Universität Heidelberg

FUNKTIONALANALYSIS

Prof. Dr. A. Marciniak-Czochra

Sommersemester 2018

Stand 10. Oktober 2018

Chris Kowall

Korrekturhinweise an kowall@math.uni-heidelberg.de

Inhaltsverzeichnis

1 Strukturen

Die Funktionalanalysis untersucht allgemeine Strukturen und Konzepte auf Funktionenräumen sowie stetige Abbildungen zwischen solchen Räumen wie beispielsweise die bekannten stetigen reellwertigen Funktionen $C(\Omega)$, die Folgenräume ℓ^p oder die Lebesgue-integrierbaren Funktionen $L^p(\Omega)$ für offene Mengen $\Omega \subset \mathbb{R}^n$. Im Folgenden werden wir uns hauptsächlich auf lineare Abbildungen konzentrieren, d.h. auf die lineare Funktionalanalysis.

1.1 Topologische Räume

Definition 1.1. Ein **topologischer Raum** ist ein Paar $(X, \mathcal{T})$ bestehend aus einer Menge X und einer Teilmenge $\mathcal{T} \subset \operatorname{Pot}(X)$ der Potenzmenge von X mit:

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

(T2) Seien $U_1, U_2 \in \mathcal{T}$. Dann ist $U_1 \cap U_2 \in \mathcal{T}$,

(T3) Sei Λ eine beliebige Menge und $U_\lambda \in \mathcal{T} \ \ \forall\, \lambda \in \Lambda$. Dann ist $\bigcup\limits_{\lambda \in \Lambda} U_\lambda \in \mathcal{T}$.

$\mathcal{T}$ heißt dann eine **Topologie** auf X und die Elemente $O \in \mathcal{T}$ nennt man **offene Mengen**. $A \subset X$ heißt **abgeschlossen**, falls $A^c := X \setminus A \in \mathcal{T}$ offen ist.

Bemerkung: Die Topologie bestimmt daher, welche Mengen offen bzw. abgeschlossen sind. Laut Definition sind endliche Durchschnitte sowie beliebige Vereinigungen offener Mengen wiederum offen. Wegen

$$\left(\bigcup_{\lambda \in \Lambda} A_\lambda \right)^c = \bigcap_{\lambda \in \Lambda} A_\lambda^c$$

sind beliebig viele Durchschnitte und endliche Vereinigungen abgeschlossener Mengen abgeschlossen.

Beispiel 1.2.

(a) Für eine beliebige Menge X bezeichnet $\mathcal{T}_1 = \{\emptyset, X\}$ die **triviale** und $\mathcal{T}_2 = \operatorname{Pot}(X)$ die **diskrete Topologie**.

(b) Für $X = \mathbb{R}^n$ sagt man: $A \subset X$ ist offen, wenn für alle $x \in A$ ein $\epsilon > 0$ mit

$$B_\epsilon(x) = \{y \in \mathbb{R}^n \mid |x - y| < \epsilon\} \subset A$$

existiert. Dies definiert eine Topologie auf $\mathbb{R}^n$.

(c) Sei $(X, \mathcal{T})$ ein topologischer Raum und $Y \subset X$. Dann kann man die **Unterraumtopologie**

$$\mathcal{T}_Y := \{Y \cap A \mid A \in \mathcal{T}\}$$

definieren. Dann ist $(Y, \mathcal{T}_Y)$ ein topologischer Raum.

(d) Seien $(X, \mathcal{T})$ und $(Y, \mathcal{S})$ topologische Räume. Dann lässt sich auf

$$X \times Y = \{(x, y) \mid x \in X, y \in Y\}$$

die **Produkttopologie** $\mathcal{T}_{\text{prod}}$ wie folgt definieren:

$$W \in \mathcal{T}_{\text{prod}}, \text{ falls } \forall (x, y) \in W \quad \exists U \in \mathcal{T}, V \in \mathcal{S}, \text{ sodass } (x, y) \in U \times V \subset W.$$

Definition 1.3. Sei $(X, \mathcal{T})$ ein topologischer Raum und $A \subset X$ eine Teilmenge. Dann ist der **Abschluss** von A definiert durch

$$\overline{A} := \bigcap \{B \mid B^c \in \mathcal{T} \text{ und } A \subset B\},$$

d.h. $\overline{A}$ ist die kleinste abgeschlossene Menge, die A enthält. Das **Innere** von A ist definiert durch

$$A^\circ := \bigcup \{B \subset A \mid B \in \mathcal{T}\},$$

d.h. A° ist die größte offene Menge, die in A enthalten ist. Es folgt $A^\circ \subset A \subset \overline{A}$.
Der **Rand** von A ist dann definiert als

$$\partial A := \overline{A} \setminus A^\circ.$$

Definition 1.4. Sei $(X, \mathcal{T})$ ein topologischer Raum. Eine Menge $A \subset X$ heißt **dicht**, falls $\overline{A} = X$. Der Raum X heißt **separabel**, falls er eine abzählbare dichte Teilmenge enthält.

Beispiel 1.5. (a) Falls $\mathcal{T}_1 = \{\emptyset, X\}$ die triviale Topologie ist, so findet man $\overline{A} = X$ und $A^\circ = \emptyset$ für $A \neq \emptyset$, d.h. jede nicht-leere Menge ist dicht und X ist immer separabel.

(b) Für die diskrete Topologie $\mathcal{T}_2 = \operatorname{Pot}(X)$ folgt $\overline{A} = A^\circ = A \quad \forall A \subset X$. Hier ist X separabel, wenn X abzählbar ist.

Für $\mathbb{R}$ mit der Topologie aus Beispiel 1.2 (b) erhält man

(c) Für $A = \mathbb{Q}$ ist $\overline{\mathbb{Q}} = \mathbb{R}$ und $\mathbb{Q}^\circ = \emptyset$, d.h. $\partial A = \overline{A}$ und $\mathbb{Q}$ liegt dicht im separablen Raum $\mathbb{R}$.

(d) Für $A = B_\varepsilon(x)$ ist $A^\circ = A$,

$$\overline{A} = \{y \in \mathbb{R} \mid |x - y| \leq \varepsilon\} \quad \text{und} \quad \partial A = \{y \in \mathbb{R} \mid |x - y| = \varepsilon\}.$$

Definition 1.6. Sei $(X, \mathcal{T})$ ein topologischer Raum und $x \in X$. Eine Menge $U \subset X$ ist eine **offene Umgebung** von x, falls $U \in \mathcal{T}$ und $x \in U$. Für eine beliebige Menge Λ ist eine Familie $\{U_\lambda\}_{\lambda \in \Lambda}$ bestehend aus offenen Umgebungen von x eine **Umgebungsbasis** von $x \in X$, falls gilt:

$$V \subset X \text{ offene Umgebung von } x \quad \Rightarrow \quad \exists \lambda \in \Lambda, \text{ sodass } U_\lambda \subset V.$$

Definition 1.7 (Konvergenz in topologischen Räumen)**.** Sei $(X, \mathcal{T})$ ein topologischer Raum und $(x_n)_{n \in \mathbb{N}}$ eine Folge in X. Dann **konvergiert** $(x_n)_{n \in \mathbb{N}}$ gegen $x \in X$ für $n \to \infty$, wenn für alle offenen Umgebungen $U \subset X$ von x ein $n_0 \in \mathbb{N}$ existiert, sodass $x_n \in U$ für alle $n \geq n_0$. Schreibweise: $x_n \to x$ für $n \to \infty$.
$x \in X$ heißt **Häufungspunkt** der Folge $(x_n)_{n \in \mathbb{N}}$, falls in jeder Umgebung $U \in \mathcal{T}$ von x stets unendlich viele Folgenglieder $x_n \in U$ liegen.

Bemerkung: In metrischen Räumen charakterisiert der Begriff der Konvergenz die Topologie. Dies ist in allgemeinen topologischen Räumen *nicht* der Fall. Ist $(X, \mathcal{T})$ ein topologischer Raum und A abgeschlossen, dann sind alle Limites von Folgen in A noch in A enthalten. Die umgekehrte Implikation gilt aber nur, wenn jeder Punkt von X eine abzählbare Umgebungsbasis besitzt. Siehe Beispiel 1.13 (b).

Definition 1.8 (Stetigkeit)**.** Seien $(X, \mathcal{T})$ und $(Y, \mathcal{S})$ zwei topologische Räume. Eine Funktion $f : X \to Y$ heißt **stetig**, falls $f^{-1}(V) \in \mathcal{T}$ für alle $V \in \mathcal{S}$ gilt, d.h. wenn das Urbild jeder offenen Menge in Y eine offene Menge in X ist. Für $x \in X$ sagen wir, dass f stetig in x ist, falls für alle offenen Umgebungen V von $f(x)$ eine offene Umgebung U von x existiert, sodass $f(U) \subset V$ gilt.

Eine Funktion ist damit genau dann stetig (d.h. auf ganz X) wenn sie in jedem Punkt des Definitionsbereichs X stetig ist.

Lemma 1.9. *Seien $(X, \mathcal{T})$ und $(Y, \mathcal{S})$ zwei topologische Räume und $f : X \to Y$ eine stetige Funktion. Dann gilt für alle konvergenten Folgen $(x_n)_{n\in\mathbb{N}}$ in X mit $x_n \to x$ auch $f(x_n) \to f(x)$ für $n \to \infty$.*

Beweis. Siehe Übungen. □

Bemerkung: Die Umkehrung von Lemma 1.9 gilt i.Allg. nicht. Sondern nur in topologischen Räumen mit einer höchstens abzählbaren Umgebungsbasis für jeden Punkt.
Es ist jedoch einfach zu zeigen, dass f in X stetig ist genau dann wenn f in jedem $x \in X$ stetig ist.

Definition 1.10 (Vergleich von Topologien)**.** Seien $\mathcal{T}_1, \mathcal{T}_2$ zwei Topologien auf einer Menge X. Dann nennt man $\mathcal{T}_1$ **feiner** oder **stärker** als $\mathcal{T}_2$ (oder $\mathcal{T}_1$ eine **Verfeinerung** von $\mathcal{T}_2$) bzw. man nennt $\mathcal{T}_2$ **gröber** oder **schwächer** als $\mathcal{T}_1$, wenn $\mathcal{T}_2 \subset \mathcal{T}_1$.

Bemerkung:

(i) Sei $\mathcal{T}_1$ feiner als $\mathcal{T}_2$. Dann gibt es in $(X, \mathcal{T}_1)$ mehr offene Mengen als in $(X, \mathcal{T}_2)$. Deswegen ist Konvergenz in $(X, \mathcal{T}_1)$ schwieriger zu erhalten als in $(X, \mathcal{T}_2)$, d.h. Konvergenz $x_n \to x$ bzgl. $\mathcal{T}_1$ impliziert $x_n \to x$ bzgl. $\mathcal{T}_2$.

(ii) Seien $\mathcal{T}_1, \mathcal{T}_2$ Topologien von X, $\mathcal{S}_1, \mathcal{S}_2$ Topologien von Y und $f : X \to Y$ eine Funktion.
Auch der Begriff der Stetigkeit hängt von der Topologie ab. Sei $\mathcal{T}_2 \subset \mathcal{T}_1$. Dann impliziert die Stetigkeit von $f : X \to Y$ bzgl. $\mathcal{T}_2$ die Stetigkeit von f bzgl. $\mathcal{T}_1$.
Andererseits, wenn $\mathcal{S}_1 \supset \mathcal{S}_2$ zwei Topologien auf Y sind, impliziert die Stetigkeit von f bzgl. $\mathcal{S}_1$ die Stetigkeit von f bzgl. $\mathcal{S}_2$.

Beispiel 1.11. Sei X eine beliebige Menge.

(a) $\mathcal{T}_1 = \operatorname{Pot}(X)$ ist die feinste Topologie auf X. Hierin ist z.B. $U = \{x\}$ für jedes $x \in X$ offen. Eine Folge konvergiert, d.h. $x_n \to x$, genau dann wenn $x_n = x$ stationär für $n \geq n_0$.
Insbesondere ist jede Funktion $f : X \to Y$ stetig für $\mathcal{T}_1 = \operatorname{Pot}(X)$.

(b) $\mathcal{T}_2 = \{\emptyset, X\}$ ist die gröbste Topologie auf X. Jeder Punkt $x \in X$ hat nur die Umgebung X und jede beliebige Folge ist konvergent. Falls X mehrere Elemente enthält, ist der Grenzwert nicht eindeutig.
Trägt Y die triviale Topologie $\mathcal{T}_2 = \{\emptyset, Y\}$, so ist jede Funktion $f : X \to Y$ stetig.

Damit ein topologischer Raum hinreichend viele offene Mengen beinhaltet, gibt es diverse Trennungseigenschaften, die an eine Topologie gestellt werden können. Eine geht dabei auf F. Hausdorff zurück.

Definition 1.12. Ein topologischer Raum $(X, \mathcal{T})$ heißt **hausdorff'sch** (**Hausdorff-Raum**), falls beliebige Punkte $x, y \in X$, $x \neq y$, disjunkte offene Umgebungen besitzen, d.h.

$$\forall\, x, y \in X, x \neq y \quad \exists\, U_x, U_y \in \mathcal{T} \text{ mit } x \in U_x,\ y \in U_y \text{ und } U_x \cap U_y = \emptyset.$$

Der topologische Raum $(X, \mathcal{T})$ heißt **kompakt**, falls er hausdorff'sch ist und jede offene Überdeckung

$$X = \bigcup_{\lambda \in \Lambda} U_\lambda \quad \text{mit} \quad U_\lambda \in \mathcal{T} \ \ \forall\, \lambda \in \Lambda$$

eine endliche Teilüberdeckung

$$X = \bigcup_{i=1}^{n} U_{\lambda_i} \quad \text{mit} \quad \lambda_1, \ldots, \lambda_n \in \Lambda$$

besitzt.

Bemerkung:

(i) Sei $(X, \mathcal{T})$ ein kompakter topologischer Raum und $A \subset X$ abgeschlossen. Dann ist A versehen mit der Unterraumtopologie $\mathcal{T}_A$ ein kompakter Raum. In der Tat ist klar, dass A hausdorff'sch ist. Wenn $\{U_\lambda\}_{\lambda \in \Lambda}, U_\lambda \in \mathcal{T}_A$, eine offene Überdeckung von A ist, so liefert $\big\{\{U_\lambda\}_{\lambda \in \Lambda}, A^c\big\}$ eine offene Überdeckung von X. Deswegen gibt es eine endliche Teilüberdeckung von X gemäß

$$X = A^c \cup \bigcup_{i=1}^{n} U_{\lambda_i},$$

also ist $\bigcup\limits_{i=1}^{n} U_{\lambda_i}$ eine endliche Teilüberdeckung von A.

(ii) Der Grenzwert einer konvergenten Folge in einem Hausdorff-Raum ist aufgrund der Trennungseigenschaft eindeutig. Jedoch folgt auf topologischen Räumen aus Kompaktheit i. Allg. *nicht* die Folgenkompaktheit, d.h. nicht jede Folge besitzt eine konvergente Teilfolge.

Beispiel 1.13.

(a) Die triviale Topologie $\mathcal{T} = \{\emptyset, X\}$, wobei X mindestens 2 Elemente besitze, ist kein Hausdorff-Raum (siehe Beispiel 1.11 (b)).

(b) Die Topologie auf $X = \mathbb{R}$ mit

$$\mathcal{T} = \{U \subset \mathbb{R} \mid U = \emptyset \text{ oder } U^c \text{ ist abzählbar}\}$$

ist nicht hausdorff'sch.

(c) Metrische Räume sind hausdorff'sch.

Satz 1.14. *Sei $(X, \mathcal{T})$ ein kompakter Raum und $(x_n)_{n\in\mathbb{N}}$ eine Folge in X. Dann existiert mindestens ein Häufungspunkt von $(x_n)_{n\in\mathbb{N}}$ in X.*

Beweis. Angenommen, $(x_n)_{n\in\mathbb{N}}$ habe keinen Häufungspunkt in X. Dann findet man für jedes $y \in X$ eine offene Umgebung $U_y \in \mathcal{T}$, sodass U_y nur endlich viele Punkte x_n enthält. $\{U_y\}_{y\in X}$ ist eine offene Überdeckung von X. Deswegen gibt es endliche viele $y_1, \ldots, y_m \in X$ mit

$$X = \bigcup_{i=1}^{m} U_{y_i}.$$

Dies widerspricht der Tatsache, dass jedes U_y nur endlich viele Punkte von x_n enthält. $\square$

Die folgenden Lemmata spiegeln hilfreiche Trennungseigenschaften wider, die normalerweise in einem Raum gelten sollten.

