch der schwach-$*$ Grenzwert. Angenommen x, y sind schwache Grenzwerte von $(x_n)_{n\in\mathbb{N}}$. Dann gilt für alle $x' \in X'$:

$$x'(x) = \lim_{n\to\infty} x'(x_n) = x'(y).$$

Also folgt $x'(x-y) = 0$ für alle $x' \in X'$. Daraus erhalten wir:

$$0 = \sup_{x' \in X'} \|x'(x-y)\| \geq \|J_X(x-y)\| = \|x-y\|,$$

also $\|x-y\| = 0$ und damit muss schon $x = y$ gelten.

(2) Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in X, die stark gegen $x \in X$ konvergiert. Dann gilt $\|x_n - x\| \to 0$ für $n \to \infty$. Damit folgt für alle $x' \in X'$:

$$\left|x'(x_n) - x'(x)\right| = \left|x'(x_n - x)\right| \leq \|x'\| \, \|x - x_n\| \to 0 \qquad \text{für } n \to \infty.$$

Dies entspricht schwacher Konvergenz.

(3) Sei $(x'_n)_{n\in\mathbb{N}}$ eine Folge in X' mit $x'_n \overset{*}{\rightharpoonup} x' \in X'$ für $n \to \infty$. Für alle $x \in X$ gilt

$$\left|x'_n(x)\right| \leq \|x'_n\| \, \|x\|$$

und durch Bilden des lim inf erhalten wir somit:

$$|x'(x)| = \lim_{n\to\infty} |x'_n(x)| = \liminf_{n\to\infty} |x'_n(x)| \leq \|x\| \liminf_{n\to\infty} \|x'_n\|.$$

Es folgt:

$$\|x'\| \leq \liminf_{n\to\infty} \|x'_n\|.$$

(4) Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in X mit $x_n \rightharpoonup x \in X$ für $n \to \infty$. Für $x = 0$ gilt die Behauptung offenbar, also sei nun o. E. $x \neq 0$. Dann gilt für alle $x' \in X'$ (mit analogen Argumenten wie zuvor):

$$\left|x'(x)\right| \leq \|x'\| \liminf_{n\to\infty} \|x_n\|.$$

Wir wählen ein $x' \in X'$ mit $\|x'\| = 1$ und $x'(x) = \|x\|$ (vgl. Konstruktion direkt vor Satz 4.18) und erhalten so die Behauptung.

(5) Sei $(x'_n)_{n\in\mathbb{N}}$ eine Folge in X' mit $x'_n \overset{*}{\rightharpoonup} x'$ in X' für $n \to \infty$. Für alle $x \in X$ ist dann $\big(x'_n(x)\big)_{n\in\mathbb{N}}$ eine konvergente, also beschränkte Folge, d. h. es gilt:

$$\sup_{n\in\mathbb{N}} \left|x'_n(x)\right| < \infty.$$

Banach-Steinhaus (Satz 5.3) liefert:

$$\sup_{n\in\mathbb{N}} \|x'_n\| < \infty.$$

Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in X mit $x_n \rightharpoonup x$ in X für $n \to \infty$, so gilt

$$J_X x_n \overset{*}{\rightharpoonup} J_X x \quad \text{in } X'' \text{ für } n \to \infty.$$

Nach dem vorherigen Absatz ist also $(J_X x_n)_{n\in\mathbb{N}}$ in X'' beschränkt und da J_X eine Isometrie ist, folgt: $(x_n)_{n\in\mathbb{N}}$ ist beschränkt in X.

(6) Seien die ersten beiden Bedinungen aus der Behauptung erfüllt. Dann gilt:

$$\begin{aligned}
\left|x'(x) - x'_n(x_n)\right| &= \left|x'(x) - x'_n(x) + x'_n(x) - x'_n(x_n)\right| \\
&\leq \underbrace{\left|x'(x) - x'_n(x)\right|}_{\to 0, \text{ da } x'_n \overset{*}{\rightharpoonup} x'} + \underbrace{\|x - x_n\| \, \|x'_n\|}_{\substack{\to 0, \text{ da } x_n \to x \text{ stark und} \\ (x'_n)_{n\in\mathbb{N}} \text{ beschränkt nach (5)}}} \\
&\to 0 \qquad \text{für } n \to \infty.
\end{aligned}$$

Sind die zweiten Bedingungen erfüllt, so gilt mit einem analogen Argument:

$$\big|x'(x) - x'_n(x_n)\big| \leq \underbrace{\big|x'(x) - x'(x_n)\big|}_{\to 0,\text{ da } x_n \rightharpoonup x} + \underbrace{\|x' - x'_n\|\,\|x_n\|}_{\substack{\to 0,\text{ da } x'_n \to x' \text{ stark und} \\ (x_n)_{n\in\mathbb{N}} \text{ beschränkt nach (5)}}}$$
$$\to 0 \qquad \text{für } n \to \infty.$$

■

Bemerkung 7.7.

(i) In einem Hilbertraum H gilt:

$$x_n \rightharpoonup x \qquad \Longleftrightarrow \qquad \forall\, y \in H\colon\ \langle x_n, y\rangle \to \langle x, y\rangle.$$

Dies gilt, da wir H' isometrisch isomorph mit H identifizieren können.

(ii) Schwache Konvergenz ist eine Verallgemeinerung der Konvergenz in allen Koordinatenrichtgungen. Ersetze „Koordinaten von x“ durch $x'(x)$ für Funktionale $x' \in X'$.

Satz 7.8. Sei X ein separabler Banachraum über $\mathbb{K}$. Dann ist die abgeschlossene Einheitskugel $\overline{B_1(0)}$ in X' schwach-$*$ folgenkompakt.

Beweis. Sei $\{x_n \mid n \in \mathbb{N}\}$ dicht in X und sei $(x'_k)_{k\in\mathbb{N}}$ eine Folge in X' mit $\|x_k\| \leq 1$ für alle $k \in \mathbb{N}$. Dann ist $\big(x'_k(x_n)\big)_{k\in\mathbb{N}}$ für alle $n \in \mathbb{N}$ eine beschränkte Folge in $\mathbb{K}$. Nach dem Diagonalverfahren gibt es eine Teilfolge $(x'_{k_i})_{i\in\mathbb{N}}$, so dass $\big(x'_{k_i}(x_n)\big)_{i\in\mathbb{N}}$ für alle $n \in \mathbb{N}$ konvergiert. Setze dann für alle $n \in \mathbb{N}$:

$$x'(x_n) \coloneqq \lim_{i\to\infty} x'_{k_i}(x_n).$$

Wir setzen x' linear fort auf $Y \coloneqq \operatorname{span}\{x_n \mid x \in \mathbb{N}\}$. Für alle $y \in Y$ gilt dann

$$|x'(y)| = \lim_{i\to\infty} |x'_{k_i}(y)| \leq \|y\|,$$

woraus folgt, dass x' auf Y gleichmäßig stetig ist und sich somit eindeutig auf $\overline{Y} = X$ fortsetzen lässt. Folglich ist $x' \in X'$ mit $\|x'\| \leq 1$. Sei $x \in X$ und $(y_j)_{j\in\mathbb{N}}$ eine Folge in Y, die gegen x konvergiert. Dann gilt:

$$\begin{aligned}
|(x' - x'_{k_i})(x)| &\leq |(x' - x'_{k_i})(y_j)| + \|x' - x'_{k_i}\|\,\|x - y_j\| \\
&\leq |(x' - x'_{k_i})(y_j)| + 2\|x - y_j\| \\
&\xrightarrow[i\to\infty]{\text{nach Konstr.}} 0 + 2\|x - y_j\| \xrightarrow[j\to\infty]{} 0
\end{aligned}$$

Da x beliebig war, folgt wie gewünscht:

$$\forall\, x \in X\colon \quad x'_{k_i}(x) \to x'(x) \qquad \text{für } i \to \infty.$$

■

Wir hätten gerne die analoge Aussage von Satz 7.8 für $\overline{B_1(0)} \subset X$ und schwache Konvergenz. Dafür brauchen wir Reflexivität und einige Aussagen über reflexive Räume.

Lemma 7.9. Sei X ein Banachraum. Dann gelten folgende Aussagen:

(1) Ist X reflexiv und $Y \subset X$ ein abgeschlossener Unterraum, so ist Y reflexiv.

(2) Sei Y ein weiterer Banachraum und gelte $X \cong Y$ mittels des Operators $T \in L(X,Y)$ (mit $T^{-1} \in L(Y,X)$). Dann ist X genau dann reflexiv, wenn Y reflexiv ist.

(3) Es ist X genau dann reflexiv, wenn X' reflexiv ist.

Beweis. (1) Sei o. E. $Y \neq X$ und sei $y'' \in Y''$. Dann definieren wir $x'' \in X''$ durch

$$x''(x') := y''(x'|_Y).$$

Sei $x := J_X^{-1} x'' \in X$. Für alle $x' \in X'$ mit $x'|_Y = 0$ gilt

$$0 = x''(x') = (J_X x)(x') = x'(x)$$

und somit folgt $x \in Y$ aus (dem Beweis von) Korollar 4.16. Sei $y' \in Y'$ und sei $x' \in X'$ eine Hahn-Banach-Fortsetzung von y' (Satz 4.6). Dann gilt

$$y''(y') = y''(x'|_Y) = x''(x') = x'(x) = y'(x) = (J_Y x)(y').$$

Es folgt $J_Y x = y''$ und somit ist J_Y surjektiv, also Y reflexiv.

(2) Aus Symmetriegründen genügt es, eine Richtung zu zeigen. Sei X reflexiv und $y'' \in Y''$. Definiere $x'' \in X''$ durch

$$x''(x') := y''(x' \circ T^{-1}).$$

Für $y' \in Y'$ gilt dann

$$y''(y') = x''(y' \circ T) = (y' \circ T)(J_X^{-1} x'') = y'(T J_X^{-1} x'') = J_Y(T J_X^{-1} x'')(y'),$$

also folgt $y'' = J_Y(T J_X^{-1} x'')$ und damit ist J_Y surjektiv, d. h. Y reflexiv.

(3) „$\Rightarrow$“: Sei X reflexiv. Ist $x''' \in X'''$, so gilt $x''' \circ J_X \in X'$. Für alle $x'' \in X''$ gilt:

$$x'''(x'') = (x''' \circ J_X)(J_X^{-1} x'') = x''(x''' \circ J_X) = J_{X'}(x''' \circ J_X)(x'').$$

Also folgt $x''' = J_{X'}(x''' \circ J_X)$ und damit ist $J_{X'}$ surjektiv.

„$\Leftarrow$“: Sei X' reflexiv. Nach dem gerade Gezeigten ist X'' reflexiv. Weil J_X eine Isometrie ist, muss $J_X(X) \subset X''$ ein abgeschlossener Unterraum sein. Mit (1) folgt, dass $J_X(X)$ reflexiv ist, und mit (2) die Reflexivität von X.

■

Lemma 7.10. Sei X ein Banachraum. Dann gilt:

$$X' \text{ separabel} \quad \Longrightarrow \quad X \text{ separabel}$$

Beweis. Sei $\{x'_k \mid k \in \mathbb{N}\}$ dicht in X'. Wähle nun für alle $k \in \mathbb{N}$ ein $x_k \in X$ mit $\|x_k\| = 1$ und $x'_k(x_k) \geq \frac{1}{2}\|x'_k\|$. Definiere dann

$$Y := \overline{\operatorname{span}\{x_k \mid k \in \mathbb{N}\}}.$$

Sei $x' \in X'$ mit $x'|_Y = 0$. Dann gilt für alle $k \in \mathbb{N}$:

$$\|x' - x'_k\| \geq |(x' - x'_k)(x_k)| = |x'_k(x_k)| \geq \frac{1}{2}\|x'_k\| \geq \frac{1}{2}(\|x'\| - \|x'_k - x'\|).$$

Daraus folgt

$$\|x'\| \leq 3 \inf_{k \in \mathbb{N}} \|x' - x'_k\| = 0,$$

da $\{x'_k \mid k \in \mathbb{N}\}$ dicht liegt in X'. Mittels Korollar 4.16 folgt $X = Y$.

■

Achtung, die Umkehrung von Lemma 7.10 gilt im Allgemeinen nicht!

Satz 7.11. Sei X ein reflexiver Banachraum. Dann ist $\overline{B_1(0)} \subset X$ schwach folgenkompakt.

Beweis. Sei $(x_k)_{k \in \mathbb{N}}$ eine Folge in $\overline{B_1(0)}$ und

$$Y := \overline{\operatorname{span}\{x_k \mid k \in \mathbb{N}\}}.$$

Dann ist Y offenbar separabel. Da Y ein abgeschlossener Unterraum des reflexiven Banachraums X ist, erhalten wir mit Lemma 7.9 (1), dass auch Y reflexiv ist. Also ist $Y'' = J_Y(Y)$ separabel. Nach Lemma 7.10 ist damit auch Y' separabel. Da $\overline{B_1(0)} \subset (Y')'$ schwach-$*$ folgenkompakt ist (Satz 7.8), existiert eine Teilfolge $(x_{k_i})_{i \in \mathbb{N}}$ und ein $y'' \in Y''$ mit

$$J_Y x_{k_i} \to y'' \quad \text{schwach-}* \text{ in } Y'' \qquad \text{für } i \to \infty.$$

Für $x := J_Y^{-1} y''$ gilt also $J_Y x_{k_i} \overset{*}{\rightharpoonup} J_Y x$ für $i \to \infty$. Somit gilt für alle $y' \in Y'$:

$$y'(x_{k_i}) = (J_Y x_{k_i})(y') \to (J_Y x)(y') = y'(x) \qquad \text{für } i \to \infty.$$

Es folgt $x_{k_i} \to x$ schwach in Y für $i \to \infty$ und damit auch

$$x_{k_i} \to x \quad \text{schwach in } X \qquad \text{für } i \to \infty.$$

■

Satz 7.12 (Anwendung auf Hilberträume)**.** Sei H ein Hilbertraum und $(x_k)_{k \in \mathbb{N}}$ eine beschränke Folge in H. Dann existiert eine Teilfolge $(x_{k_i})_{i \in \mathbb{N}}$ und ein $x \in H$, so dass für alle $y \in H$ gilt:

$$\langle x_{k_i}, y \rangle \to \langle x, y \rangle \qquad \text{für } i \to \infty.$$

Beweis. Nach Bemerkung 6.7 ist H reflexiv. Nach Satz 7.11 sind abgeschlossene Kugeln also schwach folgenkompakt. Das heißt, es existieren eine Teilfolge $(x_{k_i})_{i\in\mathbb{N}}$ und ein $x \in H$ mit:

$$\forall\, x' \in X'\colon \quad x'(x_{k_i}) \to x'(x) \qquad \text{für } i \to \infty.$$

Mit dem Rieszschen Darstellungssatz (6.5) erhalten wir:

$$\forall\, y \in H\ \exists\, x'_y \in H'\colon \quad \langle\,\cdot\,, y\rangle = x'_y$$

Dies impliziert:

$$\forall\, y \in H\colon \quad \langle x_{k_i}, y\rangle = x'_y(x_{k_i}) \to x'_y(x) = \langle x, y\rangle \qquad \text{für } i \to \infty.$$

■

Satz 7.13. Sei H ein Hilbertraum und $(x_k)_{k\in\mathbb{N}}$ eine Folge in H. Dann gilt:

$$\begin{aligned} & x_k \to x \quad \text{stark in } H \qquad \text{für } k \to \infty \\ \Longleftrightarrow \qquad & x_k \to x \quad \text{schwach in } H \quad \text{und} \quad \|x_k\| \to \|x\| \qquad \text{für } k \to \infty. \end{aligned}$$

Beweis. „$\Rightarrow$“ ist klar.
„$\Leftarrow$“: Für alle $k \in \mathbb{N}$ gilt:

$$\begin{aligned} \|x_k\|^2 &= \|x - x + x_k\|^2 = \langle x - x + x_k, x - x + x_k\rangle \\ &= \|x\|^2 + 2\operatorname{Re}\langle x_k - x, x\rangle + \|x_k - x\|^2. \end{aligned}$$

Wegen $x_k \rightharpoonup x$ für $k \to \infty$ gilt $\langle x_k - x, x\rangle \to 0$ für $k \to \infty$. Es folgt

$$\|x - x_k\| \to 0 \qquad \text{für } k \to \infty.$$

■

Satz 7.14. Sei X ein normierter Raum und $M \subset X$ abgeschlossen und konvex. Dann ist M *schwach folgenabgeschlossen*, das heißt: Ist $(x_k)_{k\in\mathbb{N}}$ eine Folge in M mit $x_k \to x$ schwach in X für $k \to \infty$, so ist auch $x \in M$.

Beweis. Sei o. E. M nicht leer und sei $(x_k)_{k\in\mathbb{N}}$ eine Folge in M mit $x_k \rightharpoonup x \in X$ für $k \to \infty$. Angenommen $x \notin M$. Dann liefert der Satz von Hahn-Banach in der zweiten geometrischen Formulierung (4.15) die Existenz eines $x' \in X'$ und eines $\alpha \in \mathbb{R}$ mit

$$\alpha < \operatorname{Re} x'(x) \qquad \text{und} \qquad \forall\, y \in M\colon\ \operatorname{Re} x'(y) \le \alpha.$$

Nach Voraussetzung gilt $x_k \rightharpoonup x$ für $k \to \infty$, also folgt

$$\operatorname{Re} x'(x_k) \to \operatorname{Re} x'(x) \qquad \text{für } k \to \infty.$$

Wegen $x_k \in M$ für alle $k \in \mathbb{N}$ gilt dann aber

$$\operatorname{Re} x'(x) = \lim_{k\to\infty} \operatorname{Re} x'(x_k) \le \alpha,$$

im Widerspruch zu $\operatorname{Re} x'(x) > \alpha$.

■

Lemma 7.15 (Lemma von Mazur). Sei X ein normierter Raum und sei $(x_k)_{k\in\mathbb{N}}$ eine Folge in X mit $x_k \to x$ schwach in X für $k \to \infty$. Dann gilt:

$$x \in \overline{\operatorname{conv}\{x_k \mid k \in \mathbb{N}\}}.$$

(Dabei sei $\operatorname{conv}(A)$ für eine Teilmenge $A \subset X$ die *konvexe Hülle von A in X*, d. h. die kleinste konvexe Teilmenge von X, die A enthält.)

Beweis. Setze

$$M := \operatorname{conv}\{x_k \mid k \in \mathbb{N}\}.$$

Dann ist M konvex und damit auch $\overline{M}$. Aus Satz 7.14 folgt, dass $\overline{M}$ schwach folgenabgeschlossen ist und da $x_k \in M \subset \overline{M}$ für alle $k \in \mathbb{N}$ erfüllt ist, erhalten wir die Behauptung. ■

Definition 7.16. Sei X ein Banachraum und

$$\varphi\colon X \to (-\infty, \infty]$$

eine Abbildung. Dann heißt φ *schwach unterhalbstetig*, falls für alle Folgen $(x_n)_{n\in\mathbb{N}}$ in X mit $x_n \to x$ schwach für $n \to \infty$ gilt:

$$\varphi(x) \le \liminf_{n\to\infty} \varphi(x_n).$$

Satz 7.17. Sei X ein Banachraum und

$$\varphi\colon X \to (-\infty, \infty]$$

konvex und unterhalbstetig in der starken Topologie. Dann ist φ schwach unterhalbstetig.