Lemma 1.15. *Sei $(X, \mathcal{T})$ ein kompakter Raum. Sei $x \in X$ und U eine offene Umgebung von x. Dann existiert eine offene Umgebung V von x mit $\overline{V} \subset U$.*

Beweis. Für jedes $y \in X \setminus U$ wähle man offene Umgebungen W_y von y und V_y von x mit $W_y \cap V_y = \emptyset$. Dann ist $\big\{\{W_y\}_{y\in X\setminus U}, U\big\}$ eine offene Überdeckung von X. Ex existieren also $n \in \mathbb{N}$ und $y_1, \ldots, y_n \in X \setminus U$ mit

$$X = U \cup \bigcup_{i=1}^{n} W_{y_i}.$$

Setze

$$V := \bigcap_{i=1}^{n} V_{y_i} \quad \text{sowie} \quad W := \bigcup_{i=1}^{n} W_{y_i},$$

dann ist V eine offene Umgebung von x und $X = U \cup W$. W ist ebenfalls offen mit $W \cap V = \emptyset$, deswegen ist W^c abgeschlossen mit $V \subset W^c$, also $\overline{V} \subset W^c$. Folglich ist $\overline{V} \cap W = \emptyset$, d.h. $\overline{V} \subset U$. $\square$

Lemma 1.16. *Sei $(X, \mathcal{T})$ ein kompakter Raum und $U, V \in \mathcal{T}$ mit $\overline{V} \subset U$. Dann gibt es ein $W \in \mathcal{T}$ mit*

$$\overline{V} \subset W \subset \overline{W} \subset U.$$

Beweis. Nach Lemma 1.15 existiert für jedes $x \in \overline{V}$ eine offene Umgebung $W_x \in \mathcal{T}$ von x mit $\overline{W}_x \subset U$. Dann ist $\big\{\{W_x\}_{x\in\overline{V}}, X \setminus \overline{V}\big\}$ eine offene Überdeckung von X. Deswegen gibt es endlich viele $x_1, \ldots, x_n \in \overline{V}$ mit

$$(X \setminus \overline{V}) \cup \bigcup_{i=1}^{n} W_{x_i} = X.$$

Setze $W := \bigcup_{i=1}^{n} W_{x_i}$, dann ist W offen und enthält $\overline{V}$. Außerdem ist bei Übergang zum Abschluss $\overline{W} \subset \bigcup_{i=1}^{n} \overline{W}_{x_i} \subset U$. $\square$

Definition 1.17. Einen topologischen Raum $(X, \mathcal{T})$ mit der Eigenschaft, dass für alle $U, V \in \mathcal{T}$ mit $\overline{V} \subset U$ ein $W \in \mathcal{T}$ existiert mit $\overline{V} \subset W \subset \overline{W} \subset U$, bezeichnet man als **normal**.

Lemma 1.16 zeigt, dass jeder kompakte Raum normal ist. Dabei ist die Kompaktheit wichtig, denn die Hausdorff-Eigenschaft allein impliziert dies nicht. 1925 konstruierte Urysohn einen Hausdorff-Raum, in welchem alle stetigen reellwertigen Funktionen konstant sind. Der folgende Satz ist eine Variante des Lemmas von Urysohn, um solche Trivialitäten auszuschließen.

Satz 1.18. *Seien $(X, \mathcal{T})$ ein normaler Raum und $A, B \subset X$ disjunkte, nicht-leere abgeschlossene Mengen. Dann existiert eine stetige Abbildung*

$$g : X \to [0,1] \quad \textit{mit} \quad A \subset g^{-1}(0),\ B \subset g^{-1}(1),$$

wobei das Intervall $[0,1] \subset \mathbb{R}$ mit der Standardtopologie versehen ist.

Beweis. Wähle per Definition eine Umgebung $U_{1/2} \in \mathcal{T}$ mit $A \subset U_{1/2} \subset \overline{U}_{1/2} \subset B^c$ sowie $U_{1/4}, U_{3/4} \in \mathcal{T}$ mit

$$A \subset U_{1/4} \subset \overline{U}_{1/4} \subset U_{1/2} \subset \overline{U}_{1/2} \subset U_{3/4} \subset \overline{U}_{3/4} \subset B^c.$$

Durch Iteration definiere man U_λ mit $\lambda \in \Lambda = \{\frac{m}{2^n} \mid m = 1, 2, \ldots, 2^{n-1},\ n \in \mathbb{N}\}$, sodass $\overline{U}_\lambda \subset U_\mu$ für $\lambda < \mu$ gilt. Setze $U_1 := B^c$ und $f : X \to [0,1]$ mit

$$f(x) = \begin{cases} 1 & \text{falls } x \in X \setminus \bigcup\limits_{\lambda \in \Lambda \cup \{1\}} U_\lambda, \\ \inf\limits_{x \in U_\lambda} \lambda & \text{falls } x \in \bigcup\limits_{\lambda \in \Lambda \cup \{1\}} U_\lambda. \end{cases}$$

Nach Konstruktion gilt

$$f(x) = \begin{cases} 0 & \text{falls } x \in A, \\ 1 & \text{falls } x \in B. \end{cases}$$

Es bleibt zu zeigen, dass f stetig ist: Sei also $x \in X$, $\epsilon > 0$ beliebig und konstruiere eine offene Umgebung U von x, sodass $f(U) \subset (f(x) - \epsilon, f(x) + \epsilon)$.

(i) $x \in A$: Wählt man $\lambda_0 \in \Lambda$ mit $\lambda_0 < \epsilon$, so ist U_{λ_0} eine gewünschte Umgebung mit $f(U_{\lambda_0}) \subset (-\epsilon, \epsilon)$.

(ii) $x \in B = X \setminus \bigcup\limits_{\lambda \in \Lambda \cup \{1\}} U_\lambda$: Wähle $\lambda_0 \in \Lambda$ mit $1 - \epsilon < \lambda_0 < 1$ und setze $U := X \setminus \overline{U}_{\lambda_0}$, dann ist U offene Umgebung mit $U \cap U_\lambda = \emptyset$ für alle $\lambda \leq 1 - \varepsilon$ und folglich $f(U) \subset (1 - \epsilon, 1 + \epsilon)$.

(iii) $x \in \bigcup\limits_{\lambda \in \Lambda \cup \{1\}} U_\lambda \setminus A$: Wegen $f(x) \in (0,1)$ wähle $\lambda_0, \lambda_1 \in \Lambda$ mit

$$f(x) - \epsilon < \lambda_0 < f(x) < \lambda_1 < f(x) + \epsilon.$$

Setzt man $U := U_{\lambda_1} \setminus \overline{U}_{\lambda_0}$, so folgt $f(U) \subset (\lambda_0, \lambda_1) \subset (f(x) - \epsilon, f(x) + \epsilon)$.

□

Folglich gibt es in normalen Räumen eine Vielzahl stetiger Funktionen die auf disjunkten, abgeschlossenen Mengen vorgegebene Werte annehmen können. Des weiteren lassen sich auf abgeschlossenen Teilmengen $A \subset X$ stetige reellwertige Funktionen auf den ganzen Raum stetig fortsetzen nach einem Satz von Tietze-Urysohn.

1.2 Metrische Räume

Neben den kompakten Räumen sind auch metrische Räume normal. Im Folgenden einige Wiederholungen aus Analysis 2 und deren Einordnung in den topologischen Kontext.

Definition 1.19 (Ana 2)**.** Sei X eine Menge, eine Abbildung $d : X \times X \to [0, \infty)$ heißt eine **Metrik** auf M genau dann wenn für alle $x, y, z \in X$ gilt:

(D1) $d(x, y) = 0 \quad \Leftrightarrow \quad x = y$ (Definitheit)

(D2) $d(x, y) = d(y, x)$ (Symmetrie)

(D3) $d(x, z) \leq d(x, y) + d(y, z)$ (Dreiecksungelichung)

Das Paar (X, d) heißt dann **metrischer Raum**.

Beispiel 1.20 (Ana 2)**.** Einige wichtige und wiederkehrende Beispiele, welche man als Übung leicht verifiziert, sind:

(a) Sei $X = \mathbb{K}^n$ für $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Die *euklidische Metrik* ist bestimmt durch

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

(b) Sei $X = \mathbb{R}^n$. Für $1 \leq p < \infty$ definieren wir

$$d_p(x, y) = \left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^{\frac{1}{p}}.$$

Ist $p = \infty$, so definieren wir

$$d_\infty(x, y) = \max_{i=1,\dots,n} |x_i - y_i|.$$

(c) Für $1 \leq p < \infty$ definiert man die Folgenräume

$$\ell^p := \{a = (a_n)_{n\in\mathbb{N}}, a_n \in \mathbb{K} \mid \|a\|_p < \infty\} \quad \text{mit} \quad \|a\|_p := \left(\sum_{k=1}^{\infty} |a_k|^p \right)^{1/p}$$

der zur p-ten Potenz absolut summierbaren Folgen bzw. den Raum der beschränkten Folgen

$$\ell^\infty := \{a = (a_n)_{n\in\mathbb{N}}, a_n \in \mathbb{K} \mid \|a\|_\infty < \infty\} \quad \text{mit} \quad \|a\|_\infty := \sup_{k\in\mathbb{N}} |a_k|.$$

So ist mit analoger Notation (ℓ^p, d_p) ein metrischer Raum für

$$d_p(x, y) = \begin{cases} \left(\sum_{k\in\mathbb{N}} |x_k - y_k|^p \right)^{\frac{1}{p}} & 1 \leq p < \infty, \\ \sup_{k\in\mathbb{N}} |x_k - y_k| & p = \infty. \end{cases}$$

(d) Sei X eine beliebige Menge. Die *charakteristische (diskrete) Metrik* ist definiert durch
$$d(x,y) = \begin{cases} 0 & \text{für } x = y, \\ 1 & \text{sonst.} \end{cases}$$

(e) Sei $X = \mathbb{R}$. Neben der euklidischen Standardmetrik auf $\mathbb{R}$ lässt sich die Metrik
$$d(x,y) = \frac{|x-y|}{1+|x-y|}$$
definieren.

(f) Der Raum $\mathbb{R}^{\mathbb{N}}$ der reellen Folgen kann mit der Metrik
$$d(x,y) = \sum_{k=1}^{\infty} 2^{-k} \frac{|x_k - y_k|}{1+|x_k-y_k|}$$
versehen werden.

Mit einer Metrik lassen sich auch Abstände zwischen Mengen messen.

Definition 1.21 (Abstände)**.** Sei (X,d) ein metrischer Raum und seien $A_1, A_2, A \subset X$ und $x \in X$, $\varepsilon > 0$. Man definiert den Durchmesser einer Menge durch
$$\operatorname{diam}(A) := \sup\{d(x,y) \mid x,y \in A\}.$$
Eine Menge $A \subset X$ heißt **beschränkt**, falls $\operatorname{diam}(A) < \infty$ erfüllt ist.
Distanzen zwischen Mengen werden wie folgt abgekürzt:
$$\operatorname{dist}(A_1,A_2) := \inf\{d(x,y) \mid x \in A_1,\ y \in A_2\}, \quad \operatorname{dist}(A,\emptyset) := \infty,$$
$$\operatorname{dist}(x,A) := \operatorname{dist}(\{x\},A) = \inf\{d(x,y) \mid y \in A\},$$
sowie die entsprechende ε-Umgebung $B_\varepsilon(A) := \{y \in X \mid \operatorname{dist}(y,A) < \varepsilon\}$.

Bemerkung (Induzierte Topologie)**:** Jeder metrische Raum (X,d) ist ein topologischer Raum $(X,\mathcal{T}_d)$ mit der von der Metrik definierten Topologie $\mathcal{T}_d$ gemäß (siehe Übungen)
$$O \in \mathcal{T}_d \quad \Leftrightarrow \quad \forall\, x \in O \ \ \exists\, \epsilon > 0 \text{ mit } B_\epsilon(x) = \{y \in X \mid d(x,y) < \epsilon\} \subset O.$$
Ebenso ist durch $\{B_\epsilon(x)\}_{\epsilon>0}$ eine Umgebungsbasis von $x \in X$ bzgl. der induzierten Topologie definiert.
Während es immer möglich ist, eine Topologie auf einem metrischen Raum zu definieren, so ist es nicht immer möglich, auf einem topologischen Raum eine Metrik so zu definieren, dass die induzierte Topologie mit der ursprünglichen Topologie übereinstimmt, d.h. nicht alle Topologien sind **metrisierbar**. Siehe Beispiel 1.13.

Nun lassen sich auch der Konvergenzbegriff einer Folge oder die Stetigkeit einer Funktion äquivalent charakterisieren. Im Folgenden sei für einen metrischen Raum stets die induzierte Topologie angenommen.

Satz 1.22. *Seien (X,d) und $(Y,\tilde{d})$ zwei metrische Räume.*

(i) *Eine Folge* $(x_n)_{n\in\mathbb{N}}$ *mit* $x_n \in X$ *konvergiert genau dann gegen ein* $x \in X$ *wenn gilt:*

$$\forall\, \epsilon > 0 \;\; \exists\, n_0 \in \mathbb{N} : \quad d(x_n, x) < \epsilon \;\; \forall\, n \geq n_0$$

(ii) $A \subset X$ *ist abgeschlossen genau dann wenn für jede konvergente Folge* $(x_n)_{n\in\mathbb{N}}$ *mit Werten in* A *der Grenzwert ebenfalls in* A *liegt. Des Weiteren ist der Abschluss gegeben durch*

$$\overline{A} = \{x \in X \mid \exists\, (x_n)_{n\in\mathbb{N}} \text{ Folge in } A \text{ mit } x_n \to x \text{ für } n \to \infty\}.$$

(iii) *Die Funktion* $f : X \to Y$ *ist genau dann stetig in* $x \in X$ *wenn*

$$(x_n)_{n\in\mathbb{N}} \text{ Folge in } X \text{ mit } x_n \to x \;\;\Rightarrow\;\; f(x_n) \to f(x).$$

Beweis. (i) Folgt direkt aus obiger Bemerkung für die induzierte Topologie.

(ii) Sei zunächst A abgeschlossen, d.h. $X \setminus A$ offen, und $x_n \in A$ für $n \in \mathbb{N}$ mit $x_n \to x$ für $n \to \infty$. Angenommen $x \notin A$, dann ist $X \setminus A$ eine Umgebung von x und mit der Definition der Konvergenz folgt für ein $n_0 \in \mathbb{N}$, dass $x_n \in X \setminus A$ für alle $n \geq n_0$. Im Widerspruch zur Annahme.
Sei nun das Folgenkriterium erfüllt und $x \in X \setminus A$ beliebig. Wir wollen zeigen, dass es ein $\varepsilon > 0$ gibt mit $B_\varepsilon(x) \subset X \setminus A$. Wäre dies nicht der Fall, gäbe es zu jedem $n \in \mathbb{N}$ ein Punkt $x_n \in A$ mit $d(x_n, x) < 1/n$. Dann konvergiert aber bereits $x_n \to x$, sodass $x \in A$ gelten müsste, im Widerspruch zur Annahme. Damit ist $X \setminus A$ offen.

(iii) Siehe Übungen.

□

Ebenso lässt sich in metrischen Räumen Kompaktheit mittels Folgen nachweisen.

Definition 1.23. Sei (X, d) ein metrischer Raum. $A \subset X$ heißt **präkompakt**, falls A für jedes $\epsilon > 0$ eine endliche Überdeckung mit ϵ-Kugeln besitzt, d.h. falls $x_1, \ldots, x_n \in A$ existieren mit

$$A \subset \bigcup_{i=1}^{n} B_\epsilon(x_i).$$

Teilmengen präkompakter Mengen sind wiederum präkompakt. Außerdem sind präkompakte Mengen beschränkt, wie man sich leicht klar macht.

Definition 1.24 (Ana2; Vollständigkeit)**.** Sei (X, d) ein metrischer Raum. Eine Folge $(x_n)_{n\in\mathbb{N}}$ in X heißt **Cauchy-Folge**, falls $d(x_n, x_m) \to 0$ für $n, m \to \infty$.
Der metrische Raum (X, d) heißt **vollständig**, falls jede Cauchy-Folge konvergent ist.

Nun lässt sich die Kompaktheit charakterisieren durch folgenden

Satz 1.25. *Sei* (X, d) *ein metrischer Raum. Für* $A \subset X$ *sind die folgenden Aussagen äquivalent:*

(i) *A ist kompakt (überdeckungskompakt)* *(ii) A ist **folgenkompakt**, d.h. jede Folge in A besitzt eine konvergente Teilfolge mit Grenzwert in A.*

(iii) (A, d) ist vollständig und A ist präkompakt.

Beweis. **(i) ⇒ (ii):** Siehe Satz 1.14.

(ii) ⇒ (iii): Nach (ii) besitzt jede Cauchy-Folge in A einen Häufungspunkt in A, der wiederum in metrischen bzw. Hausdorff-Räumen eindeutig ist. Also konvergiert die Folge in A und (A, d) ist vollständig.
Gibt es für ein $\varepsilon > 0$ keine endliche ε-Überdeckung von A, so lässt sich induktiv die Folge $(x_n)_{n\in\mathbb{N}}$ mit

$$x_1 \in A, \quad x_{n+1} \in A \setminus \bigcup_{i=1}^{n} B_\varepsilon(x_i)$$

wählen, welche offensichtlich keinen Häufungspunkt hat. Dies ist aber ein Widerspruch.

(iii) ⇒ (i): Sei $(U_i)_{i\in I}$ eine offene Überdeckung von A und

$$M := \left\{ B \subset A \;\middle|\; J \subset I, B \subset \bigcup_{i\in J} U_i \;\Rightarrow\; \text{J ist unendlich} \right\}$$

die Menge der unendlich überdeckten Teilmengen. Es genügt $A \notin M$ zu zeigen. Nehmen wir nun an, es gälte $A \in M$, so gibt es aufgrund der Präkompaktheit für beliebiges $\varepsilon > 0$ eine endliche Überdeckung

$$A \subset \bigcup_{k=1}^{n_\varepsilon} B_\varepsilon(x_k) \qquad \text{mit} \quad x_k \in A.$$

Für $\varepsilon = 1$ gibt es dann ein Index $k \in \{1, \dots, n_\varepsilon\}$, o.E. $x_1 \in A$, mit nicht-leerem Schnitt $B_1(x_1) \cap A \in M$ wegen der unendlichen Überdeckung. Als Teilmenge ist nun wieder $B_1(x_1) \cap A \in M$ präkompakt und man findet $x_2 \in B_1(x_1) \cap A$ mit $B_{1/2}(x_2) \cap A \in M$. Induktiv erhält man eine Folge von x_k für $\varepsilon = 1/k$ mit

$$\emptyset \neq B_m := \bigcap_{k=1}^{m} B_{\frac{1}{k}}(x_k) \cap A \in M \qquad \forall\, m \in \mathbb{N}.$$

Wähle für jedes $m \in \mathbb{N}$ ein $y_m \in B_m$, dann ist $y_m, y_l \in B_{\frac{1}{m}}(x_m)$ für $m \leq l$, d.h. $d(y_m, y_l) \leq 2/m$ und $(y_m)_{m\in\mathbb{N}}$ ist eine Cauchy-Folge in A. Aufgrund der Vollständigkeit gibt es ein $y \in A$ mit $\varepsilon_m := d(y_m, y) \to 0$ für $m \to \infty$.
Nun ist aber auch $y \in U_i$ für ein $i \in J$ und es folgt für große m

$$B_m \subset B_{\frac{1}{m}}(x_m) \subset B_{\frac{2}{m}}(y_m) \subset B_{\frac{2}{m}+\varepsilon_m}(y) \subset U_i,$$

d.h. $B_m \notin M$, ein Widerspruch.

□

Bemerkung: Da jede konvergente Folge auch Cauchy-Folge ist, sind kompakte Mengen abegschlossen (wegen der Vollständigkeit).
Falls (X, d) vollständig ist, so ist $A \subset X$ präkompakt genau dann wenn $\overline{A}$ kompakt ist. Dies folgt, da abgeschlossene Teilmengen vollständiger Räume wiederum vollständig sind.

Beispiel 1.26. Für die Standardmetrik auf $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ ist

(a) $\mathbb{Q}$ mit $d(x, y) = |x - y|$ weder vollständig noch präkompakt,

(b) $\mathbb{K}^n$ vollständig, aber nicht präkompakt,

(c) das reelle Intervall $(0, 1)$ präkompakt, aber nicht kompakt.

Wie die Beispiele zeigen, ist nicht jeder metrische Raum auch vollständig, jedoch lässt sich dieser vervollständigen analog zur Konstruktion der reellen Zahlen als Vervollständigung der rationalen Zahlen.

Definition 1.27. Seien (X, d_X) und (Y, d_Y) metrische Räume. Eine Funktion $\phi : X \to Y$ heißt **Isometrie**, wenn für alle $x, y \in X$ gilt:

$$d_Y(\phi(x), \phi(y)) = d_X(x, y).$$

Eine **Vervollständigung** von (X, d_X) ist ein vollständiger, metrischer Raum (Y, d_Y) mit einer zugehörigen Isometrie $\phi : X \to Y$, sodass $\overline{\phi(X)} = Y$ gilt.

Lemma 1.28. *Seien (X, d_X) und (Y, d_Y) metrische Räume und (Y, d_Y) vollständig. Sei $\tau : S \subset X \to Y$ eine Isometrie auf einer dichten Teilmenge S von X. Dann besitzt τ eine eindeutige isometrische Fortsetzung $\tilde{\tau} : X \to Y$ auf ganz X, d.h. $\tilde{\tau}_{|S} = \tau$ und*

$$d_Y(\tilde{\tau}(x), \tilde{\tau}(y)) = d_X(x, y) \qquad \forall\, x, y \in X.$$

Analog lassen sich gleichmäßig stetige Funktionen auf ihren Abschluss eindeutig gleichmäßig stetig fortsetzen.

Beweis. Siehe Übungen. □

Die Begriffe Lipschitz- sowie gleichmäßige Stetigkeit lassen sich wie bereits in Analysis 2 auf metrischen Räumen definieren (siehe Übungen).

Satz 1.29. *Sei (X, d) ein metrischer Raum. Dann gibt es eine Vervollständigung, welche bis auf einen isometrischen Isomorphismus eindeutig bestimmt ist.*

Beweis. Man konstruiere die Vervollständigung explizit. Sei C_X die Menge aller Cauchy-Folgen in X. Wir führen auf C_X folgende Äquivalenzrelation für $x = (x_n)_{n\in\mathbb{N}}, y = (y_n)_{n\in\mathbb{N}}$ ein:

$$x \sim y \quad :\Leftrightarrow \quad \lim_{k\to\infty} d(x_k, y_k) = 0.$$

Da $(d(x_n, y_n))_{n\in\mathbb{N}}$ eine reelle Cauchy-Folge ist, ist der Grenzwert tatsächlich wohldefiniert und $\sim$ eine Äquivalenzrelation auf C_X.
Betrachte den Quotientenraum $Y := C_X/_\sim$ und definiere die Abbildung

$$d_Y : Y \times Y \to [0, \infty) \qquad \text{mit} \qquad d_Y([x], [y]) := \lim_{k\to\infty} d(x_k, y_k) \qquad \forall\, [x], [y] \in Y.$$

Die Dreiecksungleichung garantiert dabei die Wohldefiniertheit von d_Y, da die rechte Seite nicht von der Wahl der Repräsentanten abhängt. Analog folgen die Metrikaxiome für d_Y.

Für die Vollständigkeit sei $(x^{(n)})_{n\in\mathbb{N}}$ eine Cauchy-Folge in Y. Dann gibt es für jedes $n \in \mathbb{N}$ ein Cauchy-Folge $(x_k^{(n)})_{k\in\mathbb{N}}$ in X mit $x^{(n)} = [(x_k^{(n)})_{k\in\mathbb{N}}]$, d.h. insbesondere für ein j_n

$$d(x_i^{(n)}, x_j^{(n)}) \leq \frac{1}{n} \qquad \forall\, i, j \geq j_n.$$

Für $m, n \in \mathbb{N}$ folgt dann für $j \geq j_n, j_m$

$$\begin{aligned} d(x_{j_n}^{(n)}, x_{j_m}^{(m)}) &\leq d(x_{j_n}^{(n)}, x_j^{(n)}) + d(x_j^{(n)}, x_j^{(m)}) + d(x_j^{(m)}, x_{j_m}^{(m)}) \\ &\leq \frac{1}{n} + d(x_j^{(n)}, x_j^{(m)}) + \frac{1}{m} \\ &\to \frac{1}{n} + d_Y(x^{(n)}, x^{(m)}) + \frac{1}{m} \qquad (\text{für } j \to \infty). \end{aligned}$$

Sei nun $y \in X$ die Diagonalfolge definiert durch $y_n := x_{j_n}^{(n)}$, dann ist y eine Cauchy-Folge in X, denn

$$d(y_n, y_m) \leq \frac{1}{n} + d_Y(x^{(n)}, x^{(m)}) + \frac{1}{m} \to 0 \qquad (\text{für } m, n \to \infty).$$

Für $n \geq j_m$ gilt außerdem

$$d(x_n^{(m)}, y_n) \leq d(x_n^{(m)}, x_{j_m}^{(m)}) + d(x_{j_m}^{(m)}, x_{j_n}^{(n)}) \leq \frac{1}{m} + d(y_n, y_m),$$

sodas die Folge $(x^{(n)})_{n\in\mathbb{N}}$ gegen $[y] \in Y$ konvergiert:

$$d_Y(x^{(m)}, [y]) = \lim_{n\to\infty} d(x_n^{(m)}, y_n) \to 0 \qquad (\text{für } n, m \to \infty).$$

Damit ist (Y, d_Y) vollständig.

Man definiere die Abbildung $\phi : X \to Y$ mit

$$\phi(x) = [(x, x, x, \ldots)],$$

d.h. $\phi(x)$ ist die Äquivalenzklasse der konstanten Folge mit Wert x und definiert damit eine Isometrie, welche insbesondere injektiv ist.

Man zeige nun, dass $\phi(X)$ dicht in Y ist. Sei dazu $x = [(x_n)_{n\in\mathbb{N}}] \in Y$. Dann gilt mit der Cauchy-Eigenschaft der Folge $(x_n)_{n\in\mathbb{N}}$ in X

$$d_Y(x, \phi(x_n)) = \lim_{m\to\infty} d(x_m, x_n) \to 0 \qquad (\text{für } n, m \to \infty),$$

d.h. für alle $x \in Y$ und $\varepsilon > 0$ findet man ein $\tilde{x} \in X$ mit $d_Y(x, \phi(\tilde{x})) < \epsilon$.

Es bleibt nun nur noch die Eindeutigkeit der Vervollständigung bis auf eine isometrische Isomorphie zu zeigen: Dazu seien $(Y, d_Y; \phi)$ und $(W, d_W; \psi)$ zwei Vervollständigungen von (X, d_X). Dann ist die Abbildung

$$\tau = \psi \circ \phi^{-1} : \phi(X) \subset Y \to W$$

ebenfalls eine Isometrie definiert auf der dichten Teilmenge $\phi(X) \subset Y$. Nach Lemma 1.28 besitzt sie eine eindeutige isometrische Fortsetzung $\tilde{\tau}$ auf ganz Y. Des weiteren ist das Bild $\tilde{\tau}(Y)$ des vollständigen Raums Y abgeschlossen und, da das Bild $\psi(X)$ dicht in W liegt, ist auch $\tilde{\tau}(Y) \subset W$ dicht. Zusammenfassend ist also $\tilde{\tau}(Y) = W$, d.h. es gibt einen eindeutigen isometrischen Isomorphismus zwischen ϕ und ψ.

□

Bemerkung: Diese Konstruktion kann benutzt werden, um die Lebesgue-Räume $L^p(\Omega)$ für offenes $\Omega \subset \mathbb{R}^n$ zu definieren. In der Tat ist $L^p(\Omega)$ isometrisch isomorph zu der Vervollständigung des Raumes $C(\Omega)$ der stetigen Funktionen bzgl. der L^p-Norm

$$\|f\|_p = \left(\int_\Omega |f(x)|^p \,\mathrm{d}\lambda(x) \right)^{\frac{1}{p}}.$$

1.3 Normierte Räume

Wichtige Beispiele von metrischen Räumen mit Vektorraumstruktur besitzen eine Norm.

Definition 1.30 (Ana 2)**.** Ein normierter Raum $(X, \|\cdot\|)$ ist ein Paar bestehend aus einem $\mathbb{K}-$Vektorraum X und einer Abbildung $\|\cdot\| : X \to [0, \infty)$ mit den Eigenschaften

(N1) $\|x\| = 0 \quad \Leftrightarrow \quad x = 0$ (Definitheit)

(N2) $\forall \lambda \in \mathbb{K}, x \in X : \quad \|\lambda x\| = |\lambda| \, \|x\|$ (Homogenität)

(N3) $\forall x, y \in X : \quad \|x + y\| \leq \|x\| + \|y\|$ (Dreiecksungleichung)

Bemerkung:

(i) Eine Norm $\|\cdot\|$ induziert auf X eine Metrik $d(x, y) = \|x - y\|$ und somit auch eine Topologie $\mathcal{T}_d$. Eine Topologie, die auf einem Vektorraum definiert ist, heißt **Vektorraum - Topologie** und X heißt ein **topologischer Vektorraum**.

(ii) Eine Metrik d auf einem Vektorraum X definiert die Norm $\|x\| = d(x, 0)$ nur dann, wenn für alle $\lambda \in \mathbb{K}$ und alle $x, y, z \in X$ gilt:

$$d(\lambda x, \lambda y) = |\lambda| \, d(x, y) \qquad \text{(Homogenität)}$$
$$d(x + z, y + z) = d(x, y). \qquad \text{(Translationsinvarianz)}$$

Beispiel 1.31 (Ana 2)**.**

(a) $(\mathbb{K}^n, \|\cdot\|_2)$ ist ein normierter $\mathbb{K}$-Vektorraum für die Euklidsche Norm

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

(b) Sei $K \subset \mathbb{K}^n$ eine kompakte Menge. Der Raum $C_\mathbb{K}(K) := \{f : K \to \mathbb{K} \mid f \text{ ist stetig}\}$ der stetigen Funktionen mit Werten in $\mathbb{K}$ ist versehen mit der Norm

$$\|f\|_\infty = \max_{x \in K} |f(x)|$$

ein normierter $\mathbb{K}$-Vektorraum.

(c) Für die Folgenräume ℓ^p für $1 \leq p \leq \infty$ erfüllt die Abbildung $\|\cdot\|_p$ aus Beispiel 1.20 aufgrund der Minkowski-Ungleichung alle Eigenschaften einer Norm (siehe Analysis 2).

(d) Sind $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ normierte Räume, so ist auch der Produktraum $X \times Y$ normiert für
$$\|(x,y)\|_{X\times Y} := \|x\|_X + \|y\|_Y \qquad \forall\, x \in X, y \in Y.$$

Definition 1.32 (Banachraum)**.** Ein normierter Raum $(X, \|\cdot\|)$ heißt **vollständig**, falls X als metrischer Raum mit der induzierten Metrik $d(x,y) = \|x - y\|$ vollständig ist. Ein solcher vollständiger normierter Raum heißt **Banachraum**.

Beispiel 1.33.

(a) Die Beispiele 1.31 (a)-(c) sind Banachräume. (Siehe Analysis 2)

(b) Sei K ein kompakter Raum. Man definiere den Raum der stetigen Funktionen
$$C_{\mathbb{K}}(K) := \{f : K \to \mathbb{K} \mid f \text{ stetig}\},$$
wobei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ mit der Standardtopologie versehen ist. Nach Lemma 2.4 ist $C_{\mathbb{K}}(K)$ ein Banachraum mit der Norm $\|f\| := \max_{x\in K} |f(x)|$. Dagegen ist $C_{\mathbb{K}}(K)$ für Kompakta $K \subset \mathbb{R}$ mit der L^p-Norm für $1 \leq p < \infty$ unvollständig.

(c) Sei (Ω, Σ, μ) ein Maßraum und $1 \leq p \leq \infty$. Dann sind die Lebesgue-Räume $L^p(\Omega, \mu)$ vollständig normiert. (Siehe Höhere Analysis bzw. Satz 2.47)

(d) Falls $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ Banach-Räume sind, so ist auch der Produktraum $(X \times Y, \|\cdot\|_X + \|\cdot\|_Y)$ ein Banachraum.