Beweis. Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in X mit $x_n \rightharpoonup x$ für $n \to \infty$. Für alle $\lambda \in \mathbb{R}$ ist die Menge

$$A_\lambda := \{y \in X \mid \varphi(y) \le \lambda\}$$

konvex (Lemma 4.27 (ii)) und abgeschlossen. Aus Satz 7.14 folgt, dass A_λ für alle $\lambda \in \mathbb{R}$ schwach folgenabgeschlossen ist. Seien

$$L := \liminf_{n\to\infty} \varphi(x_n) \quad \text{und} \quad \varepsilon \in \mathbb{R}_{>0}.$$

(Ein ähnliches Argument wie das folgende zeigt auch $-\infty < L$.) Dann sind unendlich viele Folgenglieder von $\big(\varphi(x_n)\big)_{n\in\mathbb{N}}$ kleiner als $L+\varepsilon$ und damit liegen unendlich viele Folgenglieder von $(x_n)_{n\in\mathbb{N}}$ in $A_{L+\varepsilon}$. Wir können also eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ von $(x_n)_{n\in\mathbb{N}}$ auswählen, so dass alle Folgenglieder dieser Teilfolge in $A_{L+\varepsilon}$ liegen. Weil $(x_n)_{n\in\mathbb{N}}$ schwach gegen x konvergiert, muss dies auch für $(x_{n_k})_{k\in\mathbb{N}}$ gelten. Da $A_{L+\varepsilon}$ schwach folgenabgeschlossen ist, folgt $x \in A_{L+\varepsilon}$. Weil dies für alle $\varepsilon \in \mathbb{R}_{>0}$ gilt, erhalten wir

$$x \in \bigcap_{\varepsilon\in\mathbb{R}_{>0}} A_{L+\varepsilon} = A_L,$$

also $\varphi(x) \le L = \liminf_{n\to\infty} \varphi(x_n)$ wie gewünscht. Damit ist φ unterhalbstetig. ■

Bemerkung: Der Satz gilt auch für

$$\varphi\colon M \to (-\infty, \infty]$$

mit konvexem und abgeschlossenem $M \subset X$. (Setze z. B. φ durch ∞ auf X fort.)

Satz 7.18. Sei X ein reflexiver Banachraum und $M \subset X$ nicht leer, konvex und abgeschlossen. Weiter sei

$$\varphi\colon M \to (-\infty, \infty]$$

konvex und unterhalbstetig mit $D(\varphi) \neq \emptyset$ und, falls M nicht beschränkt ist:

$$\lim_{\substack{x \in M,\\ \|x\| \to \infty}} \varphi(x) = \infty,$$

wobei wir dies wie folgt auffassen:

$$\forall K \in \mathbb{R}_{>0}\ \exists R \in \mathbb{R}_{>0}\ \forall x \in M\colon \quad \|x\| > R \implies \varphi(x) \geq K.$$

Dann nimmt φ sein Minimum an, d. h. es existiert ein $x_0 \in M$, so dass $\varphi(x_0) = \min_M \varphi$ gilt.

Beweis. Sei $m \coloneqq \inf_M \varphi$. Es gilt $m < \infty$ (da $D(\varphi) \neq \emptyset$). Nun sei $(x_k)_{k\in\mathbb{N}}$ eine Folge in M mit $\varphi(x_k) \to m$ für $k \to \infty$. Nach Voraussetzung ist aber M beschränkt oder es gilt $\lim_{\|x\|\to\infty} \varphi(x) = \infty$, also muss $(x_k)_{k\in\mathbb{N}}$ beschränkt sein. Weil X reflexiv ist, gibt es also nach Satz 7.11 eine schwach konvergente Teilfolge von $(x_k)_{k\in\mathbb{N}}$. Sei $(x_{k_i})_{i\in\mathbb{N}}$ solch eine Teilfolge mit Grenzwert x. Da M nach Satz 7.14 schwach folgenabgeschlossen ist, folgt $x \in M$. Nach Satz 7.17 ist φ schwach unterhalbstetig, also gilt:

$$\varphi(x) \leq \liminf_{i\to\infty} \varphi(x_{k_i}) = m.$$

Wegen $x \in M$ gilt dann also

$$\varphi(x) = \inf_M \varphi.$$

■

7.19 (Initialtopologie). Sei X eine Menge und $(Y_i)_{i\in I}$ eine Familie von topologischen Räumen. Weiter sei $(\varphi_i\colon X \to Y_i)_{i\in I}$ eine Familie von Abbildungen. Wir betrachten:

Problem 1: Konstruiere eine Topologie auf X, so dass für alle $i \in I$ die Abbildung φ_i stetig ist. Falls möglich, finde eine Topologie $\mathcal{T}$, die aus wenigen offenen Mengen besteht ($\mathcal{T}$ soll also „ökonomisch" sein).

Bemerkung:

(i) Wir können immer die diskrete Topologie $\mathcal{T} = \mathcal{P}(X)$ wählen, aber dies ist nicht sehr ökonomisch.

(ii) Für alle $i \in I$ muss natürlich gelten:

$$W_i \subset Y_i \text{ offen} \quad \implies \quad \varphi^{-1}(W_i) \in \mathcal{T},$$

da sonst φ_i nicht stetig wäre. Fassen wir alle diese Mengen zusammen, so erhalten wir ein Mengensystem $(U_\lambda)_{\lambda\in\Lambda}$.

Problem 2: Ist X eine Menge und $(U_\lambda)_{\lambda\in\Lambda}$ ein System von Teilmengen von X, so finde die kleinste Topologie $\mathcal{T}$ auf X, so dass U_λ offen ist für alle $\lambda \in \Lambda$.

Wir müssen ein System von Mengen finden, das stabil ist unter endlichen Durchschnitten und beliebigen Vereinigungen.

Wir benutzen folgendes Vorgehen:

(i) Bilde endliche Schnitte

$$\bigcap_{\lambda\in\Gamma} U_\lambda \qquad \text{mit} \qquad \Gamma \subset \Lambda \text{ endlich.}$$

Diese neue Familie nennen wir Φ. (Falls $\Gamma = \emptyset$, sei $\bigcap_{\lambda\in\Gamma} U_\lambda = X$.)

(ii) Bilde beliebige Vereinigungen von Elementen in Φ und nenne dieses System $\mathcal{T}$.

Lemma 7.20. Das Mengensystem $\mathcal{T}$ aus 7.19, Problem 2 (ii) ist stabil unter endlichen Durchschnitten.

(Beweis: selber oder in ein Topologiebuch schauen.)

Definition 7.21.

1. Wir nennen die Topologie, die wir durch den Prozess in 7.19 aus der Familie $(\varphi_i)_{i\in I}$ erhalten, die *die von $(\varphi_i)_{i\in I}$ erzeugte Topologie.*
2. Sei im Folgenden $(X, \mathcal{T})$ ein topologischer Raum. Wir nennen $\mathcal{F} \subset \mathcal{T}$ eine *Basis der Topologie*, falls für alle Punkte $x \in X$ und alle Umgebungen U von x ein $A \in \mathcal{F}$ existiert mit $x \in A \subset U$.
3. Sei $x \in X$. Wir nennen $\mathcal{F}_x \subset \mathcal{T}$ eine *Umgebungsbasis von x*, falls für alle Umgebungen U von x ein $A \in \mathcal{F}_x$ existiert mit $x \in A \subset U$.

Bemerkung: In der Konstruktion in 7.19 bilden die Mengen

$$\bigcap_{i\in J} \varphi_i^{-1}(W_i) \qquad \text{für } J \subset I \text{ endlich und } W_i \subset Y_i \text{ offen}$$

eine Basis der Topologie $\mathcal{T}$.

Proposition 7.22. Sei $(\varphi_i)_{i\in I}$ wie in 7.19 und $\mathcal{T}$ die erzeugte Topologie auf X. Weiter sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in X. Dann gilt:

$$x_n \to x \quad \text{für } n \to \infty \qquad \Longleftrightarrow \qquad \forall\, i \in I\colon\ \varphi_i(x_n) \to \varphi_i(x) \quad \text{für } n \to \infty.$$

(Beweis: selber.)

Definition 7.23 (Schwache Topologie)**.** Sei X ein Banachraum. Die *schwache Topologie $\mathcal{T}_{\mathrm{w}}$ auf X* ist die von $(\varphi)_{\varphi \in X'}$ erzeugte Topologie.

Bemerkung: Mittels Hahn-Banach kann man zeigen, dass $(X, \mathcal{T}_{\mathrm{w}})$ ein Hausdorffraum ist. Außerdem gilt der folgende Satz:

Satz 7.24. Sei X ein Banachraum. Sei für ein Tripel (n, z', ε) mit $n \in \mathbb{N}$, $z' \in (X')^n$ und $\varepsilon \in \mathbb{R}_{>0}$:

$$U_{n,z',\varepsilon} \coloneqq \big\{x \in X \;\big|\; \forall\, k \in \{1, \ldots, n\}\colon\; |z'_k(x)| < \varepsilon\big\}$$

(diese Mengen bilden eine Umgebungsbasis von $0 \in X$). Dann gilt:

$$\mathcal{T}_{\mathrm{w}} = \big\{A \subset X \;\big|\; \forall\, x \in A\; \exists\, (n, z', \varepsilon)\colon\; x + U_{n,z',\varepsilon} \subset A\big\}.$$

(Beweis: selber.)

Diese Umgebungen erlauben nur Kontrolle von endlich vielen Koordinaten (anders als bei der starken Topologe, bei der innerhalb eines Balls alle Koordinaten unter Kontrolle sind).

Bemerkung: Es ist $\mathcal{T}_{\mathrm{w}}$ die schwächste Topologie, so dass alle $x' \in X'$ noch stetig sind. Auf X' führe folgende Topologie ein, die wir $\mathcal{T}'_{\mathrm{w}}$ nennen: Für ein Tripel (n, z, ε) mit $n \in \mathbb{N}$, $z \in X^n$ und $\varepsilon \in \mathbb{R}_{>0}$ setze

$$U_{n,z,\varepsilon} \coloneqq \big\{x' \in X' \;\big|\; \forall\, k \in \{1, \ldots, n\}\colon\; |x'(z_k)| < \varepsilon\big\}$$

und

$$\mathcal{T}'_{\mathrm{w}} \coloneqq \big\{A \subset X' \;\big|\; \forall\, x' \in A\; \exists\, (n, z, \varepsilon)\colon\; x' + U_{n,z,\varepsilon} \subset A\big\}.$$

Es gilt: X' wird mit $\mathcal{T}'_{\mathrm{w}}$ zu einem topologischer Raum (*schwach-$*$ Topologie*).

Satz 7.25 (Satz von Alaoglu)**.** Sei X ein Banachraum. Dann ist $\overline{B_1(0)} \subset X'$ kompakt bezüglich der schwach-$*$ Topologie auf X'.

Bemerkung: Ist X nicht separabel, so ist „kompakt“ im Allgemeinen nicht dasselbe wie „folgenkompakt“ bezüglich der schwach-$*$ Topologie.

8 Spektrum für kompakte Operatoren

Ziel: Verallgemeinerung der Jordan'schen Normalform auf den unendlich-dimensionalen Fall.

Definition 8.1 (Kompakter Operator)**.** Seien X, Y Banachräume über $\mathbb{K}$. Dann heißt $T \in L(X, Y)$ *kompakter (linearer) Operator*, falls eine der folgenden Bedingungen erfüllt ist:

(1) $\overline{T\big(B_1(0)\big)}$ ist kompakt in Y

(2) $\forall\, M \subset X\colon\quad M$ beschränkt $\implies T(M)$ präkompakt in Y

(3) Für jede beschränkte Folge $(x_n)_{n\in\mathbb{N}}$ in X besitzt $(Tx_n)_{n\in\mathbb{N}}$ eine konvergente Teilfolge.

Satz: Die Bedingungen aus Definition 8.1 sind äquivalent.

Beweis.
„(1)$\Rightarrow$(2)“: Sei $R \in \mathbb{R}_{>0}$ und $M \subset B_R(0)$. Dann ist $\overline{T(B_R(0))}$ nach Voraussetzung kompakt in Y. Dann ist auch die darin enthaltene abgeschlossene Menge $\overline{T(M)}$ kompakt und somit ist $T(M)$ relativ kompakt und damit präkompakt.

„(2)$\Rightarrow$(3)“: Sei $R \in \mathbb{R}_{>0}$ und sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in X mit $\|x_n\| \le R$ für alle $n \in \mathbb{N}$. Dann ist $(Tx_n)_{n\in\mathbb{N}}$ eine Folge in $\overline{T(B_R(0))}$. Aber $T(B_R(0))$ ist nach Voraussetzung präkompakt, also auch relativ kompakt, womit $\overline{T(B_R(0))}$ kompakt und damit auch folgenkompakt sein muss.

„(3)$\Rightarrow$(1)“: Sei $(y_n)_{n\in\mathbb{N}}$ eine Folge in $\overline{T(B_1(0))}$. Für alle $n \in \mathbb{N}$ sei dann $x_n \in B_1(0) \subset X$ mit $\|y_n - Tx_n\| \le 1/n$. Nach Voraussetzung exisitiert dann eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$, so dass $(Tx_{n_k})_{k\in\mathbb{N}}$ konvergiert; bezeichne $y \in Y$ den zugehörigen Grenzwert. Aus der Konstruktion der Folge $(x_n)_{n\in\mathbb{N}}$ folgt nun $y_{n_k} \to y$ für $k \to \infty$. Damit ist $\overline{T(B_1(0))}$ folgenkompakt, also auch kompakt.

∎

Definition 8.2. Für Banachräume X, Y seien

$$K(X,Y) \coloneqq \{T \in T(X,Y) \mid T \text{ ist kompakt}\} \quad \text{und} \quad K(X) \coloneqq K(X,X).$$

In reflexiven Räumen gibt es folgende Charakterisierung kompakter Operatoren:

Lemma 8.3. Seien X, Y Banachräume und sei X reflexiv. Sei $T\colon X \to Y$ linear. Dann gilt:

$$T \in K(X,Y) \quad \Longleftrightarrow \quad T \text{ vollstetig},$$

wobei T *vollstetig* ist, falls gilt: Ist $(x_n)_{n\in\mathbb{N}}$ eine Folge in X mit $x_n \rightharpoonup x$ in X für $n \to \infty$, so gilt $Tx_n \to Tx$ stark in Y für $n \to \infty$.

Beweis. „$\Rightarrow$": Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in X mit $x_n \rightharpoonup x$ für $n \to \infty$. Da schwach konvergente Folgen nach Lemma 7.6 (5) beschränkt sind, ist $(x_n)_{n\in\mathbb{N}}$ beschränkt. Da $\overline{T(B_1(0))}$ nach Voraussetzung kompakt ist, existiert ein $y \in Y$ und eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$, so dass $(Tx_{n_k})_{k\in\mathbb{N}}$ stark gegen y konvergiert. Für $y' \in Y'$ ist $z \mapsto y'(Tz)$ eine Abbildung in X', also gilt:

$$y'(Tx_n) \to y'(Tx) \qquad \text{für } n \to \infty.$$

Daraus folgt $Tx_n \rightharpoonup Tx$ für $n \to \infty$. Da starke Konvergenz auch schwache Konvergenz (gegen denselben Grenzwert) impliziert, gilt $y = Tx$. Also gilt $Tx_{n_k} \to Tx$ stark für $k \to \infty$. Das gleiche Argument gilt für jede Teilfolge. Daraus folgt, dass die gesamte Folge konvergiert.

„$\Leftarrow$": Aus Vollstetigkeit folgt Stetigkeit. Also gilt $T \in L(X,Y)$. Sei $(x_n)_{n\in\mathbb{N}}$ eine beschränkte Folge in X. Nach Satz 7.11 existiert dann eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ mit $x_{n_k} \rightharpoonup x \in X$ für $k \to \infty$. Da T nach Voraussetzung vollstetig ist, gilt dann aber

$$Tx_{n_k} \to Tx \qquad \text{für } k \to \infty,$$

also folgt mit der dritten Charakterisierung in Definition 8.1 die Behauptung. ∎

Lemma 8.4. Seien X, Y Banachräume.

(i) Sei $T \in L(X,Y)$ mit $\dim R(T) < \infty$. Dann folgt $T \in K(X,Y)$.

(ii) Sei $P \in P(X)$ ein Projektor. Dann gilt:

$$P \in K(X) \quad \Longleftrightarrow \quad \dim R(P) < \infty.$$

Beweis. (i) Mit $R := \|T\|$ gilt:

$$\overline{T\big(B_1(0)\big)} \subset \overline{B_R(0)}.$$

Die Teilmenge $\overline{B_R(0)} \cap R(T)$ ist ein abgeschlossener Ball im endlich dimensionalen Raum $R(T)$, also folgt mit Heine-Borel (Satz 7.4), dass sie auch kompakt ist. Daraus folgt die Kompaktheit von $\overline{T\big(B_1(0)\big)}$.

(ii) „$\Leftarrow$" folgt aus (i).
„$\Rightarrow$": Es gilt

$$\overline{B_1(0)} \cap R(P) \subset \overline{P\big(B_1(0)\big)}.$$

Damit ist $\overline{B_1(0)} \cap R(P)$ kompakt und aus Heine-Borel (Satz 7.4) folgt, dass $R(P)$ endlich dimensional ist.

∎

Lemma 8.5. Seien X, Y, Z Banachräume und $T_1 \in L(X, Y)$ sowie $T_2 \in L(Y, Z)$. Ist dann T_1 oder T_2 kompakt, so ist $T_2 T_1$ kompakt.

Beweis. Sei $(x_n)_{n\in\mathbb{N}}$ eine beschränkte Folge in X. Da T_1 stetig ist, ist $(T_1 x_n)_{n\in\mathbb{N}}$ beschränkt in Y. Falls T_2 kompakt ist, existiert eine konvergente Teilfolge $(T_2 T_1 x_{n_k})_{k\in\mathbb{N}}$. Ist T_1 kompakt, so existiert eine konvergente Teilfolge $(T_1 x_{n_k})_{k\in\mathbb{N}}$. Weil T_2 stetig ist, konvergiert dann auch $(T_2 T_1 x_{n_k})_{k\in\mathbb{N}}$.

■

Definition 8.6 (Spektrum)**.** Sei X ein Banchraum über $\mathbb{K}$ und $T \in L(X)$.

(i) Die *Resolventenmenge von T* sei

$$\varrho(T) \coloneqq \big\{\lambda \in \mathbb{K} \mid N(\lambda \operatorname{Id} - T) = \{0\} \wedge R(\lambda \operatorname{Id} - T) = X\big\}.$$

(Im endlich-dimensionalen Fall folgt eine der Bedingungen aus dieser Definition aus der jeweils anderen. Im unendlich-dimensionalen muss dies *nicht* gelten!)