Da zwar nach Analysis 2 jeder endlich-dimensionale normierte Raum vollständig ist, aber nicht jeder normierte Raum, zieht man die zugehörige metrische Vervollständigung in Betracht und erhält

Korollar 1.34. *Sei $(X, \|\cdot\|)$ ein normierter $\mathbb{K}$-Vektorraum und (Y, d_Y) die entsprechende metrische Vervollständigung aus Satz 1.29. Dann lässt sich auf Y die Vektorraumstruktur fortsetzen, sodass die Isometrie $\phi : X \to Y$ $\mathbb{K}$-linear ist, mit Norm $\|y\|_Y = d_Y(0, y)$.*

Beweis. Aufgrund der Dreiecksungleichung sind die Addition $+ : X \times X \to X$ und die Norm $\|\cdot\| : X \to \mathbb{R}$ Lipschitz-stetige Funktionen. Außerdem ist die skalare Multiplikation $\cdot : \mathbb{K} \times X \to X$ stetig und gleichmäßig stetig auf beschränkten Teilmengen. Der Beweis von Lemma 1.28 zeigt, dass die eindeutige Fortsetzbarkeit (gleichmäßig stetig auf beschränkten Teilmengen) auch in diesem Fall gültig bleibt. Dies liefert eine Norm-erhaltende Fortsetzung zu einem Banach-Raum. □

Bereits in Analysis 2 wurde die Äquivalenz von Normen bzw. der Satz von Bolzano-Weierstraß für endlich-dimensionale normierte $\mathbb{K}$-Vektorräume gezeigt. Ebenso folgte

Satz 1.35 (Heine-Borel)**.** *Sei $(X, \|\cdot\|)$ ein endlich-dimensionaler normierter Raum und $A \subset X$. Dann gilt*
$$A \text{ ist kompakt} \quad \Leftrightarrow \quad A \text{ ist beschränkt und abgeschlossen.}$$

Beweis. Siehe Analysis 1 und 2. □

Im Allgemeinen gelten diese Aussagen jedoch nicht, wie nachstehender Satz zeigt. Dazu benötigen wir folgendes

Lemma 1.36 (Riesz'sches Lemma). *Sei $A \subsetneq X$ ein abgeschlossener, echter Untervektorraum eines normierten $\mathbb{K}$-Vektorraums $(X, \|\cdot\|)$. Dann existiert für jedes $\theta \in (0,1)$ ein $x_\theta \in X$ mit $\|x_\theta\| = 1$ und*

$$\|x_\theta - a\| \geq 1 - \theta \qquad \forall\, a \in A.$$

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

$$d := \operatorname{dist}(x, A) = \inf\{\|x - a\| \mid a \in A\} > 0,$$

denn sonst gäbe es eine Folge $(a_n)_{n\in\mathbb{N}}$ in A mit $\|x - a_n\| \to 0$, d.h. $x \in \overline{A} = A$.
Mit $d(1-\theta) < d$ gibt es dann ein $a_\theta \in A$ mit $\|x - a_\theta\| < d/(1-\theta)$. Definiere

$$x_\theta := \frac{x - a_\theta}{\|x - a_\theta\|} \qquad \text{mit} \quad \|x_\theta\| = 1,$$

so folgt für beliebiges $a \in A$ wegen $a_\theta + \|x - a_\theta\| a \in A$

$$\|x_\theta - a\| = \frac{1}{\|x - a_\theta\|}\|x - (a_\theta + \|x - a_\theta\| a)\| \geq \frac{d}{\|x - a_\theta\|} > 1 - \theta.$$

□

In allgemeinen normierten Räumen gibt es zwar keine Definition von Orthogonalität, jedoch liefert das obige Lemma ein fast orthogonales Element. In endlich dimensionalen Räumen, wie beispielsweise $\mathbb{R}^n$, ist dabei auch $\theta = 0$ zugelassen, wie man sich leicht in $\mathbb{R}^n$ mithilfe des Einheitsvektors auf dem abgeschlossenen Untervektorraum A klar macht. Abgesehen von dieser geometrischen Interpretation liefert das Lemma

Satz 1.37. *Für einen normierten $\mathbb{K}$-Vektorraum $(X, \|\cdot\|)$ sind äquivalent:*

(i) $\dim X < \infty$.

(ii) $\overline{B_1(0)} \subset X$ *ist kompakt.*

(iii) Jede beschränkte Folge in X besitzt eine konvergente Teilfolge.

Beweis. Siehe Übungen.

□

1.4 Hilberträume

Eine weitere Struktur, welche geometrisch zur Längen- und Winkelmessung benutzt werden kann, ist das Skalarprodukt auf einem Vektorraum.

Definition 1.38 (Ana 2). Sei H ein $\mathbb{K}$-Vektorraum. Ein **Skalarprodukt** auf H ist eine Abbildung $(\cdot,\cdot) : H \times H \to \mathbb{K}$, welche folgende Eigenschaften für alle $x, y, z \in H, \lambda \in \mathbb{K}$ erfüllt:

(S1) $(z, x+\lambda y) = (z,x) + \lambda(z,y)$ (Linearität im 2. Argument)

(S2) $(x,y) = \overline{(y,x)}$ (Symmetrie)

(S3) $(x,x) \geq 0 \quad$ und $\quad (x,x) = 0 \quad \Leftrightarrow \quad x = 0$ (Definitheit)

Das Paar $(H, (\cdot,\cdot))$ nennt man einen **Prähilbertraum**.

Für $\mathbb{K} = \mathbb{C}$ ist das Skalarprodukt linear in der zweiten Komponente, aber antilinear in der ersten, d.h. $(\lambda x, y) = \overline{\lambda}(x,y)$. Im reellen Fall ist das Skalarprodukt bilinear, d.h. linear in beiden Argumenten.
Ein grundlegendes Resultat aus Analysis 2 ist die folgende

Lemma 1.39 (Cauchy-Schwarz Ungleichung)**.** *Sei $(H, (\cdot,\cdot))$ ein Prähilbertraum. Dann gilt*

$$|(x,y)|^2 \leq (x,x)(y,y) \qquad \forall\, x,y \in H.$$

Beweis. Da die Ungleichung für $y = 0$ bereits erfüllt ist, sei o.B.d.A. $y \neq 0$. Für ein beliebiges $\alpha \in \mathbb{K}$ gilt

$$0 \leq (x+\alpha y, x+\alpha y) = (x,x) + \overline{\alpha}(y,x) + \alpha(x,y) + \alpha\overline{\alpha}(y,y).$$

Setzt man $\alpha := -(y,x)(y,y)^{-1}$ ein, so folgt

$$\begin{aligned}
0 &\leq (x,x) - \overline{(y,x)}(y,y)^{-1}(y,x) - (y,x)(y,y)^{-1}(x,y) + |(x,y)|^2(y,y)^{-1} \\
&= (x,x) - 2|(x,y)|^2(y,y)^{-1} + |(x,y)|^2(y,y)^{-1} \\
&= (x,x) - |(x,y)|^2(y,y)^{-1}.
\end{aligned}$$

□

Korollar 1.40 (Ana 2)**.** *Sei $(H, (\cdot,\cdot))$ ein Prähilbertraum. Dann definiert $\|x\| := \sqrt{(x,x)}$ eine Norm auf H.*

Beweis. Es ist nur die Dreieckungleichung nachzuweisen. $(N1), (N2)$ sind ganz klar. Für $x, y \in H$ gilt

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

□

Definition 1.41 (Ana 2)**.** Ein Prähilbertraum $(H, (\cdot,\cdot))$ heißt **Hilbertraum**, falls das Paar $(H, \|\cdot\|)$ mit der induzierten Norm $\|x\| := \sqrt{(x,x)}$ ein Banachraum ist.

Beispiel 1.42 (Ana 2)**.**

(a) Für $H = \mathbb{K}^n$ mit $(x,y) := \sum_{i=1}^n \overline{x_i} y_i$ ebenfalls ein Hilbertraum.

(b) Der Folgenraum ℓ^2 ist mit dem Skalarprodukt $(x,y) := \sum_{i=1}^\infty \overline{x_i} y_i$ ein Hilbertraum nach den Übungen.

(c) Der Raum $C([a,b])$ für $a,b \in \mathbb{R}, a < b$ versehen mit dem Skalarprodukt

$$(f,g) := \int_a^b \overline{f(x)} g(x) \, \mathrm{d}x$$

ist ein Prähilbertraum, aber kein Hilbertraum.

Hilberträume besitzen mehr Struktur als Banachräume, z.B. ist die Maximumsnorm auf $\mathbb{R}^n$ nicht skalarproduktinduziert, sonst müsste sie die folgende Parallelogrammidentität erfüllen.

Lemma 1.43 (Parallelogrammidentität)**.** *Sei $(H,(\cdot,\cdot))$ ein Prähilbertraum. Dann gilt für die induzierte Norm*

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

Beweis. Mit den Rechnungen aus Korollar 1.40 erhält man

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

Die Addition der beiden Gleichung ergibt

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

□

Korollar 1.44. *Sei $(H,(\cdot\,,\cdot))$ ein Prähilbertraum. Dann ist die Vervollständigung bzgl. der von $(\cdot\,,\cdot)$ induzierten Metrik ein Hilbertraum mit natürlicher Fortsetzung von $(\cdot\,,\cdot)$.*

Beweis. Nach Korollar 1.34 genügt es die gleichmäßige Stetigkeit des Skalarprodukts auf beschränkten Mengen zu prüfen. Dies erfolgt analog zur skalaren Multiplikation. □

In Prähilberträumen lässt sich der Begriff der Orthogonalität einführen und analog zum Riesz'schen Lemma die Projektion auf eine abgeschlossene Menge untersuchen.

Definition 1.45. Eine Menge $A \subset X$ eines $\mathbb{K}$-Vektorraums heißt **konvex**, falls für jedes reelle $\lambda \in [0,1]$ und $x,y \in A$ gilt: $\lambda x + (1-\lambda) y \in A$.

Satz 1.46 (Projektionsabbildung)**.** *Sei $(H,(\cdot\,,\,\cdot))$ ein Hilbertraum und $A \subset H$ eine nichtleere, abgeschlossene und konvexe Menge. Dann existiert genau eine Abbildung*

$$P : H \to A \qquad \textit{mit} \qquad \|x - P(x)\| = \mathrm{dist}\,(A,x) = \inf_{a \in A} \|a - x\| \quad \forall\, x \in H.$$

Man nennt die Abbildung P ***Projektionsabbildung*** *oder* ***orthogonale Projektion*** *von H auf A.*

Beweis. Für $x \in H$ definiere man $d := \mathrm{dist}\,(x,A)$ und wähle eine Minimalfolge $(x_n)_{n \in \mathbb{N}}$ in A mit $\|x - x_n\| \to d$ für $n \to \infty$. Dann folgt mit der Parallelogrammidentität

$$\begin{aligned} \|x_n - x_m\|^2 &= \|x_n - x + x - x_m\|^2 \\ &\le 2\left(\|x_n - x\|^2 + \|x_m - x\|^2\right) - 4 \cdot \left\| \frac{x_n + x_m}{2} - x \right\|^2. \end{aligned}$$

Nun ist $\frac{1}{2}(x_m + x_n) \in A$ als Konvexkombination und $(x_n)_{n\in\mathbb{N}}$ ist eine Cauchy-Folge, denn

$$\|x_n - x_m\|^2 \leq 2\left(\|x_n - x\|^2 + \|x_m - x\|^2\right) - 4\,d^2 \to 0 \qquad (\text{für } m, n \to \infty).$$

Da $A \subset H$ als abgeschlossene Teilmenge eines vollständigen Raumes selbst vollständig ist, gibt es ein $y \in A$ mit $x_n \to y$ für $n \to \infty$. Setze $P(x) := y$.

Für die Eindeutigkeit nehmen wir an, dass $y_1, y_2 \in A$ existieren, sodass

$$d = \|x - y_1\| = \|x - y_2\|.$$

Dann gilt für die Folge $(z_n)_{n\in\mathbb{N}} = (y_1, y_2, y_1, y_2, \ldots)$ gerade $\|z_n - x\| \to d$ für $n \to \infty$. Daraus folgt wie oben, dass $(z_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge ist. Wählt man $\epsilon = \frac{1}{2}\|y_1 - y_2\|$, so folgt $y_1 = y_2$. □

Im weiteren Verlauf wird sich zeigen, dass diese Minimierungsaufgabe auch in bestimmten Banachräumen lösbar ist. In einem Hilbertraum lässt sich die Lösung des obigen Variationsproblems wie folgt geometrisch über den Winkel charakterisieren (beachte dabei den Kosinussatz).

Lemma 1.47. *Sei $(H, (\cdot\,,\cdot))$ ein Hilbertraum und $A \subset H$ eine nicht-leere, abgeschlossene und konvexe Menge. Dann ist für $x \in H$ und die zugehörige Projektionsabbildung äquivalent:*

(i) $\|x - P(x)\| = \inf_{a\in A}\|a - x\|$.

(ii) $\mathrm{Re}\,(x - P(x), a - P(x)) \leq 0 \quad \forall\, a \in A$.

Beweis. Es sind stets $a, P(x) \in A$ für jedes $x \in H$.

(i) ⇒ (ii): Sei $\lambda \in (0, 1]$. Da A konvex ist, gilt $((1-\lambda)P(x) + \lambda\, a) \in A$ und folglich

$$\begin{aligned}\|x - P(x)\|^2 &\leq \|x - \big((1-\lambda)P(x) + \lambda\, a\big)\|^2\\ &= \|x - P(x) - \lambda(a - P(x))\|^2\\ &= \|x - P(x)\|^2 - 2\ \mathrm{Re}\big(x - P(x), \lambda\,(a - P(x))\big) + \|\lambda\,(a - P(x))\|^2.\end{aligned}$$

Division durch $\lambda > 0$ ergibt

$$\mathrm{Re}\big(x - P(x), a - P(x)\big) \leq \frac{\lambda}{2}\|a - P(x)\|^2 \qquad \forall\ \lambda(0, 1].$$

Für $\lambda \to 0$ folgt die Behauptung.

(ii) ⇐ (i): Für jedes $a \in A$ ist

$$\begin{aligned}\|x - a\|^2 &= \|\,(x - P(x)) + (P(x) - a)\,\|^2\\ &= \|x - P(x)\|^2 + 2\ \mathrm{Re}\big(x - P(x), P(x) - a\big) + \|P(x) - a\|^2\\ &\geq \|x - P(x)\|^2.\end{aligned}$$

Hieraus folgt $\|x - P(x)\| \leq \inf\limits_{a\in A}\|x - a\|$.

□

Ein wichtiger Spezialfall ist das folgende

Korollar 1.48 (Projektion auf Unterräume). *Sei $(H, (\cdot\,,\cdot))$ ein Hilbertraum und $Y \subset H$ ein abgeschlossener Unterraum. Dann existiert genau eine Abbildung $P : H \to Y$ mit*

$$\|x - P(x)\| = \operatorname{dist}(x, Y).$$

Die Abbildung P ist linear und äquivalent charakterisiert durch

$$(x - P(x), y) = 0 \quad \forall\, y \in Y. \tag{1.1}$$

Beweis. Aus Satz 1.46 folgt die Existenz und Eindeutigkeit von P und es gilt die Charakterisierung

$$\operatorname{Re}\big(x - P(x), a - P(x)\big) \leq 0 \quad \forall\, a \in Y. \tag{1.2}$$

Wähle $a = P(x) \pm iy$ und $a = P(x) \pm y$ und man erhält (1.1). Umgekehrt impliziert (1.1) immer (1.2) mit $y = a - P(x) \in Y$.
Für die Linearität seien $x_1, x_2 \in H$ und $\lambda \in \mathbb{K}$. Aus (1.1) folgt

$$\big(x_1 + x_2 - (P(x_1) + P(x_2)), y\big) = 0 = \big(\lambda\, x_1 - \lambda\, P(x_1), y\big) \quad \forall\, y \in Y.$$

Aus der Eindeutigkeit folgt nun $P(x_1) + P(x_2) = P(x_1 + x_2)$ und $P(\lambda\, x_1) = \lambda\, P(x_1)$. □

Definition 1.49. Sei $(H, (\cdot\,,\cdot))$ ein Hilbertraum

(i) Zwei Elemente $x, y \in H$ heißen **orthogonal**, falls $(x, y) = 0$.

(ii) Eine Familie $e = (e_i)_{i \in I}$ (wobei I höchstens abzählbar sei) heißt **Orthogonalsystem**, wenn für alle $k, l \in I$ gilt $e_k \neq 0$ und

$$(e_k, e_l) = \|e_k\|^2 \delta_{kl}.$$

Falls $\|e_k\| = 1$ für alle $k \in I$, heißt e **Orthonormalsystem** (ONS).