Das *Spektrum von T* ist

$$\sigma(T) \coloneqq \mathbb{K} \setminus \varrho(T).$$

Das Spektrum kann zerlegt werden in das *Punktspektrum*

$$\sigma_{\mathrm{p}}(T) \coloneqq \big\{\lambda \in \sigma(T) \mid N(\lambda \operatorname{Id} - T) \neq \{0\}\big\},$$

das *kontinuierliche Spektrum*

$$\sigma_{\mathrm{c}}(T) \coloneqq \big\{\lambda \in \sigma(T) \mid N(\lambda \operatorname{Id} - T) = \{0\} \wedge R(\lambda \operatorname{Id} - T) \neq X \wedge \overline{R(\lambda \operatorname{Id} - T)} = X\big\}$$

und das *Residualspektrum*

$$\sigma_{\mathrm{r}}(T) \coloneqq \big\{\lambda \in \sigma(T) \mid N(\lambda \operatorname{Id} - T) = \{0\} \wedge \overline{R(\lambda \operatorname{Id} - T)} \neq X\big\}.$$

Bemerkung 8.7. (i) Es gilt $\lambda \in \varrho(T)$ genau dann, wenn $(\lambda \operatorname{Id} - T)\colon X \to X$ bijektiv ist. Nach dem Satz von der inversen Abbildung (5.10) ist dies äquivalent zur Existenz von

$$R(\lambda, T) \coloneqq (\lambda \operatorname{Id} - T)^{-1} \in L(X).$$

Wir nennen $R(\lambda, T)$ *Resolvente von T* und die Abbildung $\lambda \mapsto R(\lambda, T)$ die *Resolventenfunktion von T*.

(ii) Zu $\lambda \in \sigma_{\mathrm{p}}(T)$ ist äquivalent: Es gibt ein $x \in X \setminus \{0\}$ mit $Tx = \lambda x$. Dann heißt λ *Eigenwert* und x *Eigenvektor zum Eigenwert λ*.

(iii) Wir nennen $N(\lambda \operatorname{Id} - T)$ den *Eigenraum von T zum Eigenwert λ*.

(iv) Wir sagen $Y \subset X$ ist T-invariant, falls $T(Y) \subset Y$ gilt. Der Eigenraum zu einem Eigenwert ist stets T-invariant.

Satz 8.8. Sei X ein Banachraum über $\mathbb{K}$ und $T \in L(X)$. Dann ist $\varrho(T)$ offen und die Resolventenfunktion $R(\,\cdot\,, T)$ ist eine analytische Abbildung von $\varrho(T)$ nach $L(X)$. Weiterhin gilt für alle $\lambda \in \varrho(T)$:

$$\|R(\lambda, T)\|^{-1} \leq \operatorname{dist}\big(\lambda, \sigma(T)\big).$$

Dass $R(\,\cdot\,, T)$ analytisch ist, bedeutet dabei: Für alle $\lambda \in \varrho(T)$ existiert ein $r \in \mathbb{R}_{>0}$ mit $B_r(\lambda) \subset \varrho(T)$ und eine Folge $(a_n)_{n\in\mathbb{N}}$ in $L(X)$, so dass gilt:

$$\forall\, \mu \in B_r(\lambda)\colon \quad R(\mu, T) = \sum_{n=0}^{\infty} a_n\,(\mu - \lambda)^n.$$

Beweis. Sei $\lambda \in \varrho(T)$. Dann gilt für alle $\mu \in \mathbb{K}$:

$$\mu \operatorname{Id} - T = (\lambda \operatorname{Id} - T)\underbrace{\big(\operatorname{Id} - (\lambda - \mu)R(\lambda, T)\big)}_{=: S(\mu)}.$$

Indem wir die Neumann'sche Reihe (Satz 3.7) benutzen, folgt, dass $S(\mu)$ invertierbar ist, falls

$$\|(\lambda - \mu)R(\lambda, T)\| < 1 \quad \text{bzw.} \quad |\lambda - \mu| < \|R(\lambda, T)\|^{-1}$$

gilt. Sei $r := \|R(\lambda, T)\|^{-1}$ und $\mu \in B_r(\lambda)$. Nach den obigen Argumenten ist dann $\mu \operatorname{Id} - T$ invertierbar und damit gilt $\mu \in \varrho(T)$. Wir erhalten mit der Reihendarstellung von $S(\mu)$ aus Satz 3.7:

$$R(\mu, T) = \big(S(\mu)\big)^{-1} R(\lambda, T) = \sum_{n=0}^{\infty} (\mu - \lambda)^n\,(-1)^n R(\lambda, T)^{n+1}.$$

Außerdem gilt $B_r(\lambda) \subset \varrho(T)$, woraus $\operatorname{dist}(\lambda, \sigma(T)) \geq r = \|R(\lambda, T)\|^{-1}$ folgt. ∎

Bemerkung: Weil $\varrho(T)$ offen ist, ist $\sigma(T)$ abgeschlossen.

Definition 8.9. Sei X ein Banachraum und $T \in L(X)$. Sei

$$r(T) := \inf_{n\in\mathbb{N}} \|T^n\|^{1/n} = \lim_{n\to\infty} \|T^n\|^{1/n}.$$

Wir nennen $r(T)$ den *Spektralradius von T*.

Die Gleichheit in dieser Definition folgt dabei aus dem folgenden Lemma:

Lemma 8.10. Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}$ mit

$$0 \leq a_{n+m} \leq a_n\, a_m$$

für alle $n, m \in \mathbb{N}$. Dann konvergiert $\big(\sqrt[n]{a_n}\big)_{n\in\mathbb{N}}$ gegen $\inf_{n\in\mathbb{N}} \sqrt[n]{a_n}$.

Beweis. Sei $a \coloneqq \inf_{n\in\mathbb{N}} \sqrt[n]{a_n}$ und sei $\varepsilon \in \mathbb{R}_{>0}$. Wähle $N \in \mathbb{N}$ mit $\sqrt[N]{a_N} < a + \varepsilon$ und setze

$$b_\varepsilon \coloneqq \max\{a_1, \ldots, a_N\}.$$

Sei $n \in \mathbb{N}$ und seien $k \in \mathbb{N}$ und $r \in \{1, \ldots, N\}$ mit $n = kN + r$. Dann gilt

$$\begin{aligned}
\sqrt[n]{a_n} &= (a_{kN+r})^{1/n} \leq \left(a_N^k\, a_r\right)^{1/n} \\
&\leq (a+\varepsilon)^{kN/n}\, b_\varepsilon^{1/n} = (a+\varepsilon)^{1+r/n}\, a_r^{1/n} \\
&\leq (a+\varepsilon)^{1+N/n}\, b_\varepsilon^{1/n} \\
&\to a + \varepsilon \qquad \text{für } n \to \infty.
\end{aligned}$$

Für $\varepsilon \to 0$ erhalten wir $\lim_{n\to\infty} \sqrt[n]{a_n} \leq a = \inf_{n\in\mathbb{N}} \sqrt[n]{a_n}$ und daraus folgt die Behauptung. ■

Satz 8.11. Sei X ein Banachraum über $\mathbb{K}$ und sei $T \in L(X)$. Dann gilt:

(a) $\forall\, \lambda \in \sigma(T)\colon \quad |\lambda| \leq r(T)$

(b) Für $\mathbb{K} = \mathbb{C}$ existiert ein $\lambda \in \sigma(T)$ mit $|\lambda| = r(T)$, d. h. es gilt:

$$\sup_{\lambda\in\sigma(T)} |\lambda| = \lim_{n\to\infty} \|T^n\|^{1/n}.$$

(c) $\sigma(T)$ ist kompakt.

(d) $\sigma(T) \neq \emptyset$ für $\mathbb{K} = \mathbb{C}$.

Beweis. (c) Sei $\lambda \in \mathbb{K} \setminus \{0\}$. Mithilfe der Neumann'schen Reihe (Satz 3.7) folgt, dass $\operatorname{Id} - T/\lambda$ invertierbar ist, falls $\|T/\lambda\| < 1$ bzw. $|\lambda| > \|T\|$ gilt. Außerdem gilt dann

$$R(\lambda, T) = \lambda^{-1}(\operatorname{Id} - T/\lambda)^{-1} = \lambda^{-1} \sum_{n=0}^{\infty} \lambda^{-n} T^n, \qquad (*)$$

was wir später benötigen werden. Wir erhalten also

$$s \coloneqq \sup_{\lambda\in\sigma(T)} |\lambda| \leq \|T\|.$$

Nach Definition gilt $\sigma(T) = \mathbb{K} \setminus \varrho(T)$ und $\varrho(T)$ ist nach Satz 8.8 offen, also ist $\sigma(T)$ abgeschlossen. Aus Abgeschlossenheit und Beschränktheit folgt nun mit dem Satz von Heine-Borel für endlich-dimensionale euklidsche Räume, dass $\sigma(T)$ kompakt ist.

(a) Seien $\lambda \in \sigma(T)$ und $m \in \mathbb{N}$ und sei s wie im Beweis von (c). Es gilt

$$\lambda^m \operatorname{Id} - T^m = (\lambda \operatorname{Id} - T)\, S_m(T) = S_m(T)\, (\lambda \operatorname{Id} - T)$$

mit

$$S_m(T) = \sum_{i=0}^{m-1} \lambda^{m-1-i} T^i.$$

Wäre nun $\lambda^m \operatorname{Id} - T^m$ bijektiv, so müsste $\lambda \operatorname{Id} - T$ nach jeweils einer der obigen Gleichheiten surjektiv und injektiv, d. h. auch bijektiv sein. Da dies nach Voraussetzung nicht der Fall ist, kann also auch $\lambda^m \operatorname{Id} - T^m$ nicht bijektiv sein. Daraus erhalten wir (mit dem Beweis von (c)):

$$\lambda^m \in \sigma(T^m) \implies |\lambda^m| \leq \|T^m\| \implies |\lambda| \leq \|T^m\|^{1/m}.$$

Dies zeigt:

$$s \leq \lim_{n\to\infty} \|T^n\|^{1/n} = r(T).$$

(d) Gelte $\mathbb{K} = \mathbb{C}$. Wir nehmen an, das $\sigma(T) = \emptyset$ gilt. Dann ist die Resolventenfunktion

$$\lambda \mapsto R_\lambda := R(\lambda, T) = (\lambda \operatorname{Id} - T)^{-1}$$

auf ganz $\mathbb{C}$ definiert und lokal in eine Potenzreihe entwickelbar (mittels der Neumann'schen Reihe). Sei $\ell \in (L(X))'$. Die Funktion $\lambda \mapsto \ell(R_\lambda)$ hat lokal um $\lambda \in \mathbb{K}$ die Gestalt

$$\mu \mapsto \ell(R_\mu) = \sum_{n=0}^{\infty} (-1)^n \ell\big(R_\lambda^{n+1}\big)\,(\mu - \lambda)^n \tag{$**$}$$

(vgl. Beweis von Satz 8.8). Die Funktion $\lambda \mapsto \ell(R_\lambda)$ ist somit analytisch. Sie ist außerdem beschränkt, denn: Für $|\lambda| > 2\|T\|$ gilt nach $(*)$

$$|\ell(R_\lambda)| \leq \|\ell\|\,\big|\lambda^{-1}\big| \sum_{n=0}^{\infty} \frac{\|T\|^n}{|\lambda|^n} \leq \|\ell\|\,\frac{1}{\|T\|}$$

und auf $\overline{B_{2\|T\|}(0)} \subset \mathbb{C}$ ist sie beschränkt, da sie stetig ist. Aus dem Satz von Liouville folgt: Die Funktion $\lambda \mapsto \ell(R_\lambda)$ ist konstant. Dies kann aber für $\lambda = 0$ in $(**)$ nur gelten, falls alle Koeffizienten bis auf den nullten verschwinden. Inbesondere gilt somit:

$$0 = \ell(R_0^2) = \ell\big((T^2)^{-1}\big).$$

Da $\ell \in L(X)'$ beliebig war, folgt nun mithilfe des Satzes von Hahn-Banach (zum Beispiel wie im Beweis von Lemma 7.6 (1))

$$\big(T^2\big)^{-1} = 0.$$

Dies ist aber ein Widerspruch, d. h. die Annahme muss falsch gewesen sein, d. h. es gilt doch $\sigma(T) \neq \emptyset$.

(b) Gelte $\mathbb{K} = \mathbb{C}$. Wir zeigen zunächst:

$$s := \sup_{\lambda\in\sigma(T)} |\lambda| = r(T).$$

Aus (a) folgt: $s \leq r(T)$. Sei $\mu \in \mathbb{C}$ mit $|\mu| > s$ und sei $\ell \in (L(X))'$. Daraus, dass $R(\,\cdot\,, T)$ analytisch ist, folgt, dass auch $\ell \circ R(\,\cdot\,, T)\colon \varrho(T) \to \mathbb{C}$ analytisch ist. Für $\tilde{\mu} \in \varrho(T)$ mit $|\tilde{\mu}| > \|T\|$ wissen wir nach $(*)$:

$$\ell\big(R(\tilde{\mu}, T)\big) = \sum_{n=0}^{\infty} \ell\big(T^n \cdot (\tilde{\mu})^{-(n+1)}\big).$$

Wählen wir nun $\tilde{\mu}$ geeignet (so dass $\mu \in B_r(\tilde{\mu}) \subset \varrho(T)$ für ein $r \in \mathbb{R}_{>0}$ gilt), so folgt mit dem Potenzreihenentwicklungssatz (bzw. aus der Cauchyformel), dass diese Reihe auch für μ (statt $\tilde{\mu}$) konvergiert. Dann muss aber $(\ell(T^n/\mu^{n+1}))_{n\in\mathbb{N}}$ eine Nullfolge in $\mathbb{C}$ sein, d. h. es gilt:

$$\lim_{n\to\infty} \ell\big(T^n/\mu^{n+1}\big) = 0.$$

Weil $\ell \in L(X)'$ beliebig war, folgt:

$$T^n/\mu^{n+1} \rightharpoonup 0 \qquad \text{für } n \to \infty.$$

Also ist $(T^n/\mu^{n+1})_{n\in\mathbb{N}}$ beschränkt (Lemma 7.6 (5)), etwa durch $K \in \mathbb{R}_{>0}$. Für alle $n \in \mathbb{N}$ gilt dann

$$\|T^n\|^{1/n} \leq K^{1/n}\, |\mu|^{(n+1)/n},$$

woraus

$$r(T) = \lim_{n\to\infty} \|T^n\|^{1/n} \leq \lim_{n\to\infty} K^{1/n}\, |\mu|^{(n+1)/n} = |\mu|$$

und damit auch $r(T) \leq s$ folgt. Weil $\sigma(T)$ nach (c) kompakt und $|\cdot|$ stetig ist, wird das Supremum in $r(T) = s = \sup_{\sigma(T)}|\cdot|$ angenommmen und damit ist alles gezeigt.

■

Bemerkungen 8.12.

(i) Ist $\dim X < \infty$, so gilt $\sigma(T) = \sigma_{\mathrm{p}}(T)$.

(ii) Ist $\dim X = \infty$ und $T \in K(X)$, so gilt $0 \in \sigma(T)$. Im Allgemeinen ist 0 aber *kein* Eigenwert.

Beweis. (i) Sei $\lambda \in \sigma(T)$. Dann ist $\lambda \operatorname{Id} - T$ nicht bijektiv. Da X endlich-dimensional ist, folgt, dass $\lambda \operatorname{Id} - T$ auch nicht injektiv ist. Damit folgt $\lambda \in \sigma_{\mathrm{p}}(T)$.

(ii) Sei $T \in K(X)$ und $0 \in \varrho(T)$. Dann gilt $T^{-1} \in L(X)$ und nach Lemma 8.5 gilt:

$$\operatorname{Id} = T^{-1}T \in K(X).$$

Heine-Borel (Satz 7.4) liefert: $\dim X < \infty$. Dies impliziert, dass $0 \in \varrho(T)$ nur für $\dim X < \infty$ möglich ist.

■

Definition 8.13 (Fredholm-Operator)**.** Seien X, Y Banachräume und sei $A \in L(X, Y)$. Dann heißt A *Fredholm-Operator*, falls gilt:

(1) $\dim N(A) < \infty$

(2) $R(A)$ ist abgeschlossen

(3) $\operatorname{codim} R(A) < \infty$

Der *Index von A* ist dann definiert als

$$\operatorname{ind} A \coloneqq \dim N(A) - \operatorname{codim} R(A).$$

Definition 8.14 (Kodimension)**.** Sei Y ein Banachraum.

(i) Sei $Z \subset Y$ ein abgeschlossener Unterraum. Wir sagen Z *besitzt endliche Kodimension*, falls ein Unterraum $Y_0 \subset Y$ existiert mit $\dim Y_0 < \infty$ und $Y = Z \oplus Y_0$.

(ii) Die *Kodimension von Z* ist dann definiert durch $\operatorname{codim} Z \coloneqq \dim Y_0$.

Lemma 8.15. Sei Y ein Banachraum und $Z \subset Y$ ein abgeschlosser Unterraum. Besitzt Z endliche Kodimension, so ist $\operatorname{codim} Z$ eindeutig bestimmt.

Beweisskizze. Sei $Y_0 \subset Y$ ein Unterraum mit $Y = Z \oplus Y_0$ und $\dim Y_0 < \infty$. Nach Voraussetzung ist Z abgeschlossen und Y_0 ist endlich-dimensional, also auch abgeschlossen. Nach Satz 5.14 existiert also ein Projektor $P \in P(Y)$ auf Y_0 mit $Z = N(P)$.

Sei nun $Y_1 \subset Y$ ein Unterraum mit $Z \cap Y_1 = \{0\}$. Dann ist $S \coloneqq P|_{Y_1} \colon Y_1 \to Y_0$ linear und injektiv. Daraus folgt: Y_1 ist endlich-dimensional mit $\dim Y_1 \leq \dim Y_0$ und es gilt Gleichheit genau dann, wenn $Y = Z \oplus Y_1$. (Siehe Übungen.)

Ist $Y = Z \oplus Y_1$, so tausche die Rollen von Y_1 und Y_0. Es folgt: $\dim Y_1 = \dim Y_0$. (Außerdem ist dann S bijektiv.)

■

Bemerkung: Eine große Klasse von Fredholm-Operatoren ergibt sich aus kompakten Störungen der Identität, d. h. $A = \operatorname{Id} - T$ für $T \in K(X)$.

Satz 8.16. Sei X ein Banachraum über $\mathbb{K}$ und sei $T \in K(X)$. Dann ist $A \coloneqq \operatorname{Id} - T$ ein Fredholm-Operator mit Index 0. Genauer ergibt sich dies aus den folgenden Einzelaussagen:

(1) $\dim N(A) < \infty$

(2) $R(A)$ ist abgeschlossen

(3) $N(A) = \{0\} \implies R(A) = X$

(4) $R(A) = X \implies N(A) = \{0\}$

(5) $\operatorname{codim} R(A) = \dim N(A)$

Beweis. (1) Für $x \in X$ gilt:

$$x \in N(A) \iff Ax = 0 \iff x = Tx.$$

Daher gilt

$$B_1(0) \cap N(A) \subset T\big(B_1(0)\big).$$

Weil T kompakt ist, ist die Einheitskugel in $N(A)$ also präkompakt. Nach Heine-Borel (Satz 7.4) ist damit $N(A)$ endlich-dimensional.

(2) Sei $x \in \overline{R(A)}$ und sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in X mit $Ax_n \to x$ für $n \to \infty$. Ohne Einschränkung können wir annehmen, dass für alle $n \in \mathbb{N}$ schon

$$\|x_n\| \leq 2d_n \quad \text{mit} \quad d_n := \operatorname{dist}\big(x_n, N(A)\big)$$

gilt, denn: Zu $n \in \mathbb{N}$ können wir $a_n \in N(A)$ mit $\|x_n - a_n\| \leq 2d_n$ wählen und dann $(x_n - a_n)_{n\in\mathbb{N}}$ betrachten. Wir zeigen unten, das $(d_n)_{n\in\mathbb{N}}$ beschränkt ist, also ist auch $(x_n)_{n\in\mathbb{N}}$ beschränkt. Wegen $T \in K(X)$ existieren dann eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ und ein $y \in X$ mit

$$Tx_{n_k} \to y \qquad \text{für } k \to \infty.$$

Nach Definition von A und Wahl von $(x_n)_{n\in\mathbb{N}}$ gilt dann:

$$x_{n_k} = Ax_{n_k} + Tx_{n_k} \to x + y \qquad \text{für } k \to \infty.$$

Stetigkeit von A und Eindeutigkeit des Grenzwerts implizieren $A(x + y) = x$, d. h. x liegt im Bild von A. Es folgt $\overline{R(A)} = R(A)$.

Es bleibt die Beschränktheit von $(d_n)_{n\in\mathbb{N}}$ zu zeigen. Angenommen es gibt eine Teilfolge von $(d_n)_{n\in\mathbb{N}}$, welche gegen ∞ konvergiert. (Um die Notation übersichtlich zu halten, bezeichnen wir diese weiterhin mit $(d_n)_{n\in\mathbb{N}}$.) Ohne Einschränkung gilt dann $d_n > 0$ für alle $n \in \mathbb{N}$ und wir können somit

$$(y_n)_{n\in\mathbb{N}} := (x_n/d_n)_{n\in\mathbb{N}}$$

setzen. Dann ist $(y_n)_{n\in\mathbb{N}}$ beschränkt und somit existiert (wegen $T \in K(X)$) eine Teilfolge $(y_{n_k})_{k\in\mathbb{N}}$ und ein $y \in X$ mit $Ty_{n_k} \to y \in X$ für $k \to \infty$. Dann gilt

$$y_{n_k} - Ty_{n_k} = Ay_{n_k} = \frac{Ax_{n_k}}{d_{n_k}} \to 0 \qquad \text{für } k \to \infty,$$

denn $(Ax_n)_{n\in\mathbb{N}}$ konvergiert und $(d_{n_k})_{k\in\mathbb{N}}$ geht gegen unendlich. Aus der Stetigkeit von A folgt $Ay = 0$ und damit gilt $\operatorname{dist}(y, N(A)) = 0$. Weil $\operatorname{dist}(\,\cdot\,, N(A))$ stetig ist, folgt:

$$1 = \operatorname{dist}\big(x_{n_k}, N(A)\big)/d_{n_k} = \operatorname{dist}\big(y_{n_k}, N(A)\big) \to \operatorname{dist}\big(y, N(A)\big) = 0,$$

ein Widerspruch.

(3) Sei A injektiv. Angenommen es existiert ein $x \in X \setminus R(A)$. Für alle $n \in \mathbb{N}$ gilt zunächst

$$A^n x \in R(A^n) \setminus R(A^{n+1}),$$

denn: Falls $A^n x = A^{n+1} y$ für ein $y \in X$ gilt, so folgt $A^n(x - Ay) = 0$ und aus der Injektivität von A folgt dann $x = Ay \in R(A)$, was der Wahl von x widerspricht. Weiter gilt

$$A^n = (\operatorname{Id} - T)^n = \operatorname{Id} - T\,(\dots)$$

und weil T kompakt ist, ist nach Lemma 8.5 auch $T\,(\dots)$ kompakt. Aus (2) folgt dann, dass $R(A^n)$ abgeschlossen ist. Aus den bisherigen Überlegungen folgt $\operatorname{dist}\big(A^n x, R(A^{n+1})\big) > 0$, weswegen wir für alle $n \in \mathbb{N}$ ein $a_{n+1} \in R(A^{n+1})$ mit

$$0 < \|A^n x - a_{n+1}\| \leq 2 \operatorname{dist}\big(A^n x, R(A^{n+1})\big)$$

wählen und

$$y_n := \frac{A^n x - a_{n+1}}{\|A^n x - a_{n+1}\|}$$

setzen können. Für alle $n \in \mathbb{N}$ gilt dann

$$\operatorname{dist}\bigl(y_n, R(A^{n+1})\bigr) \geq \frac{1}{2},$$

denn für alle $y \in R(A^{n+1})$ gilt

$$\|y_n - y\| = \frac{\bigl\|A^n x - (a_{n+1} + \|A^n x - a_{n+1}\|\, y)\bigr\|}{\|A^n x - a_{n+1}\|} \geq \frac{\operatorname{dist}\bigl(A^n x, R(A^{n+1})\bigr)}{\|A^n x - a_{n+1}\|} \geq \frac{1}{2}.$$

Für $m, n \in \mathbb{N}$ mit $m > n$ gilt dann:

$$\|Ty_n - Ty_m\| = \|y_n - \underbrace{(Ay_n + y_m - Ay_m)}_{\in R(A^{n+1})}\| \geq \frac{1}{2}.$$

Das heißt aber, dass $(Ty_n)_{n\in\mathbb{N}}$ keinen Häufungspunkt besitzen kann, im Widerspruch zur Kompaktheit von T. Also war die Annahme falsch und es gilt doch $X = R(A)$.

(4) Sei A surjektiv. Angenommen A ist nicht injektiv. Dann finden wir eine Folge $(x_n)_{n\in\mathbb{N}}$ in X mit $x_1 \in N(A) \setminus \{0\}$ und $Ax_{n+1} = x_n$ für alle $n \in \mathbb{N}$. Für diese gilt

$$x_{n+1} \in N(A^{n+1}) \setminus N(A^n),$$

denn $A^{n+1}x_{n+1} = A^n x_n = Ax_1 = 0$ und $A^n x_{n+1} = A^{n-1}x_n = Ax_2 = x_1 \neq 0$. Es gilt offensichtlich $N(A^n) \subset N(A^{n+1})$ und mit dem Satz vom fast orthogonalen Element (7.3) erhalten wir: Für alle $n \in \mathbb{N}$ exisitiert ein $y_{n+1} \in N(A^{n+1})$ mit $\|y_{n+1}\| = 1$ und $\operatorname{dist}\bigl(y_{n+1}, N(A^n)\bigr) \geq \frac{1}{2}$. Für $n, m \in \mathbb{N}$ mit $n > m$ gilt dann:

$$\|Ty_n - Ty_m\| = \|y_n - \underbrace{(Ay_n + y_m - Ay_m)}_{\in N(A^{n-1})}\| \geq \frac{1}{2},$$

im Widerspruch zur Kompaktheit von T. Also muss doch $N(A) = \{0\}$ gelten.

(5) Sei $n := \dim N(A)$. Wir zeigen nun per Induktion über n, dass $n = \operatorname{codim} R(A)$ gilt. Für $n = 0$, siehe oben. Im Fall $n \geq 1$, wähle $x_0 \in N(A) \setminus \{0\}$ und einen Unterraum $N_0 \subset N(A)$ mit $\dim N_0 = n - 1$ und

$$N(A) = \operatorname{span}\{x_0\} \oplus N_0.$$

Nach (2) und (4) gilt: $R(A)$ ist abgeschlossen und $R(A) \neq X$. Sei $y_0 \in X \setminus R(A)$ und

$$Y_0 := \operatorname{span}\{y_0\} \oplus R(A) \subset X.$$

Nach (dem Beweis von) Korollar 4.16 gibt es ein $x' \in X'$ mit $x'(x_0) = 1$ und $x'|_{N_0} = 0$. Definiere $T_0 x := Tx + x'(x)\, y_0$ und $A_0 := \mathrm{Id} - T_0$, also $A_0 x = Ax - x'(x)\, y_0$. Wir schließen:

$$\begin{aligned} x \in N(A_0) &\iff Ax = x'(x)\, y_0 \iff x \in N(A) \wedge x'(x) = 0 \\ &\iff x \in N_0. \end{aligned}$$

Außerdem gilt $R(A_0) = Y_0$, denn: $A_0 x_0 = -y_0$ und für $y \in R(A)$ mit $Ax = y$ für ein $x \in X$ gilt:

$$A_0\big(x - x'(x)x_0\big) = Ax + 0 = y.$$

Weiter ist T_0 kompakt, also gilt nach Induktionsvoraussetzung:

$$n - 1 = \dim N(A_0) = \operatorname{codim} R(A_0).$$

Es folgt:

$$\operatorname{codim} R(A) = \operatorname{codim} R(A_0) + 1 = n = \dim N(A).$$

■

Satz 8.17 (Spektralsatz für kompakte Operatoren)**.** Sei X ein ∞-dimensionaler Banachraum über $\mathbb{K}$ und sei $T \in K(X)$. Dann gilt:

(a) $0 \in \sigma(T)$

(b) $\sigma(T) \setminus \{0\}$ besteht nur aus Eigenwerten, d. h. $\sigma(T) \setminus \{0\} \subset \sigma_{\mathrm{p}}(T)$.

(c) Es tritt einer der folgenden Fälle ein:

  (1) $\sigma(T) = \{0\}$

  (2) $\sigma(T) \setminus \{0\}$ ist endlich

  (3) $\sigma(T) \setminus \{0\}$ besteht aus einer Folge, die gegen 0 konvergiert

(d) Jeder Eigenwert verschieden von 0 hat einen endlich-dimensionalen Eigenraum.

Beweis. (a) Siehe Bemerkung 8.12 (ii).

(b) Sei $\lambda \in \sigma(T) \setminus \{0\}$. Angenommen λ ist kein Eigenwert von T, dann gilt $N(\mathrm{Id} - T/\lambda) = N(\lambda \mathrm{Id} - T) = \{0\}$. Satz 8.16 liefert: $R(\mathrm{Id} - T/\lambda) = X$. Dies impliziert $\lambda \in \varrho(T)$, im Widerspruch zu $\lambda \in \sigma(T)$.

(c) Zu $\lambda \in \mathbb{K}$ definiere $A_\lambda \coloneqq \lambda \mathrm{Id} - T$. Entweder es gilt einer der ersten beiden Fälle, oder aber es gilt weder $\sigma(T) = \{0\}$ noch $|\sigma(T) \setminus \{0\}| < \infty$. In diesem Fall finden wir aber eine Folge von Eigenwerten $(\lambda_n)_{n\in\mathbb{N}}$ in $\sigma(T) \setminus \{0\}$. Zu $n \in \mathbb{N}$ wähle $e_n \in X \setminus \{0\}$ als Eigenvektor zu λ_n; dann gilt

$$A_{\lambda_n} e_n = 0.$$

Weiter sei $X_n \coloneqq \operatorname{span}\{e_1, \dots, e_n\}$ für alle $n \in \mathbb{N}$. Wir behaupten, dass für alle $n \in \mathbb{N}$ gilt:

$$\dim X_n = n, \quad \text{d. h. } e_1, \dots, e_n \text{ sind linear unabhängig.}$$

Falls $e_1, \dots, e_{n-1}$ linear unabhängig und $e_1, \dots, e_n$ nicht, so gilt für geeignete $\alpha_k \in \mathbb{K}$:

$$e_n = \sum_{k<n} \alpha_k e_k.$$

Dann folgt:

$$0 = (\lambda_n \operatorname{Id} - T)\, e_n = \sum_{k<n} \alpha_k\, (\lambda_n - \lambda_k)\, e_k$$

und somit $\alpha_k = 0$ für alle $k \in \{1, \dots, n-1\}$. Daraus ergibt sich $e_n = 0$, was nicht sein kann.

Jetzt wählen wir mit Hilfe des Satzes vom fast orthogonalen Element (7.3) für alle $n \in \mathbb{N}$ ein $x_n \in X_n$ mit

$$\|x_n\| = 1 \quad \text{und} \quad \operatorname{dist}(x_n, X_{n-1}) \geq \frac{1}{2}.$$

Nach Konstruktion gilt

$$x_n = \alpha_n\, e_n + \tilde{x}_n$$

für $\alpha_n \in \mathbb{K}$ und $\tilde{x}_n \in X_{n-1}$ und damit

$$T(x_n/\lambda_n) = x_n - \frac{1}{\lambda_n} A_{\lambda_n} x_n = x_n - \frac{1}{\lambda_n} A_{\lambda_n} \tilde{x}_n.$$

Weil X_{n-1} aber T-invariant ist, gilt $A_{\lambda_n} \tilde{x}_n \in X_{n-1}$ und für $m < n$ somit $T(x_m/\lambda_m) \in X_{n-1}$. Also erhalten wir für alle $m < n$:

$$\|T(x_n/\lambda_n) - T(x_m/\lambda_m)\| = \|x_n - \underbrace{(\dots)}_{\in X_{n-1}}\| \geq \frac{1}{2}.$$

Also kann die Folge $\big(T(x_n/\lambda_n)\big)_{n\in\mathbb{N}}$ keinen Häufungspunkt besitzen. Weil aber T kompakt ist, kann damit $(x_n/\lambda_n)_{n\in\mathbb{N}}$ keine beschränkte Teilfolge enthalten, d. h. es gilt

$$\|x_n/\lambda_n\| \to \infty \qquad \text{für } n \to \infty.$$

Für alle $n \in \mathbb{N}$ gilt $\|x_n\| = 1$, also muss $(\lambda_n)_{n\in\mathbb{N}}$ eine Nullfolge sein. Damit ist 0 der einzig mögliche Häufungspunkt von $\sigma(T) \setminus \{0\}$. Insbesondere ist $\sigma(T) \setminus B_r(0)$ endlich für alle $r \in \mathbb{R}_{>0}$. Damit ist $\sigma(T) \setminus \{0\}$ (als abzählbare Vereinigung endlicher Mengen) abzählbar.

(d) Dies folgt aus Satz 8.16, denn: Für $\lambda \in \sigma(T) \setminus \{0\}$ ist $\operatorname{Id} - T/\lambda$ ein Fredholm-Operator mit $N(\operatorname{Id} - T/\lambda) = N(\lambda \operatorname{Id} - T)$.

■

8.18 (Fredholm-Alternative)**.** Sei $\lambda \in \mathbb{K} \setminus \{0\}$ und $T \in K(X)$. Dann gilt die *Fredholm-Alternative*: Entweder ist $\lambda x - Tx = y$ eindeutig lösbar für alle $y \in X$, oder aber $\lambda x - Tx = 0$ hat nicht-triviale Lösungen (also von 0 verschiedene Lösungen).

Beweis. Es gilt:

$$\begin{aligned}
\lambda x - Tx = 0 \text{ eindeutig lösbar} &\iff N(\lambda \operatorname{Id} - T) = \{0\} \\
&\iff R(\lambda \operatorname{Id} - T) = X \\
&\iff \lambda x - Tx = y \text{ ist lösbar}
\end{aligned}$$

■

9 Spektralsatz für kompakte normale Operatoren

Wir erinnern an den Begriff der adjungierten Abbildung: Seien X und Y Banachräume und sei $T \in L(X,Y)$, so ist $T' \colon Y' \to X'$ gegeben durch $(T'y')(x) = y'(Tx)$ für alle $x \in X, y' \in Y'$. (Siehe auch Definition 5.15.)

Im Folgenden bezeichnet J_H für einen Hilbertraum H die Isometrie $H \to H'$ aus dem Rieszschen Darstellungssatz (6.5). (*Achtung:* dies ist nicht zu verwechseln mit der Isometrie aus Satz 4.18.)

Definition 9.1 (Hilbertraum-Adjungierte)**.** Seien X, Y Hilberträume. Dann heißt der Operator

$$T^* := J_X^{-1} T' J_Y \qquad \in \ L(Y,X)$$

Hilbertaum-Adjungierte (von T). Sei nun $X = Y$. Gilt $T = T^*$, so nennen wir T *selbstadjungiert.* Gilt $TT^* = T^*T$, so nennen wir T *normal.*

$$\begin{tikzcd} X \arrow[r, shift left, "T"] \arrow[d, "J_X"'] & Y \arrow[l, dashed, shift left, "T^*"] \arrow[d, "J_Y"] \\ X' & Y' \arrow[l, "T'"'] \end{tikzcd}$$

Lemma 9.2. Seien X, Y Banachräume und $T \in L(X,Y)$. Dann ist T^* charakterisiert durch folgende Eigenschaft: für alle $x \in X$, $y \in Y$ gilt

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

Beweis. Seien $x \in X$, $y \in Y$. Dann gilt:

$$\langle x, T^* y \rangle_X = \left\langle x, J_X^{-1} T' J_Y y \right\rangle_X = (T' J_Y y)(x) = (J_Y y)(Tx) = \langle Tx, y \rangle .$$

■

Lemma 9.3. Sei H ein Hilbertraum über $\mathbb{K}$ und $T \in L(H)$ normal. Ist $x \in H$ ein Eigenvektor von T zum Eigenwert $\lambda \in \mathbb{K}$, so ist x auch ein Eigenvektor von T^* zum Eigenwert $\overline{\lambda}$.

Beweis. Weil T normal ist, ist auch $(\lambda \operatorname{Id} - T)$ normal, denn $(\lambda \operatorname{Id} - T)^* = \overline{\lambda} \operatorname{Id} - T^*$, also

$$\begin{aligned} (\lambda \operatorname{Id} - T)^* (\lambda \operatorname{Id} - T) &= |\lambda|^2 \operatorname{Id} - \overline{\lambda} T - \lambda T^* + T^* T \\ &= |\lambda|^2 \operatorname{Id} - \overline{\lambda} T - \lambda T^* + T T^* = (\lambda \operatorname{Id} - T)(\lambda \operatorname{Id} - T)^* . \end{aligned}$$

Es gilt weiter $\|T^*x\| = \|Tx\|$ für alle $x \in H$, denn:

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

Wegen $(\lambda \operatorname{Id} - T)^* = \overline{\lambda} \operatorname{Id} - T^*$, folgt für $x \in H$ mit $Tx = \lambda x$:

$$0 = \|Tx - \lambda x\| = \|(\lambda \operatorname{Id} - T)^* x\| = \|(\overline{\lambda} \operatorname{Id} - T^*)x\|.$$

Dies zeigt die Behauptung.