(iii) Ein Orthonormalsystem e heißt **Orthonormalbasis** (ONB), wenn die lineare Hülle von e dicht in H liegt, d.h.

$$\overline{\operatorname{span}\{e_k \mid k \in \mathbb{N}\}} = H.$$

Eine Anwendung von Satz 1.46 ist die orthogonale Projektion gemäß

Satz 1.50. *Sei $(H, (\cdot\,,\cdot))$ ein Hilbertraum und $A \subset H$ ein linearer, abgeschlossener Unterraum. Dann ist*

$$A^\perp := \{x \in H \mid (x, a) = 0 \quad \forall\, a \in A\}$$

ein linearer, abgeschlossener Unterraum von H und jedes $z \in H$ lässt sich eindeutig schreiben als

$$z = x + y \qquad \text{mit} \quad x \in A, y \in A^\perp,$$

*d.h. $H = A \oplus A^\perp$. Dabei heißt $A^\perp$ das **orthogonale Komplement** von A.*

Beweis. Dass $A^\perp$ ein abgeschlossener Untervektorraum ist, zeigt die Stetigkeit des Skalarprodukts. Für $z \in H$ sei $x = P(z) \in A$ die Projektion auf A aus Korollar 1.48 und $y := z - x$. Wir zeigen zunächst $y \in A^\perp$:
Sei $\lambda \in \mathbb{C}$ und $a \in A$ beliebig sowie $d = \|y\| = \|z - x\|$. Da x das Variationsproblem löst, gilt

$$d^2 \leq \|z - (\lambda a + x)\|^2 = \|y - \lambda a\|^2 = d^2 - 2\,\mathrm{Re}\,(y, \lambda a) + |\lambda|^2 \|a\|^2$$

und somit

$$2\,\mathrm{Re}\,(y, \lambda a) \leq |\lambda|^2 \|a\|^2 \qquad \forall\, \lambda \in \mathbb{C}.$$

Schreibe nun $\lambda = r\mathrm{e}^{-\mathrm{i}\theta}$ für $r, \theta \in \mathbb{R}_+$ mit $\mathrm{e}^{-\mathrm{i}\theta}(y, a) = |(y, a)|$. Dann folgt

$$2r|(y, a)| \leq r^2 \|a\|^2 \qquad \forall\, r \in \mathbb{R}_+,$$

d.h. $(y, a) = 0$ für jedes $a \in A$ und somit $y \in A^\perp$.
Falls es zwei verschiedene Zerlegungen

$$z = x_1 + y_1 = x_2 + y_2 \qquad \text{für } x_i \in A, y_i \in A^\perp$$

gäbe, dann ist $x_1 - x_2 = y_2 - y_1 \in A \cap A^\perp = \{0\}$. □

Eine Verallgemeinerung des Satzes von Pythagoras ist

Lemma 1.51 (Bessel'sche Ungleichung)**.** *Sei $e_1, \ldots, e_n$ ein endliches Orthonormalsystem eines Prähilbertraums $(H, (\cdot\,,\cdot))$. Dann gilt:*

$$0 \leq \|x\|^2 - \sum_{i=1}^n |(x, e_i)|^2 = \left\| x - \sum_{i=1}^n (x, e_i) e_i \right\|^2 = \Big(\mathrm{dist}\,\big(x, \mathrm{span}\,\{e_1, \ldots, e_n\}\big)\Big)^2.$$

Beweis. Für $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ gilt:

$$\begin{aligned} 0 \leq \left\| x - \sum_{i=1}^n \alpha_i e_i \right\|^2 &= \|x\|^2 - \sum_{i=1}^n (x, e_i)\,\overline{\alpha}_i - \sum_{i=1}^n (e_i, x)\,\alpha_i + \sum_{i=1}^n |\alpha_i|^2 \\ &= \|x\|^2 - \sum_{i=1}^n |(x, e_i)|^2 + \sum_{i=1}^n |(x, e_i) - \alpha_i|^2. \end{aligned}$$

Die rechte Seite wird für $\alpha_i = (x, e_i)$ minimal und das Minimum ist gegeben durch

$$\|x\|^2 - \sum_{i=1}^n |(x, e_i)|^2.$$

□

Lemma 1.52. *Sei $(H, (\cdot\,,\cdot))$ ein Hilbertraum und $(e_k)_{k\in\mathbb{N}}$ ein Orthonormalsystem. Sei $(\alpha_k)_{k\in\mathbb{N}}$ eine Folge in $\mathbb{K}$. Dann gilt:*

(i) $\sum\limits_{k=1}^\infty \alpha_k e_k$ *ist konvergent in H* $\quad\Leftrightarrow\quad \sum\limits_{k=1}^\infty |\alpha_k|^2 < \infty.$

Im Falle $\sum\limits_{k=1}^\infty |\alpha_k|^2 < \infty$ gilt außerdem

$$\left\| \sum_{k=1}^\infty \alpha_k e_k \right\|^2 = \sum_{k=1}^\infty |\alpha_k|^2.$$

(ii) Wenn die Reihe $\sum_{k=1}^{\infty} \alpha_k e_k$ *konvergiert, dann ist der Grenzwert unabhängig von der Summationsreihenfolge, d.h. die Reihe konvergiert* ***unbedingt****.*

Beweis. (i) Siehe Übungen.

(ii) Sei $\phi : \mathbb{N} \to \mathbb{N}$ eine Bijektion. Dann gilt für $n \in \mathbb{N}$

$$\left\| \sum_{k=1}^{n} \alpha_{\phi(k)} e_{\phi(k)} \right\|^2 = \sum_{k=1}^{n} |\alpha_{\phi(k)}|^2.$$

Da $x := \sum_{k=1}^{\infty} \alpha_k e_k$ konvergiert, ist die Koeffizientenfolge $(\alpha_k)_{k\in\mathbb{N}} \in \ell^2$ quadratsummierbar. Folglich ist nach den Umordnungssätzen aus Analysis 1

$$\sum_{k=1}^{\infty} |\alpha_{\phi(k)}|^2 = \sum_{k=1}^{\infty} |\alpha_k|^2 = \|x\|^2 < \infty.$$

Nach (i) konvergieren damit die Reihen x und $y := \sum_{k=1}^{\infty} \alpha_{\phi(k)} e_{\phi(k)}$.
Es bleibt $x = y$ zu zeigen. Offensichtlich ist

$$\|x - y\|^2 = \|y\|^2 + \|x\|^2 - 2 \operatorname{Re}(x, y) = 2 \sum_{k=1}^{\infty} |\alpha_k|^2 - 2 \operatorname{Re}(x, y).$$

Andererseits gilt aufgrund der Stetigkeit des Skalarprodukts

$$(x, y) = \lim_{n,m\to\infty} \left(\sum_{k=1}^{n} \alpha_k e_k, \sum_{k=1}^{m} \alpha_{\phi(k)} e_{\phi(k)} \right).$$

Wähle nun für fixes $m \in \mathbb{N}$ ein $n_m \in \mathbb{N}$ so groß, dass $\{\phi(1), \ldots, \phi(m)\} \subset \{1, \ldots, n_m\}$.
Dann folgt die Behauptung für $n, m \to \infty$, denn

$$\left(\sum_{k=1}^{n_m} \alpha_k e_k, \sum_{k=1}^{m} \alpha_{\phi(k)} e_{\phi(k)} \right) \overset{\text{ONS}}{=} \sum_{k=1}^{m} |\alpha_{\phi(k)}|^2 \to \sum_{k=1}^{\infty} |\alpha_k|^2.$$

□

Satz 1.53. *Sei* $(H, (\cdot\,,\cdot))$ *ein Hilbertraum und* $(e_n)_{n\in\mathbb{N}}$ *ein Orthonormalsystem in* H*. Dann sind die folgende Aussagen äquivalent:*

(i) $(e_n)_{n\in\mathbb{N}}$ *ist eine Orthonormalbasis in* H*.*

(ii) $x = \sum_{k=1}^{\infty} (e_k, x) e_k \quad \forall\, x \in H$

(iii) $\|x\|^2 = \sum_{k=1}^{\infty} |(e_k, x)|^2 \quad \forall\, x \in H$ *(Parseval'sche Gleichung)*

(iv) Für $x \in H$ *mit* $(e_k, x) = 0$ *für alle* $k \in \mathbb{N}$ *folgt* $x = 0$*.*

Beweis. **(i)⇒(ii):** Sei $x \in H$, dann gibt es aufgrund der Dichtheit Elemente

$$x_n = \sum_{k=1}^{m_n} \alpha_{nk} e_k \to x \qquad (\text{für } n \to \infty).$$

Für $n, m \in \mathbb{N}$ mit $m \geq m_n$ folgt dann wegen der Bessel'schen Ungleichung

$$\begin{aligned} \left\| x - \sum_{k=1}^{m} (e_k, x) e_k \right\| &= \operatorname{dist}(x, \operatorname{span}\{e_1, \dots, e_m\}) \\ &\leq \operatorname{dist}(x, \operatorname{span}\{e_1, \dots, e_{m_n}\}) \leq \|x - x_n\|. \end{aligned}$$

(ii)⇒(iii): Dies folgt aus Lemma 1.52.

(iii)⇒(iv): Mit (iii) folgt $\|x\|^2 = \sum\limits_{k=1}^{\infty} |(e_k, x)|^2 = 0$, d.h. $x = 0$.

(iv)⇒(i): Sei $M = \overline{\operatorname{span}\{e_k \mid k \in \mathbb{N}\}}$. Wäre $M \neq H$, so gibt es ein ohne Einschränkung normiertes $x \in H \setminus M$ mit $\|x\| = 1$. Wegen Satz 1.50 ist $H = M \oplus M^\perp$ und $x \in M^\perp \setminus \{0\}$, d.h. $(x, e_k) = 0$ für alle $k \in \mathbb{N}$ und nach (iv) $x = 0$.

□

Satz 1.54. *Sei $(H, (\cdot\,,\cdot))$ ein unendlich-dimensionaler Hilbertraum. Dann sind folgende Aussagen äquivalent:*

(i) H ist separabel.

(ii) H besitzt eine Orthonormalbasis.

Gilt eine der beiden Aussagen, so ist H isometrisch isomorph zu ℓ^2.

Beweis.

(i) ⇒ (ii): Wir konstruieren eine ONB nach dem Gram-Schmidt-Verfahren (analog zur Linearen Algebra):
Sei $E \subset H$ abzählbar und dicht in H, etwas $E = \{y_n \mid n \in \mathbb{N}\}$ als Folge gegeben. Definiere induktiv $z_1 := y_1$ (o.E. $y_1 \neq 0$) und

$$z_k := y_{n_k} \quad \text{mit} \quad n_k = \min\{n \in \mathbb{N} \mid z_1, \dots, z_{k-1}, y_n \text{ linear unabhängig}\} \geq k.$$

Dann ist auch $E \subset \bigcup_{k \in \mathbb{N}} H_k$ für

$$H_k := \operatorname{span}\{z_1, \dots, z_k\} \quad \text{mit} \quad \dim(H_k) = k,$$

d.h. die linear unabhängige Teilfolge $(z_k)_{k \in \mathbb{N}}$ liegt ebenfalls dicht in H.
Für die Orthonormalisierung setze $e_1 := \frac{z_1}{\|z_1\|}$ und definiere induktiv

$$e_{n+1} = \frac{\tilde{e}_{n+1}}{\|\tilde{e}_{n+1}\|} \quad \text{mit} \quad \tilde{e}_{n+1} = z_{n+1} - \sum_{k=1}^{n} (z_{n+1}, e_k)\, e_k \quad \text{für} \quad z_{n+1} \in H_{n+1} \setminus H_n.$$

Man sieht leicht, dass $(e_n)_{n \in \mathbb{N}}$ ein Orthonormalsystem von H ist. Außerdem ist $H_n = \operatorname{span}\{e_1, \dots, e_n\}$ und $\operatorname{span}\{e_n \mid n \in \mathbb{N}\}$ liegt ebenfalls dicht in H, d.h. $(e_n)_{n \in \mathbb{N}}$ ist eine ONB.

(ii) $\Rightarrow$ (i): Sei $(e_k)_{k\in\mathbb{N}}$ eine ONB von H. Für jedes $n \in \mathbb{N}$ ist

$$A_n := \left\{ \sum_{k=1}^{n} \alpha_k e_k \,\middle|\, \alpha_k \in \mathbb{K} \text{ rational für } 1 \leq k \leq n \right\}$$

abzählbar, wobei wir im Fall $\mathbb{K} = \mathbb{C}$ rationale $\alpha_k \in \mathbb{Q} + \mathrm{i}\mathbb{Q}$ meinen. Mit

$$\overline{A_n} = \operatorname{span}\{e_k \mid 1 \leq k \leq n\}$$

ist A_n separabel und folglich auch der Abschluss $\overline{\bigcup_{n\in\mathbb{N}} A_n}$ der abzählbaren Vereinigung. Da jedoch die Dichtheit der linearen Hülle der ONB gerade $H = \overline{\bigcup_{n\in\mathbb{N}} A_n}$ impliziert, ist auch H separabel.

Sei $(e_n)_{n\in\mathbb{N}}$ eine Orthonormalbasis von H. Man definiere

$$\phi : \ell^2 \to X, \quad \alpha \mapsto \sum_{k=1}^{\infty} \alpha_k e_k.$$

Aus Satz 1.53 folgt nun, dass ϕ eine Isometrie zwischen ℓ^2 und H ist und aus der Definition einer Orthonormalbasis folgt die Surjektivität. $\square$

Im endlich-dimensionalen Fall ist jeder Hilbert-Raum H separabel, besitzt nach Linearer Algebra ebenso eine Orthonormalbasis und ist isometrisch isomorph zu $\mathbb{K}^{\dim H}$.

Beispiel 1.55. (a) Für den Folgenraum ℓ^2 ist die Menge der Einheitsvektoren $(e_k)_{k\in\mathbb{N}}$ eine Orthonormalbasis, wobei

$$e_1 = (1, 0, 0, 0, \ldots),\; e_2 = (0, 1, 0, 0, \ldots),\; e_3 = (0, 0, 1, 0, \ldots),\; \ldots$$

(b) Für den Lebesgue-Raum $L^2([-1, 1])$ mit dem Skalarprodukt aus Beispiel 1.42 (c) sind beispielsweise die Legendre-Polynome $(P_n)_{n\in\mathbb{N}_0}$ mit

$$P_n(x) := \left(\frac{2n+1}{2}\right)^{1/2} \frac{1}{2^n n!} \frac{\mathrm{d}^n}{\mathrm{d}x^n}(x^2 - 1)^n \qquad \text{für } x \in [-1, 1]$$

eine Orthonormalbasis.

2 Funktionenräume

In diesem Kapitel werden die wichtigsten Funktionenräume der Analysis tiefergehend studiert und elementare Eigenschaften aus Kapitel 1 hergeleitet.
Neben den klassischen stetigen sowie differenzierbaren Funktionen werden wir die Lebesgue-Räume der Höheren Analysis nochmals rekapitulieren und die verallgemeinerte Differentiation in sogenannten Sobolev-Räumen untersuchen. Diese Räume sind dabei fundamental für die Lösungstheorie partieller Differentialgleichungen wie wir sehen werden.

2.1 Beschränkte und stetige Funktionen

Definition 2.1. Sei X eine Menge und $(Y, \|\cdot\|)$ ein normierter Raum. Man definiere den Raum der beschränkten Funktionen $B(X,Y)$ durch

$$B(X,Y) := \{f : X \to Y \mid f(X) \text{ ist beschränkte Teilmenge von } Y\}.$$

Dieser ist mit $(f_1 + f_2)(x) := f_1(x) + f_2(x)$ und $(\alpha f)(x) = \alpha f(x)$ ein Vektorraum mit Norm

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

Lemma 2.2. *Sei X eine Menge und $(Y, \|\cdot\|)$ ein Banachraum. Dann ist $B(X,Y)$ mit der Norm $\|\cdot\|_{B(X,Y)}$ ein Banach-Raum.*

Beweis. Die Norm-Eigenschaften sind klar. Zum Beweis der Vollständigkeit sei $(f_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $B(X,Y)$. Dann folgt in Y

$$\|f_m(x) - f_n(x)\| \leq \|f_n - f_m\|_\infty \to 0 \qquad (\text{für } m,n \to \infty)$$

und die Vollständigkeit von Y liefert für jedes $x \in X$ einen Grenzwert, mit welchen wir eine Funktion $f : X \to Y$ mit $f(x) := \lim_{n\to\infty} f_n(x)$ in Y definieren können. Wegen

$$\|f(x) - f_n(x)\| = \lim_{m\to\infty} \|f_m(x) - f_n(x)\| \leq \liminf_{m\to\infty} \|f_m - f_n\|_\infty$$

ist f ebenfalls beschränkt, d.h. $f \in B(X,Y)$. Außerdem folgt damit die Konvergenz $f_n \to f$ in $B(X,Y)$:

$$\|f - f_n\|_\infty \leq \liminf_{m\to\infty} \|f_m - f_n\|_\infty \to 0 \qquad (\text{für } n \to \infty).$$

□

Offensichtlich ist $B(X,Y)$ der größte Funktionenraum, auf welchem die Supremumsnorm $\|\cdot\|_\infty$ definiert ist. Wichtige Untervektorräume sind

Definition 2.3. Sei X ein topologischer Raum und $(Y, \|\cdot\|)$ ein normierter Raum. Dann definiert man den Raum der stetigen Funktionen

$$C(X,Y) := \{f : X \to Y \mid f \text{ ist stetig}\}.$$

Im Fall $Y = \mathbb{K}$ schreibt man auch $C_{\mathbb{K}}(X) = C(X,\mathbb{K})$ bzw. im reellen Fall vereinfacht $C(X) = C(X,\mathbb{R})$.
Als wichtige Spezialfälle seien die Räume der beschränkten stetigen Funktionen

$$C_b(X,Y) := C(X,Y) \cap B(X,Y)$$

und die stetigen Funktionen mit kompaktem Träger

$$C_c(X,Y) := \{f \in C(X,Y) \mid \operatorname{supp} f \text{ ist kompakt}\}$$

definiert, wobei $\operatorname{supp} f := \overline{\{x \in X \mid f(x) \neq 0\}}$ den **Träger** von f bezeichnet.

Lemma 2.4. *Sei X ein topologischer Raum und $(Y, \|\cdot\|)$ ein Banachraum. Dann ist $C_b(X,Y)$ ein Banachraum mit der Norm $\|\cdot\|_\infty$. Ist $K \subset \mathbb{K}^n$ kompakt und $Y = \mathbb{K}$, so ist $(C_{\mathbb{K}}(K), \|\cdot\|_\infty)$ ein Banach-Raum und es gilt*

$$\|f\|_\infty = \sup_{x\in K} |f(x)| = \max_{x\in K} |f(x)| \qquad \forall\, f \in C_{\mathbb{K}}(K).$$

Beweis. Nach Lemma 2.2 genügt es für die Vollständigkeit zu zeigen, dass $C_b(X,Y)$ ein abgeschlossener Untervektorraum von $B(X,Y)$ ist. Sei also $(f_n)_{n\in\mathbb{N}}$ eine konvergente Folge mit Grenzwert $f \in B(X,Y)$. Wir zeigen mittels gleichmäßiger Konvergenz, dass auch f stetig ist:
Zu $\varepsilon > 0$ findet sich ein $n \in \mathbb{N}$ mit $\|f - f_n\|_\infty < \frac{\varepsilon}{3}$ und sei $x_0 \in X$ beliebig. Dann gibt es aufgrund der Stetigkeit von f_n eine Umgebung U von x_0 mit $\|f_n(x) - f_n(x_0)\| < \frac{\varepsilon}{3}$ für alle $x \in U$. Damit ist für $x \in U$ ebenfalls

$$\begin{aligned}\|f(x) - f(x_0)\| &\leq \|f(x) - f_n(x)\| + \|f_n(x) - f_n(x_0)\| + \|f_n(x_0) - f(x_0)\| \\ &\leq 2\|f - f_n\|_\infty + \|f_n(x) - f_n(x_0)\| < \varepsilon.\end{aligned}$$

Sei nun $f \in C_{\mathbb{K}}(K)$ für ein Kompaktum $K \subset \mathbb{K}^n$. Wir zeigen, dass dann f beschränkt ist und das Supremum und Infimum angenommen wird. Mit oben Gezeigtem folgt dann die Behauptung.
Für jedes $x \in K$ gibt es wegen der Stetigkeit von f eine offene Umgebung U_x von x mit $f(U_x) \subset B_1(f(x))$, wobei $B_1(f(x))$ die offene Kugel mit Radius 1 um $f(x)$ ist. Betrachtet man die Überdeckung $\bigcup_{x\in K} U_x = K$, so gibt es aufgrund der Kompaktheit endlich viele solcher Umgebungen $U_{x_1}, \dots, U_{x_n}$, die bereits K überdecken. Dies impliziert, dass f beschränkt ist:

$$\sup_{x\in K} |f(x)| = \max_{i=1,\dots,n} \sup_{x\in U_{x_i}} |f(x)| \leq \max_{i=1,\dots,n} (|f(x_i)| + 1) < \infty.$$

Sei nun $(x_n)_{n\in\mathbb{N}}$ eine Maximalfolge in K, d.h.

$$|f(x_n)| \to \sup_{y\in K} |f(y)| =: s \qquad (\text{für } n \to \infty).$$

Da K kompakt ist, besitzt die Folge $(x_n)_{n\in\mathbb{N}}$ nach Satz 1.25 eine konvergente Teilfolge mit Grenzwert $x \in K$. Die Stetigkeit von f impliziert dann $f(x_{n_k}) \to f(x)$ für $k \to \infty$. Die Eindeutigkeit von Grenzwerten in $(\mathbb{K}, |\cdot|)$ liefert $|f(x)| = s$. In ähnlicher Weise zeigt man, dass das Infimum angenommen wird. □

Da stetige Funktionen auf kompakten Mengen beschränkt sind, ist $C_{\mathbb{K}}(K)$ ein Banach-Raum. Funktionen in $C_c(X,Y)$ sind ebenso beschränkt, d.h. dieser Raum lässt sich mit der Supremumsnorm $\|\cdot\|_{B(X,Y)}$ normieren, allerdings ist der im Allgemeinen nicht abgeschlossen in $C_b(X,Y)$, d.h. nicht vollständig.
Falls $\Omega \subset \mathbb{R}^n$ offen ist, gibt es stetige Abbildungen $f : \Omega \to Y$, welche nicht beschränkt sein müssen. Im allgemeinen lässt sich der Raum $C(\Omega, Y)$ daher nicht normieren. Allerdings kann der Raum $C(\Omega, Y)$ mit einer Metrik versehen werden:

Definition 2.5. $\Omega \subset \mathbb{R}^n$ habe eine **kompakte Ausschöpfung** $(K_m)_{m\in\mathbb{N}}$, falls es kompakte Mengen $K_m \subset \mathbb{R}^n$ gibt mit

(1) $\emptyset \neq K_m \subset K_{m+1} \subset \Omega$ für $m \in \mathbb{N}$ mit $\Omega = \bigcup_{m\in\mathbb{N}} K_m$.

(2) $x \in \Omega \Rightarrow B_\delta(x) \cap \Omega \subset K_m$ für ein $\delta > 0$ und $m \in \mathbb{N}$.

Beispiel 2.6.

(a) Sei $\Omega = [-1,1]$ und $K_m := \{0\} \cup [\frac{1}{m}, 1] \cup [-1, -\frac{1}{m}]$ für $m \in \mathbb{N}$. Dann erfüllt K_m (1) jedoch nicht (2), denn für $0 \in \Omega$ existiert kein $\delta > 0$, sodass $B_\delta(0) \cap \Omega \subset K_m$ für ein $m \in \mathbb{N}$. Ω besitzt jedoch eine kompakte Ausschöpfung $K = \Omega$.

(b) Für jede offene Menge $\Omega \subset \mathbb{R}^n$ existiert eine solche Sequenz $(K_m)_{m\in\mathbb{N}}$, etwa

$$K_m := \{x \in \Omega \mid \operatorname{dist}(x, \mathbb{R}^n \setminus \Omega) \geq \tfrac{1}{m}, |x| \leq m\}.$$

Satz 2.7. *Sei $\Omega \subset \mathbb{R}^n$ offen, $(K_i)_{i\in\mathbb{N}}$ eine zugehörige kompakte Ausschöpfung. Definiere für $f \in C(\Omega, Y)$ auf einem normierten Raum $(Y, \|\cdot\|)$*

$$[f]_i := \sup_{x\in K_i} \|f(x)\| \quad \textit{und} \quad d(f,g) := \sum_{i=1}^{\infty} 2^{-i} \frac{[f-g]_i}{1+[f-g]_i}.$$

*Dann ist $[\cdot]_i$ eine Halbnorm auf $C(\Omega, Y)$ und d eine Metrik, die sogenannte **Frechet-Metrik**. Ist Y ein Banachraum, dann ist $(C(\Omega, Y), d)$ vollständig.*

Beweis. Dass d eine Metrik ist, folgt aus den Übungen. Sei $(f_j)_{j\in\mathbb{N}}$ eine Cauchy-Folge bzgl. d. Dann ist aufgrund der homöomorphen Funktion φ für jede Metrik

$$d_i(f,g) := [f-g]_i \qquad \forall\, f,g \in C(K_i, Y)$$

$(f_j)_{j\in\mathbb{N}}$ eine Cauchy-Folge bzgl. d_i. Wegen der Vollständigkeit von $C(K_i, Y)$ existiert ein $g_i \in C(K_i, Y)$, sodass

$$\lim_{j\to\infty} \left(\sup_{x\in K_i} \|f_j(x) - g_i(x)\| \right) = 0.$$

Wegen $K_i \subset K_{i+1}$ sieht man, dass für $k < i$ gerade $g_k = g_i|_{K_k}$ gilt. Also kann man $g : X \to Y$ durch

$$g(x) := g_i(x) \quad \text{für } x \in K_i$$

definieren und erhält $\lim_{j\to\infty} d_i(f_j, g) = 0$ für jedes $i \in \mathbb{N}$. Dann gibt es zu $\epsilon > 0$ ein $N \in \mathbb{N}$, sodass $\frac{1}{2^N} < \frac{\epsilon}{2}$ und ein $j_0 \in \mathbb{N}$ mit

$$d_i(f_j, g) < \frac{\epsilon}{2} \qquad \forall\, j \geq j_0,\ i \in \{1, \ldots, N\}.$$

Dann ist für $j \geq j_0$

$$d(f_j, g) = \sum_{i=1}^{N} \frac{1}{2^i} \frac{d_i(f_j,g)}{1+d_i(f_j,g)} + \sum_{i=N+1}^{\infty} \frac{1}{2^i} \frac{d_i(f_j,g)}{1+d_i(f_j,g)} < \frac{\epsilon}{2} + \frac{1}{2^N} < \epsilon.$$

Dies impliziert $d(f_j, g) \to 0$. Die Stetigkeit von g folgt aus der Stetigkeit von $g|_{K_i}$, denn für jedes $x \in \Omega$ gibt es eine offene Kugel in einer der Ausschöpfungsmengen K_i. $\square$

Bemerkung: Man kann zeigen, dass es auf $C(\Omega, Y)$ keine Norm gibt, die die gleiche Topologie erzeugt wie die Metrik d. Für offenes $\Omega \subset \mathbb{R}^n$ können nämlich Funktionen aus $C(\Omega, \mathbb{R}^n)$ beliebig bis zum Rand der Menge Ω anwachsen, z.B. $x \mapsto \mathrm{e}^{\frac{1}{x}}$ in $C(0,1)$.

Im Folgenden werden wir die Approximierbarkeit stetiger Funktionen genauer studieren. Das Analogon zur Einheitsbasis in $\mathbb{R}^n$ ist die folgende Zerlegung der Eins:

Satz 2.8. *Sei $(U_i)_{i\in I}$ eine Familie nicht-leerer, offener Mengen $U_i \subset \mathbb{R}^n$ mit*

$$U := \bigcup_{i\in I} U_i.$$

Dann gibt es eine Funktionenfolge $(\eta_k)_{k\in\mathbb{N}}$ mit folgenden Eigenschaften:

(i) $\eta_k : U \to [0,1]$ erfüllt $\eta_k \in C^\infty(U)$.

(ii) Zu jedem $k \in \mathbb{N}$ gibt es ein $i_k \in I$ mit kompaktem Träger $\operatorname{supp} \eta_k \subset U_{i_k}$, d.h. $\eta_k \in C_c^\infty(U_{i_k})$.

(iii) Jedes $x \in U$ besitzt eine offene Umgebung, auf der alle bis auf endliche viele η_k identisch Null sind.

(iv) Für jedes $x \in U$ gilt

$$\sum_{k\in\mathbb{N}} \eta_k(x) = 1.$$

Die Folge $(\eta_k)_{k\in\mathbb{N}}$ heißt eine zur Überdeckung $(U_i)_{i\in I}$ zugehörige ***Partition der Eins****.*

Beweis. Da $U \subset \mathbb{R}^n$ offen ist, lässt sich analog zur kompakten Ausschöpfung aus Beispiel 2.6 für $m \in \mathbb{N}$

$$K_m := \{x \in \mathbb{R}^n \mid \operatorname{dist}(x, \mathbb{R}^n \setminus U) \geq \tfrac{1}{m}, |x| \leq m\}$$

definieren. Dann ist K_m kompakte Teilmenge von U mit

$$U = \bigcup_{m\in\mathbb{N}} K_m \qquad \text{sowie} \qquad K_m \subset \mathring{K}_{m+1}.$$

Setze $K_{-1} = K_0 = \emptyset$. Dann ist $\mathring{K}_{m+1} \setminus K_{m-2}$ für jedes $m \in \mathbb{N}$ eine offene Menge, welche die kompakte Menge $K_m \setminus \mathring{K}_{m-1}$ enthält. Da auch jedes U_i offen ist, gibt es für jedes $x \in K_m \setminus \mathring{K}_{m-1}$ eine abgeschlossene Kugel und ein $r_x > 0, i_x \in I$ mit

$$\overline{B_{r_x}(x)} \subset (\mathring{K}_{m+1} \setminus K_{m-2}) \cap U_{i_x}.$$

Betrachtet man nun die offene Überdeckung von $K_m \setminus \mathring{K}_{m-1}$ mit diesen Kugeln $B_{r_x}(x)$, so liefert eine endliche Teilüberdeckung offene Kugeln $\{B_{r_i}(x_i)\}_{i=1}^{n_m}$ mit

$$K_m \setminus \mathring{K}_{m-1} \subset \bigcup_{i=1}^{n_m} B_{r_i}(x_i) \subset \mathring{K}_{m+1} \setminus K_{m-2}$$

wobei jeweils $\overline{B_{r_i}(x_i)} \subset U_{j_i}$ für ein $j_i \in I$ enthalten ist. Für jedes $m \in \mathbb{N}$ erhält man somit endlich viele offene Kugeln, welche nach Durchnummerierung die abzählbare Folge $(V_k)_{k\in\mathbb{N}}$ mit folgenden Eigenschaften liefert:

Für jedes $k \in \mathbb{N}$ gibt es ein $i \in I$ mit $\overline{V_k} \subset U_i$, d.h. $(V_k)_{k\in\mathbb{N}}$ ist eine Verfeinerung der Überdeckung $(U_i)_{i\in I}$. Die Kugeln überdecken U gemäß

$$\bigcup_{k\in\mathbb{N}} V_k = \bigcup_{m\in\mathbb{N}} (K_m \setminus \mathring{K}_{m-1}) = U.$$

Außerdem trifft jedes K_m nur endlich viele $\overline{V_k}$. Da jedes $x \in U$ in einem $\mathring{K}_m$ liegt, besitzt es eine Umgebung, die nur endlich viele der $\overline{V_k}$ trifft, d.h. $(V_k)_{k\in\mathbb{N}}$ ist lokal endlich.
Nun zur Konstruktion der glatten Funktionen η_k. Die Funktion

$$\psi : \mathbb{R}_+ \to \mathbb{R}_+, \quad x \mapsto \begin{cases} \mathrm{e}^{-1/x} & x > 0, \\ 0 & x \leq 0 \end{cases}$$

ist nach Analysis 1 unendlich oft differenzierbar. Jede abgeschlossene Kugel $\overline{V_k}$ lässt sich als $\overline{B_{r_k}(x_k)}$ darstellen und die Funktion

$$\psi_k : \mathbb{R}^n \to [0,1], \quad x \mapsto \psi(r_k^2 - \|x - x_k\|^2)$$

hat kompakten Träger $\operatorname{supp}(\psi_k) = \overline{V_k}$, d.h. (ii) und (iii) sind erfüllt. Definiere zur Normierung

$$\eta_k := \frac{\psi_k}{\sum_{k\in\mathbb{N}} \psi_k} : U \to [0,1],$$

wobei die Summe stets endlich ist. Damit folgen (i) und (iv). □

Korollar 2.9. *Sei $K \subset \mathbb{R}^n$ kompakt und $(U_i)_{i=1}^m$ eine endliche offene Überdeckung von K. Dann gibt es eine endliche Zerlegung der Eins, d.h. $\eta_i \in C_c^\infty(U_i, [0,1])$ für $i = 1, \dots, m$ mit*

$$\sum_{i=1}^m \eta_i(x) = 1 \qquad \forall\, x \in K.$$

Gibt es eine offene Menge $\Omega \subset \mathbb{R}^n$ mit $K \subset \Omega$, dann existiert eine ***Abschneidefunktion*** *$\eta \in C_c^\infty(\Omega, [0,1])$ mit $\eta_{|K} = 1$.*