■

Bemerkung: Im vorangehenden Beweis haben wir gesehen: Ist H ein Hilbertraum und $T \in L(H)$ normal, so gilt:

$$\forall\, x \in H\colon \quad \|T^*x\| = \|Tx\|.$$

Lemma 9.4. Sei H ein Hilbertraum über $\mathbb{C}$ und $T \in L(H)$ normal. Dann gilt:

$$\sup_{\lambda \in \sigma(T)} |\lambda| = \lim_{m \to \infty} \|T^m\|^{1/m} = \|T\|.$$

Beweis. Wir wissen schon, dass

$$\sup_{\lambda \in \sigma(T)} |\lambda| = \lim_{m \to \infty} \|T^m\|^{1/m} \leq \|T\|$$

gilt (Satz 8.11). Wir zeigen nun für alle $m \in \mathbb{N}$:

$$\|T^m\| \geq \|T\|^m,$$

woraus dann durch Wurzelziehen und Grenzwertbildung die Behauptung folgt. Für $m = 1$ ist die Ungleichung klar. Sei $m \in \mathbb{N}_{\geq 2}$. Dann gilt:

$$\begin{aligned}
\|T^m x\|^2 &= \langle T^m x, T^m x\rangle = \left\langle T^{m-1}x, T^*T^m x\right\rangle \\
&\overset{\text{CSU}}{\leq} \|T^{m-1}x\|\,\|T^*T^m x\| = \|T^{m-1}x\|\,\|T^{m+1}x\| \\
&\leq \|T^{m-1}\|\,\|T^{m+1}\|\,\|x\|^2
\end{aligned}$$

Damit gilt also

$$\|T^m\|^2 \leq \|T^{m-1}\|\,\|T^{m+1}\| \leq \|T\|^{m-1}\,\|T^{m+1}\|.$$

Mit $\|T^1\| \geq \|T\|^1$ erhalten wir so induktiv:

$$\|T^{m+1}\| \geq \frac{\|T^m\|^2}{\|T\|^{m-1}} \geq \|T\|^{2m-(m-1)} = \|T\|^{m+1}.$$

■

Satz 9.5. Sei H ein Hilbertraum und $T \in L(H)$. Dann gilt:

(a) Ist T selbstadjungiert und kompakt, so gilt $\sigma(T) \subset \mathbb{R}$.

(b) Ist T normal, so haben verschiedene Eigenwerte zueinander orthogonale Eigenvektoren.

Beweis. (a) Sei $\lambda \in \sigma(T) \setminus \{0\}$. Dann liefert der Spektralsatz (8.17), dass λ ein Eigenwert sein muss. Also existiert ein $x \in H \setminus \{0\}$ mit $Tx = \lambda x$. Somit folgt:

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

Wegen $x \neq 0$ gilt $\langle x, x\rangle > 0$ und damit erhalten wir $\lambda = \overline{\lambda}$, also muss $\lambda \in \mathbb{R}$ gelten.

(b) Seien $\lambda, \mu \in \sigma(T)$ verschiedene Eigenwerte von T mit Eigenvektoren x bzw. y aus H. Dann gilt:

$$\begin{aligned}\lambda \langle x, y\rangle &= \langle \lambda x, y\rangle = \langle Tx, y\rangle \\ &= \langle x, T^* y\rangle = \langle x, \overline{\mu} y\rangle = \mu \langle x, y\rangle.\end{aligned}$$

Wegen $\lambda \neq \mu$ muss also $\langle x, y\rangle = 0$ gelten.

■

Bemerkung: Satz 9.5 (a) gilt auch ohne die Voraussetzung, dass T kompakt ist, ist dann allerdings schwieriger zu beweisen.

Satz 9.6. Sei H ein Hilbertraum über $\mathbb{K}$ und $T \in L(H)$ selbstadjungiert. Dann gilt:

$$\|T\| = \sup_{\substack{x\in H,\\ \|x\|\leq 1}} |\langle Tx, x\rangle|.$$

Beweis. „$\geq$“ folgt aus $\forall\, x \in H\colon\ |\langle Tx, x\rangle| \leq \|Tx\|\,\|x\| \leq \|T\|\,\|x\|\,\|x\|$.

„$\leq$“: Setze $M := \sup_{x\in H,\, \|x\|\leq 1} |\langle Tx, x\rangle|$. Seien $x, y \in H$. Aus $T = T^*$ folgt:

$$\langle T(x+y), x+y\rangle - \langle T(x-y), x-y\rangle = 2\langle Tx, y\rangle + 2\langle Ty, x\rangle = 4\,\mathrm{Re}\langle Tx, y\rangle.$$

Die Parallelogrammidentität (Satz 2.8 (3)) liefert:

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

Daraus folgt für $x, y \in H$ mit $\|x\|, \|y\| \leq 1$:

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

Indem wir y mit einem skalaren Faktor (aus $\mathbb{K}$) multiplizieren, können wir annehmen, dass $\mathrm{Re}\langle Tx, y\rangle = |\langle Tx, y\rangle|$ gilt. Für $Tx \neq 0$ ergibt dies mit $y = Tx/\|Tx\|$ also

$$M \geq |\langle Tx, y\rangle| = \frac{|\langle Tx, Tx\rangle|}{\|Tx\|} = \|Tx\|.$$

Daraus folgt $\|T\| \leq M$.

■

Korollar 9.7. Sei H ein Hilbertraum und $T \in L(H)$ selbstadjungiert. Falls T außerdem positiv semidefinit ist, d. h. es gilt $\langle Tx, x\rangle \geq 0$ für alle $x \in H$, so gilt:

$$\sup_{\lambda \in \sigma(T)} \lambda = \sup_{\substack{x \in H,\\ \|x\| \leq 1}} |\langle Tx, x\rangle|.$$

Beweis. Folgt direkt aus Lemma 9.4 und Satz 9.6.

■

Lemma 9.8. Sei H ein Hilbertraum und $T \in L(H)$.

(a) Gilt $\mathbb{K} = \mathbb{C}$ und ist T normal, so existiert ein $\lambda \in \sigma(T)$ mit $|\lambda| = \|T\|$.

(b) Gilt $\mathbb{K} = \mathbb{R}$ und ist T selbstadjungiert und kompakt, so ist $\|T\|$ oder $-\|T\|$ ein Eigenwert von T.

Beweis. (a) Nach Satz 8.11 existiert ein $\lambda \in \sigma(T)$ mit

$$|\lambda| = \lim_{n\to\infty} \|T^n\|^{1/n}.$$

Die Behauptung folgt nun aus Lemma 9.4.

(b) Nach Satz 9.6 existiert eine Folge $(x_n)_{n\in\mathbb{N}}$ in $\overline{B_1(0)} \subset H$ mit

$$|\langle Tx_n, x_n\rangle| \to \|T\| \qquad \text{für } n \to \infty.$$

Gehe im Folgenden ggf. (vermöge der Kompaktheit von T) zu Teilfolgen über, um die Existenz der Grenzwerte zu erhalten. Setze

$$\lambda := \lim_{n\to\infty} \langle Tx_n, x_n\rangle \quad \text{und} \quad y := \lim_{n\to\infty} Tx_n.$$

Es gilt:

$$\begin{aligned}
\|Tx_n - \lambda x_n\|^2 &= \langle Tx_n - \lambda x_n,\, Tx_n - \lambda x_n\rangle \\
&= \|Tx_n\|^2 - 2\lambda \langle Tx_n, x_n\rangle + \lambda^2 \|x_n\|^2 \\
&\leq 2\lambda^2 - 2\lambda \langle Tx_n, x_n\rangle \\
&\to 0 \qquad \text{für } n \to \infty
\end{aligned}$$

Daher gilt $\lambda x_n \to y$ für $n \to \infty$ und damit

$$Ty = \lambda \lim_{n\to\infty} Tx_n = \lambda y.$$

Wegen $|\lambda| = \|T\|$ folgt die Behauptung, falls $y \neq 0$ gilt. Falls $y = 0$ gilt, so ist $(Tx_n)_{n\in\mathbb{N}}$ eine Nullfolge und somit erhalten wir

$$\|T\| = \lim_{n\to\infty} |\langle Tx_n, x_n\rangle| = 0,$$

d. h. $T = 0$ und dafür ist die Behauptung trivialerweise erfüllt.

■

Theorem 9.9 (Spektralsatz für kompakte, normale bzw. selbstadjungierte Operatoren)**.** Sei H ein Hilbertraum über $\mathbb{K}$ und $T \in K(H)$. Für $\mathbb{K} = \mathbb{C}$ sei T außerdem normal und für $\mathbb{K} = \mathbb{R}$ sei T selbstadjungiert. Dann existiert ein (eventuell endliches) Orthonormalsystem $e_1, e_2, \dots$ sowie eine (eventuell endliche) Nullfolge $\lambda_1, \lambda_2, \dots$ in $\mathbb{K} \setminus \{0\}$, so dass

$$H = N(T) \bigoplus \overline{\operatorname{span}\{e_1, e_2, \dots\}}$$

sowie

$$Tx = \sum_k \lambda_k \langle x, e_k \rangle e_k$$

für alle $x \in H$ gilt. Dabei sind die λ_k die von 0 verschiedenen (aber nicht notwendigerweise unterschiedlichen) Eigenwerte von T und für alle k ist e_k ist ein Eigenvektor zu λ_k. Weiter gilt:

$$\|T\| = \max_k |\lambda_k|.$$

Bemerkung: Vergleiche LinAlg: symmetrische Matrizen sind diagonalisierbar bezüglich einer Orthonormalbasis aus Eigenvektoren.

Beweis. Sei $\mu_1, \mu_2, \dots$ die Folge der paarweise verschiedenen Eigenwerte von T, die nicht verschwinden (dies sind höchstens abzählbar viele nach dem Spektralsatz für kompakte Operatoren 8.17). Sei d_i die (endliche) Dimension des Eigenraums zum Eigenwert μ_i. Definiere nun

$$(\lambda_1, \lambda_2, \lambda_3, \dots) := (\underbrace{\mu_1, \dots, \mu_1}_{d_1\text{-mal}}, \underbrace{\mu_2, \dots, \mu_2}_{d_2\text{-mal}}, \dots).$$

Weil die μ_k eine Nullfolge bilden, gilt dies auch für die λ_k. Zu jedem Eigenraum $N(\mu_i \operatorname{Id} - T)$ wähle eine Orthonormalbasis $\{e_1^i, \dots, e_{d_i}^i\}$ und definiere

$$(e_1, e_2, e_3, \dots) := (e_1^1, \dots, e_{d_1}^1, e_1^2, \dots, e_{d_2}^2, \dots).$$

Nach Satz 9.5 bilden die e_k nun ein Orthonormalsystem und es gilt: $Te_k = \lambda_k e_k$ für alle k. Mit dem gleichen Argument folgt

$$N(T) \perp e_k$$

für alle k (da ein Element aus $N(T) \setminus \{0\}$ Eigenvektor zum Eigenwert 0 ist). Der Raum

$$H_1 := N(T) \bigoplus \overline{\operatorname{span}\{e_1, e_2, \dots\}}$$

ist ein abgeschlossener Unterraum von H. Es bleibt $H_1 = H$ zu zeigen. Wir setzen $H_2 := H_1^\perp$ und behaupten, dass H_2 ein T-invarianter Unterraum ist. Sei dazu $y \in H$ mit $\langle y, e_k \rangle = 0$ für alle k. Dann gilt

$$\langle Ty, e_k \rangle = \langle y, T^* e_k \rangle = \langle y, \overline{\lambda}_k e_k \rangle = \lambda_k \langle y, e_k \rangle = 0,$$

also $Ty \perp H_1$. Analog zeigt man $Ty \perp N(T)$ für $y \in N(T)^\perp$. Damit können wir $T_2 := T|_{H_2}$ als Operator aus $K(H_2)$ auffassen. Angenommen T_2 ist nicht der Nulloperator. Dann gilt $\|T_2\| \neq 0$ und somit sichert Lemma 9.8 die Existenz eines Spektralwerts $\lambda \in \mathbb{K} \setminus \{0\}$, welcher nach dem Spektralsatz für kompakte Operatoren (Satz 8.17) ein Eigenwert sein muss, d. h. es gibt außerdem ein $x \in H_2 \setminus \{0\}$ mit $T_2 x = \lambda x$. Daraus folgt aber auch $\lambda \in \sigma(T)$ und somit $x \in \operatorname{span}\{e_1, e_2, \dots\} \subset H_2^\perp$. Also ergibt sich

$$x \in H_2 \cap H_2^\perp = \{0\},$$

ein Widerspruch. Die Annahme war also falsch und es gilt doch $T_2 = 0$, und daher auch $H_2 \subset N(T) \subset H_2^\perp$, also $H_2 = \{0\}$. Dies zeigt den ersten Teil der Behauptung. Sei nun $x \in H$. Dann gibt es also ein $y \in N(T)$, so dass

$$x = y + \sum_k \langle x, e_k \rangle \, e_k$$

gilt (für den rechten Summanden, siehe Satz 6.17). Aus der Stetigkeit von T folgt:

$$Tx = Ty + \sum_k \langle x, e_k \rangle \, Te_k = \sum_k \lambda_k \langle x, e_k \rangle \, e_k.$$

Im Beweis von Satz 8.11 haben wir gesehen, dass stets $\sup_{\lambda \in \sigma(T)} |\lambda| \leq \|T\|$ gilt. Aus Lemma 9.8 folgt, dass dieses Supremum unter den gegebenen Voraussetzungen sowohl im Fall $\mathbb{K} = \mathbb{C}$ als auch im Fall $\mathbb{K} = \mathbb{R}$ angenommen wird. Daraus erhalten wir die letzte Behauptung.

$\blacksquare$

10 L^p-Räume

10.1 (Einige Begriffe und Resultate über Maß- und Integrationstheorie, die jeder Bürger wissen sollte)**.** Maßraum, σ-Algebra, Maß, messbare Menge/Funktion.

Sei (Ω, S, μ) ein Maßraum (also Ω eine Menge, $S \subset \mathcal{P}(\Omega)$ eine σ-Algebra und μ ein Maß).

Definition: Ω heißt σ-finit, falls eine Folge $(\Omega_n)_{n\in\mathbb{N}}$ in S existiert, so dass $\Omega = \bigcup_{n=1}^{\infty} \Omega_n$ gilt und für alle $n \in \mathbb{N}$ das Maß von Ω_n endlich ist, d h. $\mu(\Omega_n) < \infty$.

Definition:

(i) Die Menge $L^1(\Omega, \mu)$ (kurz auch $L^1(\Omega)$ oder nur L^1) bezeichnet den Raum aller integrierbaren Funktionen $\Omega \to \mathbb{R}$.

(ii) $\|f\|_{L^1} = \|f\|_1 = \int_\Omega |f| \, \mathrm{d}\mu = \int |f|$

Bemerkung:

(i) Identifiziere Funktionen, die sich nur auf einer Nullmenge unterscheiden.

(ii) Wir sagen eine Eigenschaft gilt fast überall (f. ü.), falls eine Nullmenge N existiert, so dass die betrachtete Eigenschaft für alle $x \in \Omega \setminus N$ gilt.

Satz (Satz von Beppo-Levi/über monotone Konvergenz)**.** Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge in $L^1(\Omega, \mu)$ mit

(a) $f_1(x) \le f_2(x) \le f_3(x) \le \cdots$ f. ü.

(b) $\sup_{n\in\mathbb{N}} \int_\Omega f_n < \infty$.

Dann konvergiert $f_n(x)$ für $n \to \infty$ fast überall in Ω gegen einen endlichen Grenzwert, den wir mit $f(x)$ bezeichnen. Es gilt

$$\|f_n - f\|_1 \to 0 \qquad \text{für } n \to \infty.$$

Satz (Satz von Lebesgue/über dominierte Konvergenz)**.** Sei $g \in L^1(\Omega, \mu)$ und sei $f \colon \Omega \to \mathbb{R}$ messbar sowie $(f_n)_{n\in\mathbb{N}}$ eine Folge messbarer Funktionen. Es gelte $|f_n| \le g$ fast überall für alle $n \in \mathbb{N}$ und $f_n \to f$ punktweise fast überall für $n \to \infty$. Dann gilt auch $f, f_n \in L^1(\Omega, \mu)$ (für alle $n \in \mathbb{N}$) und es gilt $f_n \to f$ in $L^1(\Omega, \mu)$ für $n \to \infty$.

Definition 10.2. Sei $p \in [1, \infty)$. Wir definieren

$$L^p(\Omega) := \big\{ f \colon \Omega \to \mathbb{R} \;\big|\; f \text{ ist messbar und } |f|^p \in L^1(\Omega) \big\}$$

und

$$\|f\|_{L^p} := \|f\|_p := \bigg(\int_\Omega |f|^p \,\mathrm{d}\mu \bigg)^{1/p}.$$

Bemerkung: Wir sehen später, dass $\|f\|_{L^p}$ tatsächlich eine Norm ist.

Definition 10.3. Wir definieren

$$L^\infty(\Omega) := \big\{ f \colon \Omega \to \mathbb{R} \;\big|\; f \text{ ist messbar und es existiert ein } c \in \mathbb{R}_{>0} \text{ mit } |f(x)| \le c \text{ f. ü.} \big\}$$

und

$$\|f\|_{L^\infty} := \|f\|_\infty := \inf\big\{ c \in \mathbb{R}_{>0} \;\big|\; |f(x)| \le c \text{ f. ü.} \big\}.$$

Bemerkungen 10.4.

(i) Für $\Omega = \mathbb{N}$ und das Zählmaß μ auf $\mathbb{N}$ gilt $\ell^p = L^p(\mathbb{N}, \mu)$.

(ii) Für $f \in L^\infty(\Omega)$ gilt $|f(x)| \le \|f\|_\infty$ f. ü.

Beweis von (ii). Sei $f \in L^\infty(\Omega)$. Dann existiert eine Folge $(c_n)_{n\in\mathbb{N}}$ in $\mathbb{R}_{>0}$ mit

$$c_n \to \|f\|_\infty \quad \text{und} \quad \forall\, n \in \mathbb{N}\colon \quad |f(x)| \le c_n \quad \text{f. ü..}$$

Für $n \in \mathbb{N}$ sei E_n eine Nullmenge mit $|f(x)| \le c_n$ für alle $x \in \Omega \setminus E_n$. Setze $E := \bigcup_{n\in\mathbb{N}} E_n$. Dann ist (bekannterweise) auch E eine Nullmenge und es gilt:

$$\forall\, n \in \mathbb{N}\; \forall\, x \in \Omega \setminus E\colon \quad |f(x)| \le c_n.$$

Es folgt $|f(x)| \le \|f\|_\infty$ für alle $x \in \Omega \setminus E$.

∎

Notation: Zu $p \in [1, \infty]$ bezeichne $p' \in [1, \infty]$ den *konjugierten Exponenten* mit

$$\frac{1}{p} + \frac{1}{p'} = 1.$$

D. h. es gilt $p' = p/(p-1)$ für $p \in (1, \infty)$, $p' = \infty$ für $p = 1$ und $p' = 1$ für $p = \infty$.

Theorem 10.5 (Hölder'sche Ungleichung)**.** Sei $p \in [1, \infty]$. Für $f \in L^p(\Omega)$ und $g \in L^{p'}(\Omega)$ gilt:

$$fg \in L^1(\Omega) \quad \text{und} \quad \|fg\|_1 \le \|f\|_p \, \|g\|_{p'}.$$

Beweisskizze. Geht analog zum Beweis bei ℓ^p: vgl. Satz 2.14. Ersetze dabei jeweils x_k, y_k durch $f(x), g(x)$ und Summen durch Integrale.