Beweis. Siehe Übungen. □

Satz 2.10. *Sei $X \subset \mathbb{R}^n$ kompakt und $(Y, \|\cdot\|)$ ein separabler Banachraum. Dann ist $C(X,Y)$ mit der Norm $\|\cdot\|_\infty$ auch separabel.*

Beweis. Man konstruiere für jedes $\delta > 0$ eine abzählbare Menge $\Sigma_\delta \subset C(X,Y)$ mit folgender Eigenschaft:

$$\forall\, f \in C(X,Y),\ \varepsilon > 0 \ \ \exists\, \delta > 0 \text{ und } f_\sigma \in \Sigma_\delta, \text{ sodass } \|f - f_\delta\|_\infty < \varepsilon. \tag{2.1}$$

Dann ist auch

$$\Sigma := \bigcup_{k\in\mathbb{N}} \Sigma_{\frac{1}{k}}$$

abzählbar und aus (2.1) folgt, dass Σ dicht in $C(X,Y)$ ist. Dies zeigt, dass $C(X,Y)$ separabel ist. Man konstruiere nun Σ_δ:

Da X (prä)kompakt ist, gibt es für jedes $\delta > 0$ eine entsprechende δ-Überdeckung, d.h. es gibt $x_1, \dots, x_{k_\delta} \in X$, sodass

$$X \subset \bigcup_{i=1}^{k_\delta} B_\delta(x_i).$$

Wähle nach Korollar 2.9 eine zugehörige Zerlegung der Eins, d.h.

$$\sum_{i=1}^{k_\delta} \eta_i(x) = 1 \quad \forall\, x \in X \qquad \text{und} \qquad \eta_i = 0 \quad \text{auf } \mathbb{R}^n \setminus B_\delta(x_i), \quad i = 1, \dots, k_\delta.$$

Sei $D \subset Y$ eine abzählbare dichte Teilmenge von Y. Man definiere die abzählbare Menge

$$\Sigma_\delta := \left\{ g : X \to Y \;\middle|\; g = \sum_{i=1}^{k_\delta} \eta_i \cdot \varrho \quad \text{für} \quad \varrho \in D \right\}.$$

Es bleibt (2.1) zu verifizieren. Sei dazu $f \in C(X, Y)$ und $\varepsilon > 0$. Die Kompaktheit von X impliziert, dass f gleichmäßig stetig ist, d.h. es gibt ein $\delta_f > 0$, sodass gilt

$$\|f(x) - f(y)\| < \frac{\varepsilon}{2} \qquad \forall\, x, y \in X \;\text{ mit }\; |x - y| < \delta_f.$$

Wähle δ, sodass $0 < \delta < \delta_f$, und $\varrho_i \in D$ mit

$$\|\varrho_i - f(x_i)\| < \frac{\varepsilon}{2} \qquad \text{für} \quad i = 1, \dots, k_\delta$$

und definiere $g \in \Sigma_\delta$ mit

$$g(x) := \sum_{i=1}^{k_\delta} \varrho_i \, \eta_i(x).$$

Wegen $\eta_i(x) = 0$ für $|x_i - x| \geq \delta_f$ erhält man für alle $x \in X$

$$\|f(x) - g(x)\| \leq \sum_{i=1}^{k_\delta} \|f(x) - f(x_i)\| \, \eta_i(x) + \sum_{i=1}^{k_\delta} \|f(x_i) - \varrho_i\| \, \eta_i(x) < \varepsilon.$$

□

Während obiger Satz eine qualitative Aussage macht, lassen sich im Spezialfall für $C_\mathbb{K}(K)$ auf einem kompakten Raum K mit nachfolgendem Satz abzählbar dichte Teilmengen testen.

Definition 2.11. Sei K ein kompakter Raum. $\mathcal{A} \subset C(K)$ heißt eine **Unteralgebra** von $C(K)$, falls $\mathcal{A}$ ein linearer Unterraum ist und mit $f, g \in \mathcal{A}$ auch $f \cdot g \in \mathcal{A}$ gilt. Man sagt, dass die Unteralgebra $\mathcal{A}$ die Punkte von K trennt, falls für alle $x, y \in K$ mit $x \neq y$ ein $f \in \mathcal{A}$ mit $f(x) \neq f(y)$ existiert.

Beispiel 2.12.

(a) Betrachtet man den Vektorraum $\mathbb{R}^2$ mit der komponentenweisen Multiplikation ist eine Algebra, welche nur die Unteralgebren

$$\mathbb{R}^2, \{(x, x) \mid x \in \mathbb{R}\}, \{(x, 0) \mid x \in \mathbb{R}\}, \{(0, x) \mid x \in \mathbb{R}\} \quad \text{und} \quad \{(0, 0)\}$$

besitzt.

(b) Sei $K \subset \mathbb{R}^n$ kompakt und $\mathcal{A}$ die Menge aller reellwertigen Polynome, d.h.

$$\mathcal{A} = \left\{ p : K \to \mathbb{R} \;\middle|\; p(x) = \sum_{\alpha \in \mathbb{N}_0^n,\, |\alpha| \le n} a_\alpha x^\alpha \text{ für ein } n \in \mathbb{N}_0 \text{ und } a_\alpha \in \mathbb{R} \right\},$$

wobei Multiindex-Schreibweise

$$\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n, \quad |\alpha| = \sum_{k=1}^n \alpha_k, \quad x^\alpha = x_1^{\alpha_1} \cdot \ldots \cdot x_n^{\alpha_n}.$$

benutzt wurde. Dann ist $\mathcal{A}$ eine Unteralgebra von $C(K)$, welche die Punkte in K trennt.

(c) $C(K)$ ist selbst eine Punkte trennende Unteralgebra von $C(K)$.

Korollar 2.13. *Sei K ein kompakter Raum. Dann separiert $C_\mathbb{K}(K)$ die Punkte von K, d.h.*

$$\forall\, x, y \in K \text{ mit } x \neq y \;\; \exists\, f \in C_\mathbb{K}(K) \text{ mit } f(x) \neq f(y).$$

Beweis. Seien $x, y \in K$, $x \neq y$. Da K hausdorff'sch ist, findet man zwei offene Umgebungen U_x, U_y von x bzw. y mit $U_x \cap U_y = \emptyset$. Nach Lemma 1.15 existieren dann abgeschlossene $A, B \subset K$ mit $A \subset U_x$, $B \subset U_y$ und $x \in A$, $y \in B$. Insbesondere ist $A \cap B = \emptyset$. Nach Satz 1.18 existiert ein $f \in C_\mathbb{K}(K)$ mit $f(x) = 0$ und $f(y) = 1$. □

Um den nachstehenden allgemeinen Satz von Stone-Weierstrass zu beweisen, benötigen wir folgenden Spezialfall.

Lemma 2.14. *Die Betragsfunktion $|\cdot| : [-1,1] \to \mathbb{R}$ lässt sich gleichmäßig durch Polynome approximieren. Genauer gilt: Zu jedem $\varepsilon > 0$ gibt es ein Polynom $p : [-1,1] \to \mathbb{R}$ mit $p(0) = 0$ und*

$$\max_{x \in [-1,1]} \big| |x| - p(x) \big| < \varepsilon.$$

Beweis. Nach Analysis 1 konvergiert die Taylorreihe der Funktion $f : [0,1] \to \mathbb{R}$ mit $f(t) = (1-t)^{1/2}$ gleichmäßig und stellt f auf diesem Intervall dar:

$$f(t) = \sum_{k \in \mathbb{N}_0} (-1)^k \binom{1/2}{k} t^k.$$

Damit gibt es für jedes $\varepsilon > 0$ ein Taylorpolynom $T : [0,1] \to \mathbb{R}$ mit

$$\max_{x \in [0,1]} \big| f(x) - T(x) \big| < \frac{\varepsilon}{2}.$$

Substituiere $t = 1 - x^2$, so folgt die Aussage für $p(x) := T(1-x^2) - T(1)$ mit $p(0) = 0$ und

$$\max_{x \in [-1,1]} \big| |x| - p(x) \big| \le \max_{x \in [-1,1]} \big| f(1-x^2) - T(1-x^2) \big| + |T(1)| < \varepsilon.$$

□

Satz 2.15 (Stone-Weierstrass). *Sei K ein kompakter Raum und $\mathcal{A}$ eine Unteralgebra von $C(K)$, die die Punkte von K trennt. Dann gilt entweder*

$$\overline{\mathcal{A}} = C(K)$$

oder es existiert ein $x_0 \in K$, sodass

$$\overline{\mathcal{A}} = \{f \in C(K) \mid f(x_0) = 0\}.$$

Beweis. Aus den Übungen folgt, dass mit $\mathcal{A}$ auch $\overline{\mathcal{A}}$ eine Unteralgebra ist und mit $f, g \in \mathcal{A}$ auch $|f|, \max\{f, g\}, \min\{f, g\} \in \overline{\mathcal{A}}$ gilt. Sei nun $f \in C(K)$. Wir behaupten, dass gilt:

$$\forall\, x, y \in K \quad \exists\, g_{xy} \in \mathcal{A} \quad \text{mit} \quad g_{xy}(x) = f(x),\ g_{xy}(y) = f(y) \qquad \Rightarrow \qquad f \in \overline{\mathcal{A}}. \tag{2.2}$$

Zum Beweis sei $\varepsilon > 0$ beliebig und $x, y \in K, x \neq y$ und definiere zu obigem $g_{xy} \in \mathcal{A}$

$$U_-^{xy} := \{z \in K \mid f(z) > g_{xy}(z) - \varepsilon\} \quad \text{und} \quad U_+^{xy} := \{z \in K \mid f(z) < g_{xy}(z) + \varepsilon\}.$$

Dann ist $x, y \in U_\pm^{xy}$ und für festes $y \in K$ bildet $\bigcup_{x \in K \setminus \{y\}} U_+^{xy}$ eine offene Überdeckung der kompakten Menge K. Folglich gibt es endlich viele $x_1, \ldots, x_m \in K, x_i \neq y$. Setze

$$g_y := \max\{g_{x_1 y}, \ldots, g_{x_m y}\} \in \overline{\mathcal{A}},$$

dann gilt $f < g_y + \varepsilon$ in K und $f > g_y - \varepsilon$ auf dem Schnitt $V_y := \bigcap_{i=1}^m U_-^{x_i y}$.
Auf dieselbe Weise ist $\bigcup_{y \in K} V_y$ eine offene Überdeckung von K und es gibt eine endliche Teilüberdeckung zu $y_1, \ldots, y_k \in K$. Definiere

$$g := \min\{g_{y_1}, \ldots, g_{y_k}\} \in \overline{\mathcal{A}},$$

so ist $\|f - g\|_\infty < \varepsilon$, d.h. auch $f \in \overline{\mathcal{A}}$, da $\overline{\mathcal{A}}$ abgeschlossen ist.
Seien nun $x, y \in K$ mit $x \neq y$ und setze $B := \{(f(x), f(y)) \mid f \in \mathcal{A}\} \subset \mathbb{R}^2$. Dann ist B auch eine Unteralgebra von $\mathbb{R}^2$ und B stimmt nach Beispiel 2.12 mit einer der Unteralgebren

$$\mathbb{R}^2, \{(x, x) \mid x \in \mathbb{R}\}, \{(x, 0) \mid x \in \mathbb{R}\}, \{(0, x) \mid x \in \mathbb{R}\} \quad \text{oder} \quad \{(0, 0)\}$$

überein. Da $\mathcal{A}$ Punkte in K trennt, kann B nicht mit $\{(x, x) \mid x \in \mathbb{R}\}$ oder $\{(0, 0)\}$ übereinstimmen. Wäre $B = \mathbb{R}^2$, so ist nach eben bewiesener Behauptung (2.2) und Satz 1.18 gerade $\overline{\mathcal{A}} = C(K)$. In den übrigen beiden Fällen gibt es aber ein $x_0 \in K$ mit $f(x_0) = 0$ für alle $f \in \mathcal{A}$ und es folgt wiederum

$$\overline{\mathcal{A}} = \{f \in C(K) \mid f(x_0) = 0\}.$$

□

Beispiel 2.16. (a) Da die Menge aller Polynome aus Beispiel 2.12 (b) konstante Funktionen enthält, kann $\overline{\mathcal{A}} = \{f \in C(K) \mid f(x_0) = 0\}$ nicht gelten. Das heißt jede stetige Funktion in $C(K)$ lässt sich durch Polynome approximieren.

(b) Analog lässt sich zeigen, dass die Menge der trigonometrischen Polynome

$$\mathcal{A} = \left\{ p : \mathbb{R} \to \mathbb{R} \;\middle|\; p(x) = \sum_{k=0}^n a_k \sin(kx) + b_k \cos(kx) \ \text{ für ein } n \in \mathbb{N}_0,\ a_k, b_k \in \mathbb{R} \right\}$$

eine dichte Unteralgebra von $C_{2\pi}(\mathbb{R}) := \{f \in C(\mathbb{R}) \mid f(0) = f(2\pi)\}$ bzgl. $\|\cdot\|_\infty$ ist.

Für Funktionen auf offenen Teilmengen $\Omega \subset \mathbb{R}^n$ definiert man den Begriff Differenzierbarkeit analog zur Analysis 2:
Sei $f : \Omega \to \mathbb{K}$ und $e_1, \dots, e_n$ die Standardbasis des $\mathbb{R}^n$. Existiert der Limes

$$\partial_i f(x) := \lim_{h \to 0} \frac{f(x + h\, e_i) - f(x)}{h},$$

so heißt $\partial_i f(x)$ die **partielle Ableitung** von f nach x in Richtung $e_i = (0, \dots, 1, \dots, 0)$. $f : \Omega \to \mathbb{K}$ heißt **partiell differenzierbar**, falls alle partiellen Ableitungen in allen Punkten existieren und **stetig differenzierbar**, falls diese zusätzlich stetige Abbildungen von Ω nach $\mathbb{K}$ definieren. Der **Gradient** von f in x ist definiert durch:

$$\nabla f(x) := (\partial_1 f(x), \dots, \partial_n f(x)) .$$

Für $m \geq 2$ heißt $f : \Omega \to \mathbb{K}$ m-mal stetig differenzierbar, falls $\partial_{i_1}\partial_{i_2} \dots \partial_{i_k} f$ für alle Indizes $i_1, \dots, i_k \in \{1, \dots, n\}$ und $k \leq m$ existiert und eine stetige Abbildung von Ω nach $\mathbb{K}$ ist (diese hängen nicht von der Reihenfolge der einzelnen partiellen Ableitungen ab). Verwendet man Multiindex-Schreibweise, so ist für $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$

$$\partial^\alpha f := \partial_1^{\alpha_1} \partial_2^{\alpha_2} \cdots \partial_n^{\alpha_n} f.$$

Definition 2.17. Sei $\Omega \subset \mathbb{R}^n$ offen. Definiere

$$C^m_{\mathbb{K}}(\Omega) := \{ f : \Omega \to \mathbb{K} \mid f \text{ ist } m\text{-mal stetig differenzierbar in } \Omega \}$$

und falls Ω zusätzlich beschränkt ist

$$C^m_{\mathbb{K}}(\overline{\Omega}) := \left\{ f : \Omega \to \mathbb{K} \;\middle|\; \begin{array}{l} f \text{ ist } m\text{-mal stetig differenzierbar in } \Omega \text{ und} \\ \partial^\alpha f \text{ ist auf } \overline{\Omega} \text{ stetig fortsetzbar } \;\forall\, \alpha \in \mathbb{N}_0^n,\; |\alpha| \leq m \end{array} \right\}.$$

mit Norm

$$\|f\|_{C^m_{\mathbb{K}}(\overline{\Omega})} := \sum_{\alpha \in \mathbb{N}_0^n, |\alpha| \leq m} \|\partial^\alpha f\|_{C_{\mathbb{K}}(\overline{\Omega})}.$$

Lemma 2.18. *Sei $\Omega \subset \mathbb{R}^n$ offen und beschränkt. Dann ist $C^m_{\mathbb{K}}(\overline{\Omega})$ ist ein Banachraum.*

Beweis. Wir zeigen den Fall $m = 1$. Der allgemeine Fall folgt dann per Induktion. Sei also $(f_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $C^1_{\mathbb{K}}(\overline{\Omega})$, dann ist sind die Folgen $(f_n)_{n\in\mathbb{N}}$ und $(\partial_i f_n)_{n\in\mathbb{N}}$ für jedes $i = 1, \dots, n$ ebenfalls Cauchy-Folgen in $C_{\mathbb{K}}(\overline{\Omega})$ und es gibt stetige Grenzfunktionen f sowie g_i für jedes $i = 1, \dots, n$ mit

$$\|f_n - f\|_\infty \to 0, \quad \|\partial_i f_n - g_i\|_\infty \to 0 \qquad (\text{für } n \to \infty).$$

Sei $x \in \Omega$, dann gibt es ein $r > 0$ mit $B_r(x) \subset \Omega$. Für $y \in B_r(x)$ folgt nach dem Mittelwertsatz für $x_t = (1-t)x + ty$

$$f_n(x) - f_n(y) = \int_0^1 \nabla f_n(x_t)(y - x)\, \mathrm{d}t.$$

Dies impliziert

$$\begin{aligned} |f_n(x) - f_n(y) - (y - x)\nabla f_n(x)| &\leq \int_0^1 |\nabla f_n(x_t) - \nabla f_n(x)|\, |y - x|\, \mathrm{d}t \\ &\leq |y - x| \left(2\|\nabla f_n - g\|_\infty + \sup_{t\in[0,1]} |g(x_t) - g(x)| \right). \end{aligned}$$

Für $n \to \infty$ erhält man dann

$$|f(x) - f(y) - (y-x)g(x)| \leq \sup_{t \in [0,1]} |g(x_t) - g(x)||x-y|$$

und die Stetigkeit von g liefert die Differenzierbarkeit von f mit $\nabla f = g$ und $f_n \to f$ bezüglich $\|\cdot\|_{C^1_{\mathbb{K}}(\overline{\Omega})}$ für $n \to \infty$. □

Definition 2.19. Man definiere $C^0_{\mathbb{K}}(\Omega) := C_{\mathbb{K}}(\Omega)$ und den Raum der glatten Funktionen

$$C^\infty_{\mathbb{K}}(\Omega) := \bigcap_{m \in \mathbb{N}} C^m_{\mathbb{K}}(\Omega)$$

sowie

$$C^\infty_{\mathbb{K}}(\overline{\Omega}) := \bigcap_{m \in \mathbb{N}} C^m_{\mathbb{K}}(\overline{\Omega}).$$

Dann lassen sich diese Räume mit der Metrik aus Satz 2.7 vervollständigen. Analog definiert man die Funktionen mit kompaktem Träger

$$C^m_{c\mathbb{K}}(\Omega) := \{f \in C^m_{\mathbb{K}}(\Omega) \mid \operatorname{supp}(f) \text{ ist kompakt}\}.$$

Weitere wichtige stetige Funktionen sind die Hölder-stetigen Funktionen.