$\blacksquare$

Bemerkungen:

(i) Es gibt eine Erweiterung der Höler'schen Ungleichung: Sei $k \in \mathbb{N}$, seien $p_1, \dots, p_k \in [1, \infty]$k und für $i \in \{1, \dots, k\}$ sei $f_i \in L^{p_i}(\Omega)$. Weiter sei $p \in [1, \infty]$ mit

$$\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} + \dots + \frac{1}{p_k}.$$

Dann gilt für $f \coloneqq f_1 \cdots f_k$:

$$f \in L^p(\Omega) \quad \text{und} \quad \|f\|_p \leq \|f\|_{p_1} \cdots \|f\|_{p_k}.$$

(ii) Insbesondere gilt für $f \in L^p \cap L^q$ mit $1 \leq p \leq q \leq \infty$ auch $f \in L^r$ für alle $r \in [p, q]$. Weiter gilt

$$\|f\|_r \leq \|f\|_p^\alpha \|f\|_q^{1-\alpha}$$

für

$$\frac{1}{r} = \frac{\alpha}{p} + \frac{1-\alpha}{q}.$$

Diese Ungleichung nennt man *Interpolationsungleichung*, welche wichtig ist, um Funktionen in L^p-Räumen zu kontrollieren.

Satz 10.6. Sei $p \in [1, \infty]$. Dann ist $L^p(\Omega)$ ein Vektorraum und $\|\cdot\|_{L^p}$ ist eine Norm.

Beweisskizze. Die Fälle $p \in \{1, \infty\}$ sind klar. Sei also $p \in (1, \infty)$. Wir gehen vor wie im Fall für ℓ^p (Satz 2.15). (Im Wesentlichen ist die $\triangle$-Ungleichung zu zeigen, alles andere ist klar.)

$\blacksquare$

Satz 10.7 (Fischer-Riesz)**.** Sei $p \in [1, \infty]$. Dann ist $L^p(\Omega)$ ein Banachraum.

Beweis. Fall $p = \infty$. Sei $(f_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in L^∞. Für jedes $k \in \mathbb{N}$ existiert ein $N_k \in \mathbb{N}$ mit

$$\forall\, m, n \in \mathbb{N}_{\geq N_k}\colon \quad \|f_m - f_n\|_\infty \leq \frac{1}{k}.$$

Also existiert für jedes $k \in \mathbb{N}$ eine Nullmenge E_k mit

$$\forall\, m, n \in \mathbb{N}_{\geq N_k}\ \forall\, x \in \Omega \setminus E_k\colon \quad |f_m(x) - f_n(x)| \leq \frac{1}{k}.$$

Setze dann $E \coloneqq \bigcup_{k\in\mathbb{N}} E_k$, dann ist auch $E \subset \Omega$ eine Nullmenge. Da $(f_n(x))_{n\in\mathbb{N}}$ für $x \in \Omega \setminus E$ eine Cauchy-Folge in $\mathbb{R}$ ist, existiert ein $f(x) \in \mathbb{R}$ mit $f_n(x) \to f(x)$ für $n \to \infty$. Aus der obigen Ungleichung folgt für $m \to \infty$:

$$\forall\, n \in \mathbb{N}_{\geq N_k}\ \forall\, x \in \Omega \setminus E\colon \quad |f(x) - f_n(x)| \leq \frac{1}{k}.$$

Daraus ergibt sich $f \in L^\infty(\Omega)$ und $\|f - f_n\|_\infty \leq 1/k$ für alle $n \in \mathbb{N}_{\geq N_k}$. Daraus folgt

$$f_n \to f \quad \text{in } L^\infty(\Omega) \text{ für } n \to \infty.$$

Fall $p \in [1, \infty)$. Sei $(f_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $L^p(\Omega)$. Es genügt zu zeigen, dass eine Teilfolge dieser Folge konvergiert. Sei $(f_{n_k})_{k\in\mathbb{N}}$ eine Teilfolge, so dass für alle $k \in \mathbb{N}$ gilt:

$$\|f_{n_{k+1}} - f_{n_k}\|_p \leq \frac{1}{2^k}.$$

Im Folgenden schreiben wir wieder einfach f_k für f_{n_k}. Wir behaupten nun, dass $(f_k)_{k\in\mathbb{N}}$ in $L^p(\Omega)$ konvergiert. Für alle $n \in \mathbb{N}$ sei

$$g_n(x) := \sum_{k=1}^{n} |f_{k+1}(x) - f_k(x)|.$$

Dann gilt $\|g_n\|_p \leq 1$ für alle $n \in \mathbb{N}$, was aus der obigen Ungleichung und dem Konvergenzverhalten der geometrischen Reihe folgt. Der Satz von der monotonen Konvergenz liefert ein $g \in L^p(\Omega)$ mit

$$g_n(x) \to g(x) \quad \text{f. ü.} \qquad \text{für } n \to \infty$$

und für alle $m, n \in \mathbb{N}$ mit $m \geq n \geq 2$ gilt:

$$\begin{aligned} |f_m(x) - f_n(x)| &\leq |f_m(x) - f_{m-1}(x)| + \cdots + |f_{n+1}(x) - f_n(x)| \\ &\leq g(x) - g_{n-1}(x) \to 0 \quad \text{f. ü.} \qquad \text{für } n \to \infty. \end{aligned}$$

Dies zeigt aber, dass $(f_n(x))_{n\in\mathbb{N}}$ f. ü. konvergiert mit Grenzwert $f(x)$. Es gilt

$$|f(x) - f_n(x)| \leq g(x) \quad \text{f. ü.},$$

insbesondere folgt also:

$$f = \underbrace{f - f_n}_{\in L^p} + \underbrace{f_n}_{\in L^p} \in L^p(\Omega).$$

Der Satz von Lebesgue liefert nun:

$$\|f - f_n\|_p \to 0 \qquad \text{für } n \to \infty,$$

denn:

$$\begin{gathered} |f(x) - f_n(x)|^p \to 0 \quad \text{f. ü.} \qquad \text{für } n \to \infty \\ \text{und} \quad |f - f_n|^p \leq g^p \in L^1(\Omega) \end{gathered}$$

■

Satz 10.8. Sei $p \in [1, \infty]$ und $(f_n)_{n\in\mathbb{N}}$ eine Folge in $L^p(\Omega)$. Sei weiter $f \in L^p(\Omega)$ mit $\|f_n - f\|_p \to 0$ für $n \to \infty$. Dann existiert eine Teilfolge $(f_{n_k})_{k\in\mathbb{N}}$ und eine Funktion $h \in L^p(\Omega)$, so dass gilt:

(a) $f_{n_k}(x) \to f(x)$ f. ü. für $k \to \infty$

(b) $\forall\, k \in \mathbb{N}\colon \quad |f_{n_k}| \le h(x)$ f. ü.

Beweisskizze. Wähle die Teilfolge wie im vorangehenden Beweis und zeige die gewünschten Aussagen mit ähnlichen Argumenten wie dort ...

■

Das Ziel ist es nun, den Dualraum von $L^p(\Omega)$ zu beschreiben. Wir brauchen dazu etwas Maßtheorie.

Definition 10.9. Sei Ω eine Menge und S eine σ-Algebra auf Ω. Eine σ-additive Abbildung $\mu\colon S \to \mathbb{R}$ heißt *signiertes Maß* und eine σ-additive Abbildung $\mu\colon S \to \mathbb{C}$ heißt *komplexes Maß*.

Bemerkungen:

(i) Ist μ ein signiertes oder komplexes Maß, so gilt $\mu(\emptyset) = 0$ wegen

$$\mu(\emptyset) = \mu(\emptyset \cup \emptyset) = \mu(\emptyset) + \mu(\emptyset).$$

(ii) Wir betrachten $\Omega \subset \mathbb{R}^d$ und das Lebesgue-Maß $\mathcal{L}^d$. Sei $f \in L^1(\Omega)$. Dann definiert

$$\mu\colon S \to \mathbb{R}, \quad E \mapsto \int_E f \,\mathrm{d}\mathcal{L}^d$$

ein signiertes Maß. Im Gegensatz zu Dichtefunktionen aus der Wahrscheinlichkeitstheorie kann f hier auch negative Werte annehmen.

Definition 10.10. Sei μ ein signiertes oder komplexes Maß auf S. Für alle $E \in S$ setze

$$|\mu|(E) := \sup\left\{ \sum_{k=1}^{n} |\mu(E_k)| \;\middle|\; \begin{array}{c} n \in \mathbb{N},\ E_1, \dots, E_n \text{ aus } S \text{ und} \\ \text{paarweise disjunkt mit } E = \bigcup_{k=1}^{n} E_k \end{array} \right\}.$$

Dann heißt $|\mu|$ *Variationsmaß zu* μ. Weiter sei $\|\mu\|_{\mathrm{var}} := |\mu|(\Omega)$ die *Totalvariation von* μ. Wir nennen μ *beschränkt*, falls $\|\mu\|_{\mathrm{var}} < \infty$ gilt.

Bemerkung: Aus der Definition ergibt sich leicht, dass das Variationsmaß additiv ist.

Satz 10.11 (Satz von Radon-Nikodym)**.** Sei (Ω, S, μ) ein σ-finiter Maßraum und sei $\nu\colon S \to \mathbb{K}$ ein signiertes oder komplexes Maß mit $\|\nu\|_{\mathrm{var}} < \infty$. Weiter sei ν *absolut stetig bezüglich* μ, d. h. es gilt

$$\forall\, E \in S\colon \quad \mu(E) = 0 \implies \nu(E) = 0.$$

Dann gibt es genau eine Funktion $f \in L^1(\Omega, \mu)$ mit

$$\forall\, E \in S\colon \quad \nu(E) = \int_E f \,\mathrm{d}\mu.$$

Einen Beweis findet man beispielsweise im Buch von Alt oder in vielen Büchern zur Maßtheorie; beispielsweise bei Halmos, *Measure Theory.*

Bemerkung: In der Situation von Satz 10.11 nennt man die Funktion f die *Radon-Nikodym-Ableitung von ν bezüglich μ*, welche auch oft mit $\frac{\mathrm{d}\nu}{\mathrm{d}\mu}$ bezeichnet wird.

Beispiel: Wir betrachten $\mathcal{L}^d$, das Lebensgue-Maß, und δ_0, das Dirac-Maß bei 0, d. h.

$$\delta_0(E) = \begin{cases} 1, & 0 \in E \\ 0, & 0 \notin E. \end{cases}$$

Dann ist δ_0 *nicht* absolut stetig bezüglich $\mathcal{L}^d$.

Bemerkung: Fast alle Aussagen der Integrationstheorie gelten entsprechend für Funktionen $f\colon \Omega \to \mathbb{C}$. Für $f = f_1 + i f_2$ gilt beispielsweise

$$\int_\Omega f = \int_\Omega f_1 + i \int_\Omega f_2$$

und (mit Hilfe der Hölder'schen Ungleichung)

$$\left| \int_\Omega f \right| \leq \int_\Omega |f|.$$

Für $p \in [1, \infty]$ seien die Funktionen in $L^p(\Omega)$ im Folgenden stets $\mathbb{K}$-wertig mit $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

Satz 10.12 (Dualraum von $L^p(\Omega)$)**.** Sei (Ω, S, μ) ein σ-finiter Maßraum und sei $p \in [1, \infty)$ mit konjugiertem Exponenten $p' \in (1, \infty]$. Dann ist

$$\begin{aligned} J\colon L^{p'}(\Omega) &\to \big(L^p(\Omega)\big)' \\ f &\mapsto \Big(g \mapsto \int_\Omega g\,\overline{f}\,\mathrm{d}\mu\Big) \end{aligned}$$

ein konjugiert linearer, isometrischer Isomorphismus.

Beweis. Aus der Hölder'schen Ungleichung (Theorem 10.5) folgt:

$$\|Jf\|_{p'} \leq \|f\|_{p'}.$$

Jetzt wählen wir g so, dass $g\overline{f} = |f|^{p'}$ gilt, d. h. für $x \in \Omega$ sei $g(x) \coloneqq |f|^{p'-2} f$, falls $f(x) \neq 0$, und $g(x) \coloneqq 0$ sonst. Dann erhalten wir:

$$(Jf)(g) = \int_\Omega |f|^{p'} = \|f\|_{p'}^{p'}.$$

Außerdem gilt

$$\|g\|_p = \Big(\int_\Omega |g|^p\Big)^{1/p} = \|f\|_{p'}^{p'/p}.$$

Damit hat $g/\|f\|_{p'}^{p'/p} \in L^p(\Omega)$ Norm 1 und es folgt $\|Jf\| \geq \|f\|_{p'}$. Ingesamt erhalten wir also: J eine ist eine Isometrie und somit insbesondere injektiv.

Wir zeigen nun, dass J auch surjektiv ist. Sei dazu $F \in (L^p(\Omega))'$. Weil der Ausgangsraum σ-finit ist, finden wir eine Folge $(\Omega_n)_{n\in\mathbb{N}}$ in S mit $\Omega_n \subset \Omega_{n+1}$ für alle $n \in \mathbb{N}$, $\Omega = \bigcup_{n\in\mathbb{N}} \Omega_n$ und $\mu(\Omega_n) < \infty$ für alle $n \in \mathbb{N}$. Die Strategie ist nun, aus F ein Maß zu bauen, indem wir F bei charakteristischen Funktionen auswerten. Definiere für alle $k \in \mathbb{N}$:

$$\nu_k \colon S \to \mathbb{C}, \quad E \mapsto F(\chi_{E\cap\Omega_k}).$$

Wir zeigen für alle $k \in \mathbb{N}$:

(i) ν_k ist σ-additiv,

(ii) $\|\nu_k\|_{\mathrm{var}} < \infty$,

(iii) ν_k ist absolut stetig bezüglich μ.

Zu (i): Dass ν_k endlich additiv ist, folgt sofort aus den Eigenschaften charakteristischer Funktionen und der Linearität von F. Sei nun $(E_i)_{i\in\mathbb{N}}$ eine Folge paarweise disjunkter Mengen aus S und sei $E \coloneqq \bigcup_{i\in\mathbb{N}} E_i$. Sei außerdem $E'_j \coloneqq \bigcup_{i=1}^{j} E_i$ für alle $j \in \mathbb{N}$. Dann folgt aus der σ-Stetigkeit des Maßes μ:

$$\chi_{E'_j\cap\Omega_k} \to \chi_{E\cap\Omega_k} \quad \text{in } L^p(\Omega) \qquad \text{für } j \to \infty.$$

Weil F stetig ist, folgt:

$$\sum_{i=1}^{\infty} \nu_k(E_i) = \lim_{j\to\infty} \sum_{i=1}^{j} \nu_k(E_i) = \lim_{j\to\infty} \nu_k(E'_j) = \lim_{j\to\infty} F(\chi_{E'_j\cap\Omega_k}) = F(\chi_{E\cap\Omega_k}) = \nu_k(E).$$

Zu (ii): Seien $E_1, \dots, E_\ell \in S$ paarweise disjunkt mit $\Omega = E_1 \cup \dots \cup E_\ell$. Ohne Einschränkung sei keine der Menge $E_1, \dots, E_\ell$ eine Nullmenge bezüglich ν_k. Für alle $i \in \{1, \dots, \ell\}$ sei $\sigma_i \coloneqq \overline{\nu_k(E_i)}/|\nu_k(E_i)|$. Dann gilt:

$$\sum_{i=1}^{\ell} |\nu_k(E_i)| = \sum_{i=1}^{\ell} \sigma_i\, \nu_k(E_i) = F\Big(\underbrace{\sum_{i=1}^{\ell} \sigma_i \chi_{E_i\cap\Omega_k}}_{=:g} \Big) \le \|F\|\, \|g\|_p < \infty,$$

denn

$$\|g\|_p^p = \int_\Omega |g|^p = \int_\Omega \textstyle\sum_{i=1}^{\ell} |\sigma_i|^p \chi_{E_i\cap\Omega_k} = \displaystyle\int_\Omega \chi_{\Omega\cap\Omega_k} = \mu(\Omega_k) < \infty.$$

Zu (iii): Sei $N \in S$ eine Nullmenge bezüglich μ. Dann verschwindet $\chi_{N\cap\Omega_k}$ f. ü. auf Ω, d. h. $\chi_{N\cap\Omega_k} = 0 \in L^p(\Omega)$. Es folgt:

$$\nu_k(N) = F(0) = 0.$$

Der Satz von Radon-Nikodym (10.11) liefert nun für alle $k \in \mathbb{N}$ ein $f_k \in L^1(\Omega)$, so dass gilt:

$$\forall\, k \in \mathbb{N}\; \forall\, E \in S\colon \quad \nu_k(E) = \int_E f_k \,\mathrm{d}\mu.$$

Es gilt dann für alle $k \in \mathbb{N}$:

$$\begin{aligned} f_k &= 0 \quad \mu\text{-f. ü.} \quad \text{auf } \Omega \setminus \Omega_k \qquad \text{und} \\ f_{k+1} &= f_k \quad \mu\text{-f. ü.} \quad \text{auf } \Omega_k. \end{aligned}$$

Sei $f\colon \Omega \to \mathbb{C}$ definiert durch

$$f(x) := \overline{f_k(x)}, \quad \text{falls } k = \min\{k' \in \mathbb{N} \mid x \in \Omega_{k'}\}.$$

Aus den obigen Eigenschaften der Folge $(f_k)_{k\in\mathbb{N}}$ folgt dann:

$$\forall\, E \in S\colon \quad F(\chi_E) = \int_E \overline{f}\, \mathrm{d}\mu = \int_\Omega \chi_E \overline{f}\, \mathrm{d}\mu.$$

Außerdem gilt

$$F(g) = \int_\Omega g\overline{f}\, \mathrm{d}\mu$$

für alle $g \in L^\infty(\Omega)$ mit $g|_{\Omega\setminus\Omega_k} = 0$ für ein $k \in \mathbb{N}$, denn solche Funktionen lassen sich gleichmäßig durch Treppenfunktionen approximieren (siehe Maßtheorie). Der Raum

$$\big\{g \in L^\infty(\Omega) \;\big|\; \exists\, k \in \mathbb{N}\colon\; g|_{\Omega\setminus\Omega_k} = 0\big\}$$

liegt dicht in $L^p(\Omega)$, denn für ein $g \in L^p(\Omega)$ gilt

$$\chi_{\Omega_k}\chi_{\{|g|\le m\}}g \;\to\; g \quad \text{in } L^p(\Omega) \qquad \text{für } k, m \to \infty,$$

wie man sich mithilfe des Satzes über monotone Konvergenz überlegen kann.

Es bleibt zu zeigen: $f \in L^{p'}(\Omega)$. Denn dann stimmen die stetigen Funktionen F und $J(f)$ auf einem dichten Teilraum von $L^p(\Omega)$ überein, müssen also schon gleich sein.

Für $k, m \in \mathbb{N}$ und $q \in [1, \infty)$ seien

$$\Omega_{km} := \Omega_k \cap \{|f| \le m\} \qquad \text{und} \qquad g_{km} := \chi_{\Omega_{km}}|f|^{q-2}f.$$

Dann gilt:

$$\int_{\Omega_{km}} |f|^q\, \mathrm{d}\mu = F(g) \le \|F\|\,\|g\|_p = \|F\| \Big(\int_{\Omega_{km}} |f|^{p(q-1)}\Big)^{1/p}. \tag{$\star$}$$

1. Fall: $p > 1$. Wir setzen $q = p'$ und erhalten mit $p(q-1) = p(p'-1) = p'$ aus ($\star$):

$$\Big(\int_\Omega |f|^{p'}\, \mathrm{d}\mu\Big) \le \|F\| \Big(\int_{\Omega_{km}} |f|^{p'}\, \mathrm{d}\mu\Big)^{1/p} \quad \text{bzw.} \quad \Big(\int_{\Omega_{km}} |f|^{p'}\, \mathrm{d}\mu\Big)^{1/p'} \le \|F\|$$

Im Grenzwert $k, m \to \infty$ folgt dann aus dem Satz über monotone Konvergenz:

$$\|f\|_{p'} \le \|F\| \quad \text{und insbesondere} \quad f \in L^{p'}(\Omega)$$

2. Fall: $p = 1$. Für alle $q \in \mathbb{N}$ erhalten wir induktiv aus ($\star$):

$$\int_{\Omega_{km}} |f|^q\; \mathrm{d}\mu \;\le\; \|F\| \int_{\Omega_{km}} |f|^{q-1}\; \mathrm{d}\mu \;\le\; \|F\|^q \int_{\Omega_{km}} 1\; \mathrm{d}\mu \;=\; \|F\|^q\, \mu(\Omega_{km}).$$

Nun kann man zeigen, dass $\|f\|_{L^q(\Omega_{km})} \to \|f\|_{L^\infty(\Omega_{km})}$ für $q \to \infty$ gilt. Zusammen mit der obigen Ungleichung folgt daraus

$$\|f\|_{L^\infty(\Omega_{km})} = \lim_{q\to\infty} \|f\|_{L^q(\Omega_{km})} \leq \lim_{q\to\infty} \|F\| \left(\mu(\Omega_{km})\right)^{1/q} = \|F\|$$

(falls $\mu(\Omega_{km}) > 0$). Indem wir geeignete Nullmengen bezüglich μ betrachten, erhalten wir dann aus $\|f\|_{L^\infty(\Omega_{km})} \leq \|F\|$ für alle $k, m \in \mathbb{N}$ auch $\|f\|_{L^\infty(\Omega)} \leq \|F\|$ und damit $f \in L^{p'}(\Omega)$. ■

Bemerkungen:

(i) Für $p \in [1, \infty)$ hat $L^p(\mathbb{R}^n)$ den Dualraum $L^{p'}(\mathbb{R}^n)$, da $\mathbb{R}^n$ vermöge $\mathbb{R}^n = \bigcup_{n\in\mathbb{N}} B_n(0)$ ein σ-finiter Raum ist.

(ii) Für $p \in (1, \infty)$ ist die Voraussetzung, dass der Raum σ-finit ist, nicht notwendig. (Siehe beispielsweise Alt.)

(iii) $L^1(\Omega)$ und $(L^\infty(\Omega))'$ sind i. A. nicht isomorph.

Satz 10.13. Für $p \in (1, \infty)$ ist $L^p(\Omega)$ reflexiv. Es gilt weiter: Ist $(f_k)_{k\in\mathbb{N}}$ eine beschränkte Folge in $L^p(\Omega)$, so gibt es eine Teilfolge $(f_{k_i})_{i\in\mathbb{N}}$ und ein $f \in L^p(\Omega)$ mit

$$\forall\, g \in L^{p'}(\Omega)\colon \qquad \int_\Omega g f_{k_i} \,\mathrm{d}\mu \to \int_\Omega g f \,\mathrm{d}\mu \qquad \text{für } i \to \infty.$$

Beweis. Die Isometrien

$$J_p\colon L^p \to \left(L^{p'}\right)' \quad \text{und} \quad J_{p'}\colon L^{p'} \to \left(L^p\right)'$$

aus Satz 10.12 haben die Eigenschaft

$$\forall\, f \in L^p,\, g \in L^{p'}\colon \quad \overline{(J_{p'}g)(f)} = \int_\Omega \overline{f} g \,\mathrm{d}\mu = (J_p f)(g).$$

Sei $f'' \in (L^p)''$. Dann ist durch

$$g \mapsto f'(g) := \overline{f''(J_{p'}g)}$$

ein Funktional $f' \in (L^{p'})'$ gegeben. Setze jetzt

$$f := J_p^{-1} f' \quad \in L^p.$$

Für $g \in L^{p'}$ gilt nun:

$$f'(g) = (J_p f)(g) = \overline{(J_{p'}g)(f)} = \overline{(J_{L^p} f)(J_{p'}g)},$$

wobei $J_{L^p}\colon L^p \to (L^p)''$ die Isometrie aus Satz 4.18 ist. Somit gilt:

$$f''(J_{p'}g) = (J_{L^p} f)(J_{p'}g)$$

für alle $g \in L^{p'}$. Da $J_{p'}$ surjektiv ist, folgt $f'' = J_{L^p} f$, womit die Reflexivität von L^p gezeigt ist. Die zweite Behauptung folgt Satz 7.11 und Satz 10.12.

■

Bemerkung: Im reellen Fall (also $\mathbb{K} = \mathbb{R}$) gilt insbesondere

$$J_{L^p}^{-1} = J_p^{-1}(J_{p'})',$$

wobei $(J_{p'})' \colon (L^p)'' \to \left(L^{p'}\right)'$ der adjungierte Operator zu $J_{p'}$ ist (Definition 5.15).

Bemerkung: Im Allgemeinen ist $(L^\infty)'$ „größer" als L^1. Man kann $(L^\infty)'$ als Raum von Maßen interpretieren. Im Allgemeinen ist weder L^1 noch L^∞ reflexiv.

Satz 10.14. (i) Sei $S \subset \mathbb{R}^n$ kompakt. Dann ist $C^0(S)$ separabel.

(ii) Für $p \in [1, \infty)$ ist $L^p(\mathbb{R}^n)$ separabel und $C_0^0(\mathbb{R}^n)$ liegt dicht in $L^p(\mathbb{R}^n)$.

(iii) Sei $S \subset \mathbb{R}^n$ Lebesgue-messbar und $\mathcal{L}^n(S) > 0$. Dann ist $L^\infty(S)$ nicht separabel.

Beweis. (i) Idee: Überdecke S mit einem ε-Gitter. Sei $\varepsilon \in \mathbb{R}_{>0}$. Für $z \in \varepsilon\mathbb{Z}^n$ sei

$$Q_{\varepsilon,z} := \{x \in \mathbb{R}^n \mid z_i \leq x_i \leq z_i + \varepsilon\}$$

und

$$M_\varepsilon := \{z \in \varepsilon\mathbb{Z}^n \mid Q_{\varepsilon,z} \cap S \neq \emptyset\}.$$

Da S kompakt ist, besteht M_ε stets aus endlich vielen Punkten. Zu $y \in M_\varepsilon$ wähle $x_{\varepsilon,y} \in S$ mit

$$\|x - x_{\varepsilon,y}\|_\infty \leq \varepsilon.$$

Zu $f \in C^0(S)$ definiere

$$g_\varepsilon(y) := f(x_{\varepsilon,y})$$

und setze g_ε durch multilineare Interpolation fort, d. h. für $x = z + \varepsilon \sum_{i=1}^n t_i e_i \in Q_{\varepsilon,z}$ mit $z \in M_\varepsilon$ und $t_i \in [0,1]$ setze

$$g_\varepsilon(x) := \sum_{\gamma \in \{0,1\}^n} \Big(\prod_{j,\, \gamma_j = 0} (1 - t_j) \prod_{j,\, \gamma_j = 1} t_j \Big) g_\varepsilon(z + \varepsilon\gamma).$$

(In jedem Quader des ε-Gitters ist g_ε ein Polynom n-ten Grades in den t_j. Quader $Q_{\varepsilon,z_1}, Q_{\varepsilon,z_2}$, deren Schnittmenge nicht leer ist, liefern auf $Q_{\varepsilon,z_1} \cap Q_{\varepsilon,z_2}$ dieselbe Funktion; Bew. z. B. per Induktion.) Nun gilt $g_\varepsilon \in C^0(S)$ und für $\varepsilon \to 0$:

$$\|g_\varepsilon - f\|_{C^0(S)} \leq \sup\big\{|f(x_1) - f(x_2)| \;\big|\; x_1, x_2 \in S, \|x_1 - x_2\|_\infty \leq 3\varepsilon\big\} \to 0,$$

da f gleichmäßig stetig auf S ist. Damit folgt: Die abzählbare Menge

$$\big\{g_{1/k} \;\big|\; k \in \mathbb{N},\ \forall\, y \in M_\varepsilon\colon\ g_{1/k}(y) \in \mathbb{Q}\big\}$$

liegt dicht in $C^0(S)$.

(ii) Nach Analysis III liegt $C_0^0(\mathbb{R}^n)$ dicht in $L^p(\mathbb{R}^n)$. (Idee: Approximation zunächst durch Treppenfunktionen bezüglich Quadern. Dann approximiere charakteristische Funktionen auf Quadern durch stetige Funktionen.) Es gilt

$$C_0^0(\mathbb{R}^n) = \bigcup_{n\in\mathbb{N}} C_0^0\big(B_n(0)\big) \subset \bigcup_{n\in\mathbb{N}} C^0\big(\overline{B_n(0)}\big)$$

und für alle $n \in \mathbb{N}$ ist $C^0(\overline{B_n(0)})$ separabel bezüglich der C^0-Norm nach dem ersten Teil. Wegen

$$\|g\|_{L^p(\overline{B_n(0)})} \;\leq\; \big(\mathcal{L}^n(B_n(0))\big)^{1/p} \sup_{\overline{B_n(0)}} |g|$$

ist $C^0(\overline{B_n(0)})$ auch separabel bezüglich der L^p-Norm.

(iii) selbst!

■

Häufig müssen wir Funktionen durch glatte Funktionen approximieren. Dafür brauchen wir die Faltung:

Definition 10.15 (Faltung). Sei $p \in [1,\infty]$, sei $\varphi \in L^1(\mathbb{R}^n)$ und sei $f \in L^p(\mathbb{R}^n)$. Dann heißt die Abbildung

$$\mathbb{R}^n \to \mathbb{K}, \quad x \mapsto \int_{\mathbb{R}^n} \varphi(x-y)\, f(y)\; \mathrm{d}y$$

die *Faltung von f und φ* und wird mit $\varphi * f$ oder $f * \varphi$ bezeichnet.

Lemma 10.16 (Allgemeine Faltungsabschätzung). Sei $\varphi \in L^1(\mathbb{R}^n)$, sei $K\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{K}$ Lebesgue-messbar und sei $p \in [1,\infty]$. Dann gilt für

$$F(x) := \int_{\mathbb{R}^n} \varphi(x-y)\, K(x,y)\; \mathrm{d}y = \int_{\mathbb{R}^n} \varphi(y)\, K(x,x-y)\; \mathrm{d}y$$

die Ungleichung

$$\|F\|_p \;\leq\; \|\varphi\|_1 \sup_{h\in\operatorname{supp}(\varphi)} \|K(\,\cdot\,,\cdot-h)\|_p,$$

d. h. falls die rechte Seite endlich ist, so ist

$$y \mapsto \varphi(x-y)\, K(x,y)$$

in $L^1(\mathbb{R}^n)$ für fast alle $x \in \mathbb{R}^n$ und es gilt $F \in L^p(\mathbb{R}^n)$.

Beweis. Die Fälle $p \in \{1, \infty\}$ sind einfach. Wir zeigen daher den Fall für $p \in (1, \infty)$, wobei p' den konjugierten Exponenten bezeichne. Dann gelten folgende Abschätzungen:

$$\begin{aligned}
\int_{\mathbb{R}^n} |F(x)|^p \,\mathrm{d}x &\leq \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} \underbrace{|\varphi(y)|\,|K(x, x-y)|}_{= |\varphi(y)|^{1/p'} \left(|\varphi(y)|^{1/p}\,|K(x,x-y)|\right)} \,\mathrm{d}y \Big)^p \,\mathrm{d}x \\
&\overset{\mathrm{H}}{\leq} \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |\varphi(y)|^{p'/p'} \,\mathrm{d}y \right)^{p/p'} \left(\int_{\mathbb{R}^n} |\varphi(y)|\,|K(x, x-y)|^p \,\mathrm{d}y \right) \mathrm{d}x \\
&\overset{\mathrm{F}}{=} \|\varphi\|_1^{p/p'} \int_{\mathbb{R}^n} |\varphi(y)| \int_{\mathbb{R}^n} |K(x, x-y)|^p \,\mathrm{d}x \,\mathrm{d}y \\
&\leq \|\varphi\|_1^{p/p'+1} \sup_{y \in \operatorname{supp}(\varphi)} \|K(\,\cdot\,, \cdot - y)\|_p^p
\end{aligned}$$

Bei H geht dabei die Hölderungleichung ein und bei F der Satz von Fubini. Die Existenz der Integrale rechtfertigt man dabei „von unten nach oben“.

■

Korollar 10.17 (Faltungsabschätzung)**.** Für $f \in L^p(\mathbb{R}^n)$ und $\varphi \in L^1(\mathbb{R}^n)$ folgt

$$f * \varphi \in L^p(\mathbb{R}^n) \quad \text{mit} \quad \|f * \varphi\|_p \leq \|\varphi\|_1 \, \|f\|_p.$$

Beweis. Dies folgt direkt aus Lemma 10.16.

■

Lemma 10.18. Für $f \in L^p(\mathbb{R}^n)$ und $\varphi \in C_0^\infty(\mathbb{R}^n)$ gilt:

$$\varphi * f \in C^\infty(\mathbb{R}^n) \qquad \text{und} \qquad \partial^s(\varphi * f) = (\partial^s \varphi) * f.$$

Beweisskizze. Sei $R \in \mathbb{R}_{>0}$ mit $\operatorname{supp} \varphi \subset B_R(0)$. Dann gilt:

$$\frac{(\varphi * f)(x + he_i) - (\varphi * f)(x)}{h} = \int_{B_{2R}(0)} \underbrace{\frac{\varphi(x + he_i - y) - \varphi(x - y)}{h}}_{\to \partial_i \varphi(x-y) \text{ glm. in } y \text{ für } h \to 0} f(y) \,\mathrm{d}y.$$

Mit dem Satz von Lebesgue folgt:

$$\big(\partial_i(\varphi * f)\big)(x) = \partial_i F(x) = \int_{\mathbb{R}^n} (\partial_i \varphi)(x - y)\, f(y) \,\mathrm{d}y = \big((\partial_i \varphi) * f\big)(x).$$

Die Aussage über höhere Ableitungen erhält man mit Induktion über $|s|$.

■

Das Ziel ist es nun, eine Funktion f mit anderen Funktionen φ_ε zu falten, so dass $\varphi_\varepsilon * f$ für $\varepsilon \to 0$ gegen f konvergiert. Wenn die φ_ε gutartig genug sind, hat $\varphi_\varepsilon * f$ „gute Eigenschaften“. Die Grundidee ist dabei, die φ_ε so zu bauen, dass sie für $\varepsilon \to 0$ wie das Dirac-Maß im Punkt 0 wirken.

Definition 10.19 (Dirac-Folge). (1) Eine Folge $(\varphi_k)_{k\in\mathbb{N}}$ in $L^1(\mathbb{R}^n)$ heißt *(allgemeine) Dirac-Folge*, falls folgende Bedingungen gelten:

– Für alle $k \in \mathbb{N}$ gilt $\varphi_k \geq 0$ und $\int_{\mathbb{R}^n} \varphi_k \,\mathrm{d}\mathcal{L}^n = \|\varphi_k\|_1 = 1$.

– Für alle $\varrho \in \mathbb{R}_{>0}$ gilt $\int_{\mathbb{R}^n\setminus B_\varrho(0)} \varphi_k \,\mathrm{d}\mathcal{L}^n \to 0$ für $k \to \infty$.

(2) Sei $\varphi \in L^1(\mathbb{R}^n)$ mit $\varphi \geq 0$ und $\int_{\mathbb{R}^n} \varphi = 1$. Für $\varepsilon \in \mathbb{R}_{>0}$ definiere $\varphi_\varepsilon \colon \mathbb{R}^n \to \mathbb{R}$ durch

$$\varphi_\varepsilon(x) \coloneqq \frac{1}{\varepsilon^n}\,\varphi\left(\frac{x}{\varepsilon}\right).$$

Bemerkung: Man rechnet leicht nach: Ist $(\varepsilon_k)_{k\in\mathbb{N}}$ eine Nullfolge, so ist die Folge $(\varphi_{\varepsilon_k})_{k\in\mathbb{N}}$ eine Dirac-Folge. Im Folgenden nennen wir auch eine Familie $(\varphi_\varepsilon)_{\varepsilon\in\mathbb{R}_{>0}}$ eine Dirac-Folge (auch wenn es sich dabei nicht um eine Folge im üblichen Sinne handelt).

Bemerkung: Für $x \in \mathbb{R}^n$ bezeichne im Folgenden $|x| \coloneqq \|x\|_2$ die euklidsche Norm von x.

10.20 (Standardbeispiel für eine Dirac-Folge). Sei $\psi \in C^\infty(\mathbb{R},\mathbb{R})$ mit $\operatorname{supp}\psi \subset \mathbb{R}_{\leq 1}$. Zum Beispiel:

$$\psi(x) \coloneqq \begin{cases} 1, & x \leq 0 \\ \exp\left(\frac{1}{|x|^2-1}+1\right), & 0 \leq x \leq 1 \\ 0, & x \geq 1. \end{cases}$$

Setze dann $\varphi(x) \coloneqq \alpha\,\psi(|x|)$ für alle $x \in \mathbb{R}^n$ mit $\alpha = \left(\int_{B_1(0)} \psi\circ|\cdot|\right)^{-1}$, so dass also $\int_{\mathbb{R}^n}\varphi = 1$ gilt. Es gilt dann

$$\varphi \in C^\infty(\mathbb{R}^n) \quad \text{und} \quad \operatorname{supp}\varphi \subset \overline{B_1(0)}.$$

Weiter bildet $(\varphi_\varepsilon)_{\varepsilon\in\mathbb{R}_{>0}}$ (gemäß Definition 10.19) eine Dirac-Folge, für die offenbar

$$\operatorname{supp}(\varphi_\varepsilon) \subset B_\varepsilon(0)$$

gilt. Diese nennen wir *Standard-Dirac-Folge*.

Lemma 10.21.

(1) Sei $(\varphi_\varepsilon)_{\varepsilon\in\mathbb{R}_{>0}}$ eine Standard-Dirac-Folge und sei $f \in C^0(\mathbb{R}^n)$. Dann gilt

$$f * \varphi_\varepsilon \to f \qquad \text{für } \varepsilon \to 0 \quad \text{lokal gleichmäßig,}$$

d. h. für alle $x \in \mathbb{R}^n$ existiert eine Umgebung $U \subset \mathbb{R}^n$ von x auf welcher $((f*\varphi_\varepsilon)|_U)_{\varepsilon\in\mathbb{R}_{>0}}$ gleichmäßig gegen $f|_U$ konvergiert.

(2) Sei $(\varphi_k)_{k\in\mathbb{N}}$ eine Dirac-Folge und $f \in L^p(\mathbb{R}^n)$ für $p \in [1,\infty)$. Dann gilt:

$$f * \varphi_k \to f \quad \text{in } L^p(\mathbb{R}^n) \qquad \text{für } k \to \infty.$$

Beweis. (1) Sei $\tilde{x} \in \mathbb{R}^n$ und $B \coloneqq B_1(\tilde{x})$. Dann gilt für alle $x \in B$:

$$\begin{aligned}
\big|\big((f * \varphi_\varepsilon) - f\big)(x)\big| &= \left|\int_{\mathbb{R}^n} \varphi_\varepsilon(x-y)\big(f(y) - f(x)\big)\,\mathrm{d}y\right| \\
&= \left|\int_{\mathbb{R}^n} \varphi_\varepsilon(y)\big(f(x-y) - f(x)\big)\,\mathrm{d}y\right| \\
&\leq \int_{B_\varepsilon(0)} |\varphi_\varepsilon(y)|\,\big|f(x-y) - f(x)\big|\,\mathrm{d}y \\
&\leq \sup_{y \in B_\varepsilon(0)} |f(x-y) - f(x)|.
\end{aligned}$$

Als stetige Funktion ist f auf der kompakten Menge $\overline{B}$ sogar gleichmäßig stetig, d. h. für alle $\alpha \in \mathbb{R}_{>0}$ existiert ein $\delta_\alpha \in \mathbb{R}_{>0}$, so dass gilt:

$$\forall\, x \in B,\, y \in B_{\delta_\alpha}(0)\colon \quad \big|f(x-y) - f(x)\big| < \alpha.$$

Daraus folgt nun wie gewünscht die gleichmäßige Konvergenz auf B:

$$\sup_{x \in B} \big|\big((f * \varphi_\varepsilon) - f\big)(x)\big| \leq \sup_{x \in B} \sup_{y \in B_\varepsilon(0)} |f(x-y) - f(x)| \;\to\; 0 \qquad \text{für } \varepsilon \to 0.$$

(2) Für alle $k \in \mathbb{N}$, $x \in \mathbb{R}^n$ und $\delta \in \mathbb{R}_{>0}$ gilt:

$$\begin{aligned}
(f * \varphi_k - f)(x) &= \int_{\mathbb{R}^n} \varphi_k(y)\big(f(x-y) - f(x)\big)\,\mathrm{d}y \\
&= \int_{\mathbb{R}^n} (\chi_{B_\delta(0)}\varphi_k)(y)\big(f(x-y) - f(x)\big)\,\mathrm{d}y \\
&\quad + \int_{\mathbb{R}^n} (\chi_{\mathbb{R}^n \setminus B_\delta(0)}\varphi_k)(y)\big(f(x-y) - f(x)\big)\,\mathrm{d}y
\end{aligned}$$

Lemma 10.16 liefert dann:

$$\begin{aligned}
\|f * \varphi_k - f\|_p &\leq \|\chi_{B_\delta(0)}\varphi_k\|_1 \sup_{h \in B_\delta(0)} \|f(\,\cdot\, - h) - f\|_p \\
&\quad + \|\chi_{\mathbb{R}^n \setminus B_\delta(0)}\varphi_k\|_1 \sup_{h \in \mathbb{R}^n} \|f(\,\cdot\, - h) - f\|_p \\
&\leq \sup_{h \in B_\delta(0)} \|f(\,\cdot\, - h) - f\|_p + 2\|f\|_p \int_{\mathbb{R}^n \setminus B_\delta(0)} \varphi_k \\
&\xrightarrow[k \to \infty]{} \sup_{h \in B_\delta(0)} \|f(\,\cdot\, - h) - f\|_p \xrightarrow[\delta \to 0]{} 0
\end{aligned}$$

Die Konvergenzaussage bezüglich δ folgt dabei aus dem nächsten Lemma (10.22).

∎

Lemma 10.22 (Stetigkeit im L^p-Mittel)**.** Sei $f \in L^p(\mathbb{R}^n)$ für $p \in [1, \infty)$. Dann gilt:

$$f(\,\cdot\, - h) \to f \quad \text{in } L^p(\mathbb{R}^n) \qquad \text{für } |h| \to 0.$$

Beweis. Da $C^0_0(\mathbb{R}^n)$ dicht in $L^p(\mathbb{R}^n)$ liegt (Satz 10.14 (ii)), gibt es eine Folge $(f_k)_{k\in\mathbb{N}}$ in $C^0_0(\mathbb{R}^n)$ mit $f_k \to f$ in $L^p(\mathbb{R}^n)$ für $k \to \infty$. Für alle $k \in \mathbb{N}$ ist f_k gleichmäßig stetig und es gilt $f_k(\,\cdot\, - h) - f_k \in C^0_0(\mathbb{R}^n)$ für alle $h \in \mathbb{R}^n$, wobei wir den jeweils zugehörigen Träger dieser Funktionen mit $K_{k,h}$ bezeichnen. Man überlegt sich außerdem leicht, dass für festes $k \in \mathbb{N}$ und eine beschränke Menge $A \subset \mathbb{R}^n$ auch $\{\mathcal{L}^n(K_{k,h}) \mid h \in A\}$ in $\mathbb{R}$ beschränkt ist. Mit diesen Vorüberlegungen erhalten wir nun:

$$\begin{aligned}
\|f(\,\cdot\, - h) - f\|_p &\leq \|f_k(\,\cdot\, - h) - f_k\|_p + \|f - f_k\|_p + \|f(\,\cdot\, - h) - f_k(\,\cdot\, - h)\|_p \\
&\leq \mathcal{L}^n(K_{k,h})^{1/p} \sup_{x\in\mathbb{R}^n} |f_k(x-h) - f_k(x)| + 2\|f - f_k\|_p \\
&\xrightarrow[|h|\to 0]{} 2\|f - f_k\|_p \xrightarrow[k\to\infty]{} 0.
\end{aligned}$$

■

Satz 10.23. Sei $\Omega \subset \mathbb{R}^n$ offen und sei $p \in [1, \infty)$. Dann gilt:

$$C^\infty_0(\Omega) \text{ liegt dicht in } L^p(\Omega).$$

Beweis. Sei $f \in L^p(\Omega)$. Für $\delta \in \mathbb{R}_{>0}$ setze

$$\Omega_\delta := \big\{x \in \Omega \,\big|\, \operatorname{dist}(x, \partial\Omega) > \delta\big\} \cap B_{1/\delta}(0).$$

Dann gilt

$$\chi_{\Omega_\delta} f \to f \quad \text{in } L^p(\Omega) \qquad \text{für } \delta \to 0. \tag{$*$}$$

Dies ergibt sich aus $\Omega = \bigcup\{\Omega_\delta \,|\, \delta\in\mathbb{R}_{>0}\}$ und dem Satz über monotone Konvergenz. Sei nun $(\varphi_\varepsilon)_{\varepsilon\in\mathbb{R}_{>0}}$ eine Standard-Dirac-Folge. Wegen Lemma 10.18, $\operatorname{supp}(\varphi_\varepsilon) \subset B_\varepsilon(0)$ und der speziellen Wahl der Ω_δ gilt

$$(\chi_{\Omega_\delta} f) * \varphi_\varepsilon \in C^\infty_0(\Omega)$$

für alle $\varepsilon < \delta$. Wegen

$$(\chi_{\Omega_\delta} f) * \varphi_\varepsilon \to \chi_{\Omega_\delta} f \quad \text{in } L^p(\Omega) \qquad \text{für } \varepsilon \to 0$$

nach Lemma 10.21 (2) folgt zusammen mit $(*)$ die Behauptung.

■

Das nächste Ziel ist, eine Charakterisierung von kompakten Mengen in $C^0(K)$ (für ein Kompaktum $K \subset \mathbb{R}^n$) und in $L^p(\mathbb{R}^n)$ zu finden.

Satz 10.24 (Arzelà-Ascoli)**.** Sei $K \subset \mathbb{R}^n$ kompakt und $A \subset C^0(K, \mathbb{R}^m)$. Dann ist A genau dann präkompakt, wenn folgende Eigenschaften erfüllt sind:

(i) A ist beschränkt, also $\sup_{f\in A} \|f\|_{C^0(K,\mathbb{R}^n)} < \infty$.

(ii) A ist *gleichgradig stetig*, d. h.

$$\sup_{f\in A} |f(x) - f(y)| \;\to\; 0 \qquad \text{für } |x - y| \to 0.$$

Beweis. „$\Rightarrow$“: Sei $\varepsilon \in \mathbb{R}_{>0}$ und (nach Präkompaktheit) $A \subset \bigcup_{i=1}^{m_\varepsilon} B_\varepsilon(f_{\varepsilon,i})$. Dann gilt für jedes $f \in A$: Es existiert ein i mit $f \in B_\varepsilon(f_{\varepsilon,i})$. Daraus folgt

$$\sup_{f\in A} \|f\|_{C^0} \;\leq\; \varepsilon + \max_{j\in\{1,\ldots,m_\varepsilon\}} \|f_{\varepsilon,j}\|_{C^0} \;<\; \infty,$$

also (i). Mithilfe der $\triangle$-Ungleichung erhalten wir für alle $f \in A$, $x, y \in K$:

$$|f(x) - f(y)| \;\leq\; 2\|f - f_{\varepsilon,i}\|_{C^0} + |f_{\varepsilon,i}(x) - f_{\varepsilon,i}(y)|.$$

Daraus folgt:

$$\sup_{f\in A} |f(x) - f(y)| \;\leq\; 2\varepsilon + \underbrace{\max_{j\in\{1,\ldots,m_\varepsilon\}} |f_{\varepsilon,j}(x) - f_{\varepsilon,j}(y)|}_{\to 0 \text{ für } |x-y|\to 0},$$

wobei wir die Konvergenz im hinteren Summanden erhalten, da alle $f_{\varepsilon,i}$ sogar gleichmäßig stetig sind. Da $\varepsilon \in \mathbb{R}_{>0}$ beliebig war, folgt (ii).

„$\Leftarrow$“: Seien $\varepsilon \in \mathbb{R}_{>0}$, $R \coloneqq \sup_{f\in A}\|f\|_{C^0}$ und (nach Kompaktheit von K und $\overline{B_R(0)} \subset \mathbb{R}^m$)

$$K \subset \bigcup_{i=1}^{n_\varepsilon} B_\varepsilon(x_i) \subset \mathbb{R}^n, \qquad \overline{B_R(0)} \subset \bigcup_{j=1}^{m_\varepsilon} B_\varepsilon(\xi_j) \subset \mathbb{R}^m.$$

Sei M_ε die Menge aller Abbildungen $\{1,\ldots,n_\varepsilon\} \to \{1,\ldots,m_\varepsilon\}$ und für $\tau \in M_\varepsilon$ sei

$$A_\tau \coloneqq \big\{ f \in A \;\big|\; \forall\, i \in \{1,\ldots,n_\varepsilon\}\colon\; |f(x_i) - \xi_{\tau(i)}| \leq \varepsilon \big\}.$$

Aus der bisherigen Konstruktion ergibt sich, dass für alle $f \in A$ ein $\tau(f) \in M_\varepsilon$ existiert mit $f \in A_{\tau(f)}$. Es gilt also $A = \bigcup_{\tau\in M_\varepsilon} A_\tau$. Wir fixieren außerdem für alle $\tau \in M$ mit $A_\tau \neq \emptyset$ ein $f_\tau \in A_\tau$. Seien nun $f \in A$, $\tau \coloneqq \tau(f)$, $x \in K$ und $i \in \{1,\ldots,n_\varepsilon\}$ mit $x \in B_\varepsilon(x_i)$. Dann gilt:

$$\begin{aligned} |(f - f_\tau)(x)| &\leq |f(x) - f(x_i)| + |f(x_i) - f_\tau(x_i)| + |f_\tau(x) - f_\tau(x_i)| \\ &\leq 2\underbrace{\sup_{|y_1-y_2|<\varepsilon}\ \sup_{f\in A} |f(y_1) - f(y_2)|}_{=:\,\delta_\varepsilon} + \underbrace{|f(x_i) - \xi_{\tau(i)}| + |f_\tau(x_i) - \xi_{\tau(i)}|}_{\leq 2\varepsilon}. \end{aligned}$$

Somit erhalten wir

$$\|f - f_\tau\|_{C^0} \leq 2\delta_\varepsilon + 2\varepsilon < 2\delta_\varepsilon + 3\varepsilon.$$

Es folgt:

$$A \subset \bigcup_{\tau\in M} B_{2\delta_\varepsilon + 3\varepsilon}(f_\tau).$$

Weil A gleichgradig stetig ist, gilt $\delta_\varepsilon \to 0$ für $\varepsilon \to 0$. Also wird $2\delta_\varepsilon + 3\varepsilon$ beliebig klein, woraus die Präkompaktheit von A folgt.

■

Bemerkung: Für einen metrischen Raum (X, d) und eine Teilmenge $A \subset X$ gilt:

$$\big(\forall\, \varepsilon \in \mathbb{R}_{>0}\ \exists\, A_\varepsilon \subset X \text{ präkompakt}\colon\; A \subset B_\varepsilon(A_\varepsilon)\big) \quad\Longrightarrow\quad A \text{ präkompakt},$$

wobei

$$B_\varepsilon(A_\varepsilon) \coloneqq \{x \in X \mid \operatorname{dist}(x, A_\varepsilon) < \varepsilon\}.$$

Satz 10.25 (Kompakte Mengen in $L^p(\mathbb{R}^n)$, Riesz-Fréchet-Kolmogorov)**.** Sei $p \in [1, \infty)$ und sei $A \subset L^p(\mathbb{R}^n)$. Dann ist A präkompakt genau dann, wenn folgende Bedingungen erfüllt sind:

(i) $\sup_{f\in A} \|f\|_p < \infty$ (Beschränktheit von A)

(ii) $\sup_{f\in A} \|f(\,\cdot\, - h) - f\|_p \to 0$ für $|h| \to 0$ (Gleichgradige Stetigkeit im L^p-Mittel)

(iii) $\sup_{f\in A} \|f\|_{L^p(\mathbb{R}^n \setminus B_R(0))} \to 0$ für $R \to \infty$.

Beweis. „$\Rightarrow$“: Diese Richtung ist einfach. Nutze, dass die entsprechenden Aussagen für endlich viele f erfüllt sind, und die Präkompaktheit. (Vgl. Beweis von Satz 10.24.)

„$\Leftarrow$“: Sei $(\varphi_\varepsilon)_{\varepsilon\in\mathbb{R}_{>0}}$ eine Standard-Dirac-Folge. Wähle $(R_\varepsilon)_{\varepsilon\in\mathbb{R}_{>0}}$ in $\mathbb{R}_{>0}$ mit $R_\varepsilon \to \infty$ für $\varepsilon \to 0$. Für $\varepsilon \in \mathbb{R}_{>0}$ und $f \in A$ definiere

$$\begin{aligned} B_\varepsilon &\coloneqq B_{R_\varepsilon}(0), \qquad B_\varepsilon^+ \coloneqq B_{R_\varepsilon+\varepsilon}(0), \\ T_\varepsilon f &\coloneqq (\chi_{B_\varepsilon} f) * \varphi_\varepsilon \ \in C_0^\infty(B_\varepsilon^+), \\ \text{und} \qquad A_\varepsilon &\coloneqq \{T_\varepsilon f \mid f \in A\}. \end{aligned}$$

Mit $\chi_{B_\varepsilon} = \chi_{\mathbb{R}^n} - \chi_{\mathbb{R}^n\setminus B_\varepsilon}$ erhalten wir für alle $f \in A$:

$$\begin{aligned} \|T_\varepsilon f - f\|_p &= \|f * \varphi_\varepsilon - f - (\chi_{\mathbb{R}^n\setminus B_\varepsilon} f) * \varphi_\varepsilon\|_p \\ &\le \sup_{f\in A} \big(\|f * \varphi_\varepsilon - f\|_p + \|(\chi_{\mathbb{R}^n\setminus B_\varepsilon} f) * \varphi_\varepsilon\|_p\big) \\ &\le \sup_{f\in A} \Big(\sup_{h\in B_\varepsilon(0)} \|f(\,\cdot\, - h) - f\|_p + \|f\|_{L^p(\mathbb{R}^n\setminus B_\varepsilon)} \Big) \\ &\to 0 \qquad \text{für } \varepsilon \to 0 \quad \text{wg. (ii) und (iii)}, \end{aligned}$$

wobei bei der letzten Unlgeichung Lemma 10.16 und Korollar 10.17 eingehen. Nach der Bemerkung vor dem Beweis genügt es nun also zu zeigen, dass A_ε für alle $\varepsilon \in \mathbb{R}_{>0}$ präkompakt in $L^p(B_\varepsilon^+)$ ist. Sei $D_\varepsilon \coloneqq \overline{B_\varepsilon^+}$. Für alle $g \in A_\varepsilon$ gilt

$$\|g\|_{L^p(B_\varepsilon^+)} \le \mathcal{L}^n(B_\varepsilon^+)^{1/p} \|g\|_{C^0(D_\varepsilon)},$$

also genügt es zu zeigen: A_ε ist präkompakt in $C^0(D_\varepsilon)$. Nach dem Satz von Arzelà-Ascoli (10.24) können wir äquivalent zeigen: A_ε ist beschränkt und gleichgradig stetig. Für alle $f \in A$ und $x \in D_\varepsilon$ gilt nach der Hölder'schen Ungleichung und (i):

$$|T_\varepsilon f(x)| \le \|\varphi_\varepsilon\|_{p'} \|f\|_p \le \|\varphi_\varepsilon\|_{p'} \sup_{g\in A} \|g\|_p < \infty$$

sowie

$$|\nabla(T_\varepsilon f)(x)| = |(\chi_{B_\varepsilon} f) * (\nabla\varphi_\varepsilon)| \le \|\nabla\varphi_\varepsilon\|_{p'} \sup_{g\in A} \|g\|_p =: C < \infty.$$

Aus der ersten Ungleichung folgt die Beschränktheit von A_ε und aus der zweiten, dass für alle $f \in A,\ x, y \in D_\varepsilon$ gilt:

$$\big|(T_\varepsilon f)(x) - (T_\varepsilon f)(y)\big| \le \|\nabla(T_\varepsilon f)\|_{C^0(D_\varepsilon)} |x - y| \le C\,|x - y|.$$

Daraus folgt, dass A_ε gleichgradig stetig ist.

■