Definition 2.20. Sei $\alpha \in (0,1]$, $\Omega \subset \mathbb{R}^n$ offen und beschränkt. $f : \Omega \to \mathbb{R}$ heißt **Hölder-stetig** mit Exponent α, falls gilt:

$$[f]_{\alpha,\Omega} := \sup_{x,y \in \Omega, x \neq y} \frac{|f(x) - f(y)|}{|x-y|^\alpha} < \infty.$$

Im Fall $\alpha = 1$ heißt f **Lipschitz-stetig**.
Für $k \in \mathbb{N}_0$ sind die Räume der Hölder- bzw. Lipschitz-stetigen Funktionen definiert durch

$$C^{k,\alpha}(\overline{\Omega}) := \{f \in C^k(\overline{\Omega}) \mid \max_{\gamma \in \mathbb{N}_0^n, |\gamma| = k} [\partial^\gamma f]_{\alpha,\Omega} < \infty\}.$$

Lemma 2.21. *Sei $\alpha \in (0,1], k \in \mathbb{N}_0$, $\Omega \subset \mathbb{R}^n$ offen und beschränkt. Dann ist $C^{k,\alpha}(\overline{\Omega})$ mit der Norm*

$$\|f\|_{C^{k,\alpha}(\overline{\Omega})} := \|f\|_{C^k(\overline{\Omega})} + \max_{\gamma \in \mathbb{N}_0^n, |\gamma| = k} [\partial^\gamma f]_{\alpha,\Omega}$$

ein Banach-Raum. Insbesondere gilt für $0 < \alpha \leq \beta \leq 1$

$$C^{k,\beta}(\overline{\Omega}) \subsetneq C^{k,\alpha}(\overline{\Omega}) \subsetneq C^k(\overline{\Omega}).$$

Beweis. Siehe Übungen. □

2.2 Lebesgue-Räume

Im Folgenden wird die Maß- und Integrationstheorie, welche zur Definition des Lebesgue-Integrals und der sogenannten Lebesgue-Räume nötig ist, nochmals analog zur Höheren Analysis entwickelt und teilweise erweitert. Die Beweise wurden größtenteils in der Vorlesung Höhere Analysis bereits vorgeführt und sind gekennzeichnet.
Im zweiten Abschnitt werden grundlegende Eigenschaften der L^p-Räume bewiesen, die insbesondere als Grundlage für die Lösungstheorie partieller Differentialgleichungen unabdingbar sind.

2.2.1 Maßtheorie

Um einen sinnvollen Maßbegriff zu definieren, benötigen wir eine gewisse Struktur.

Definition 2.22 (σ-Algebra, Ana 3)**.** Sei X eine Menge. $\Sigma \subset \operatorname{Pot}(X)$ heißt **σ-Algebra**, falls

(i) $X \in \Sigma$

(ii) $A \in \Sigma \;\Rightarrow\; X \setminus A \in \Sigma$

(iii) Sei $(A_k)_{k\in\mathbb{N}}$ eine Folge von Mengen in Σ, so ist auch

$$\bigcup_{k\in\mathbb{N}} A_k \in \Sigma.$$

Beispiel 2.23 (Ana 3)**.** (a) $\Sigma = \{\emptyset, X\}$ ist die kleinste und $\Sigma = \operatorname{Pot}(X)$ die größte σ-Algebra.

(b) Ist $(A_k)_{k\in\mathbb{N}}$ eine Folge von Mengen in einer σ - Algebra, dann ist auch $\emptyset \in \Sigma$ sowie

$$\bigcap_{k\in\mathbb{N}} A_k \in \Sigma.$$

Definition 2.24 (Ana 3)**.** Sei X eine Menge und $\mathcal{F} \subset \operatorname{Pot}(X)$ eine Familie von Teilmengen von X. Man definiere die durch $\mathcal{F}$ erzeugte σ-Algebra durch

$$\Sigma(\mathcal{F}) := \bigcap \{\mathcal{A} \mid \mathcal{F} \subset \mathcal{A} \text{ und } \mathcal{A} \text{ ist eine } \sigma\text{ - Algebra}\}.$$

$\Sigma(\mathcal{F})$ ist die kleinste σ-Algebra, die $\mathcal{F}$ enthält.
Falls Σ eine σ-Algebra ist, so induziert Σ auf einer beliebigen Teilmenge $A \subset X$ die **relative σ-Algebra**

$$\Sigma_A := \{A \cap B \mid B \in \Sigma\}.$$

Bemerkung: Auf $\mathbb{R}^n$ definiert man die **Borelsche σ-Algebra $\mathcal{B}(\mathbb{R}^n)$** als die σ-Algebra, die aus der Familie $\mathcal{F}$ aller offenen Mengen erzeugt wird.

(i) Man kann zeigen, dass $\mathcal{B}(\mathbb{R}^n) \neq \operatorname{Pot}(\mathbb{R}^n)$. Es existieren nämlich Teilmengen des $\mathbb{R}^n$, die nicht borelsch sind (nicht-triviale Konstruktion).

(ii) Die borelsche σ - Algebra kann ebenso auf topologischen Räumen definiert werden.

Definition 2.25 (Maß, Maßraum, Ana 3)**.** Sei X eine Menge und Σ eine σ - Algebra auf X. Ein **Maß** auf Σ ist eine Abbildung $\mu : \Sigma \to [0, \infty]$ mit folgenden Eigenschaften:

(i) $\mu(\emptyset) = 0$.

(ii) Ist $(A_n)_{n\in\mathbb{N}}$ eine Folge in Σ und $A_k \cap A_l = \emptyset$ für $\forall\, k \neq l$, dann ist

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n) \qquad (\sigma\text{ - Additivität})$$

Das Tripel (X, Σ, μ) nennt man dann **Maßraum** und die Elemente von Σ heißen **messbare Mengen**. $A \in \Sigma$ ist eine **Nullmenge**, wenn $\mu(A) = 0$ gilt.

Lemma 2.26 (Ana 3)**.** *Für jeden Maßraum (X, Σ, μ) und $(A_k)_{k\in\mathbb{N}} \subset \Sigma$ gilt:*

(i) Sind $A, B \in \Sigma$ mit $A \subset B$, so ist $\mu(A) \leq \mu(B)$ *(Monotonie)*

(ii) Ist $(A_n)_{n\in\mathbb{N}}$ monton steigend, d.h. $A_k \subset A_{k+1}$ für alle $k \in \mathbb{N}$ gilt, so folgt

$$\lim_{n\to\infty} \mu(A_n) = \mu\left(\bigcup_{n\in\mathbb{N}} A_n\right).$$

(iii) Fällt $(A_n)_{n\in\mathbb{N}}$ monoton, d.h. $A_k \supset A_{k+1}$ für alle $k \in \mathbb{N}$, so gilt für $\mu(A_1) < \infty$

$$\lim_{n\to\infty} \mu(A_n) = \mu\left(\bigcap_{n\in\mathbb{N}} A_n\right).$$

Beweis. Siehe z.B. J. Elstrodt: Maß- und Integrationstheorie, Satz II.1.7 und II.1.10. □

Definition 2.27 (Ana 3)**.** Sei (X, Σ, μ) ein Maßraum. Ein Maß $\mu : \Sigma \to \mathbb{R}_+$ heißt **σ-finit**, falls es eine abzählbare Überdeckung $(A_k)_{k\in\mathbb{N}} \subset \Sigma$ von X gibt, d.h. $X = \bigcup_{k\in\mathbb{N}} A_k$, mit $\mu(A_k) < \infty$ für jedes $k \in \mathbb{N}$. Im Fall $\mu(X) < \infty$ heißt μ **endlich**, für $\mu(X) = 1$ **Wahrscheinlichkeitsmaß**.

Beispiel 2.28 (Ana 3)**.** Sei X eine beliebige Menge und $\Sigma = \text{Pot}\,(X)$.

(a) Für $A \in \Sigma$ setze

$$\mu(A) := \begin{cases} \#A & \text{falls } A \text{ endlich ist,} \\ \infty & \text{sonst.} \end{cases}$$

Dann ist μ das sogenannte **Zählmaß**. μ ist endlich, wenn X endlich ist und σ-finit, wenn X abzählbar ist.

(b) Für festes $x \in X$ und $A \in \Sigma$ ist

$$\mu(A) := \delta_x(A) := \begin{cases} 1 & \text{falls } x \in A, \\ 0 & \text{sonst} \end{cases}$$

das sogenannte **Dirac - Maß**. Offensichtlich ist δ_x ein Wahrscheinlichkeitsmaß.

Satz 2.29 (Lebesgue-Maß, Ana 3)**.** *Es existiert ein eindeutig bestimmtes Maß*

$$\lambda^n : \mathcal{B}(\mathbb{R}^n) \to [0,\infty] \quad \textit{mit} \quad \lambda^n\left([a_1,b_1) \times \ldots \times [a_n,b_n)\right) := \prod_{i=1}^{n} |b_i - a_i|$$

für beliebige Quader[1] *in* $\mathbb{R}^n$ *gilt.* λ^n *heißt n-dimensionales* ***Lebesgue-Maß****. Das Lebesgue-Maß kann man zu einem Maß auf der* σ*-Algebra aller Lebesgue-messbaren Mengen* $\mathcal{M}(\lambda^n)$ *erweitern. Es gilt außerdem*

$$\mathcal{B}(\mathbb{R}^n) \subsetneq \mathcal{M}(\lambda^n) \subsetneq \mathrm{Pot}\,(\mathbb{R}^n).$$

Beweis. Siehe z.B. J. Elstrodt: Maß- und Integrationstheorie, Korollar II.6.5 und Korollar II.8.6 sowie der Satz III.3.1 von Vitaly. □

Bemerkung: Manchmal ist es nützlich, λ^n als Maß auf $\mathcal{M}(\lambda^n)$ zu betrachten (statt auf $\mathcal{B}(\mathbb{R}^n)$). Der Grund ist dieser, dass $(\mathbb{R}^n, \mathcal{M}(\lambda^n), \lambda^n)$ ein vollständiger Maßraum ist, d.h.

$$A \subset B \subset \mathbb{R}^n,\ B \in \mathcal{M}(\lambda^n) \ \text{ mit } \ \lambda^n(B) = 0 \quad \Rightarrow \quad A \in \mathcal{M}(\lambda^n),\ \lambda^n(A) = 0.$$

Dasselbe gilt nicht für $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \lambda^n)$, denn es gibt Teilmengen von Lebesgue-Nullmengen, die nicht borelsch sind.

Lemma 2.30 (Ana 3)**.** *Es gelten folgende Eigenschaften des Lebesgue-Maßes*

(i) λ^n *ist* ***von außen regulär****, d.h. für jedes* $A \in \mathcal{M}(\lambda^n)$ *gilt*

$$\lambda^n(A) = \inf\left\{\lambda^n(\Omega) \mid A \subset \Omega,\ \Omega \textit{ offen}\right\}.$$

(ii) λ^n *ist* ***von innen regulär****, d.h. für jedes* $A \in \mathcal{M}(\lambda^n)$ *gilt*

$$\lambda^n(A) = \sup\left\{\lambda^n(K) \mid K \subset A,\ K \textit{ kompakt}\right\}.$$

(iii) λ^n *ist* σ*-finit, d.h. es existiert eine Folge* $(A_i)_{i\in\mathbb{N}}$ *in* $\mathcal{B}(\mathbb{R}^n)$*, sodass*

$$\lambda^n(A_i) < \infty \quad \forall\, i \in \mathbb{N} \quad \textit{und} \quad \mathbb{R}^n = \bigcup_{i\in\mathbb{N}} A_i.$$

Beweis. Siehe z.B. Jürgen Elstrodt: Maß- und Integrationstheorie, Korollar II.7.2.
Für (iii) betrachte eine Überdeckung des $\mathbb{R}^n$ mit Quadern. □

Definition 2.31 (Ana 3)**.** Sei (X, Σ, μ) ein Maßraum. Eine Aussage gilt μ**-fast überall** (μ-f.ü.), falls es für die Menge

$$A = \{x \in X \mid x \text{ erfüllt die Aussage nicht}\}$$

ein $B \in \Sigma$ gibt mit $A \subset B$ und $\mu(B) = 0$. Die Aussage gilt dann für μ**-fast alle** $x \in X$.

Da im Folgenden meist $\mathbb{K}$-wertige Funktionen untersucht werden, beschränken wir uns auf den Messbarkeitsbegriff auf solchen Messräumen.

[1] Anstatt halb-offene Intervalle können auch abgeschlossene oder offene Quader verwendet werden.

Definition 2.32 (Ana 3)**.** Sei X eine Menge, Σ eine σ-Algebra auf X. Man sagt, dass $f : X \to \mathbb{R}$ **messbar** ist (bzgl. Σ), falls für jedes $t \in \mathbb{R}$ gilt:

$$f^{-1}((t, \infty)) = \{x \in X \mid f(x) > t\} \in \Sigma.$$

Eine Funktion $f : X \to \mathbb{C}$ ist messbar, falls $\mathrm{Re}(f)$ und $\mathrm{Im}(f)$ messbar sind. Eine vektorwertige Funktion ist genau dann messbar wenn jede Komponente messbar ist.

Lemma 2.33 (Ana 3)**.** *Seien X eine Menge, Σ eine σ-Algebra auf X und $f, g : X \to \mathbb{K}$ messbar. Dann gilt*

(i) $f + \lambda g$ *für* $\lambda \in \mathbb{K}, fg, |f|, \min\{f, g\}, \max\{f, g\}$ *sind messbar.*

(ii) Falls $\phi : \mathbb{K} \to \mathbb{K}$ messbar ist, so ist $\phi \circ f : X \to \mathbb{K}$ messbar.

(iii) Stetige Funktionen sind messbar.

(iv) Falls $f_i : X \to \mathbb{K}, i \in \mathbb{N}$, eine Folge messbarer Funktionen bildet, sodass $(f_i(x))_{i\in\mathbb{N}}$ einen Limes $f(x)$ besitzt, so ist auch f messbar.

Beweis. Siehe z.B. Jürgen Elstrodt: Maß- und Integrationstheorie, Satz III.4.3 und folgende Korollare. □

Definition 2.34 (Einfache Funktionen, Ana 3)**.** Sei (X, Σ, μ) ein Maßraum. Eine messbare Funktion $\varphi : X \to \mathbb{R}$ heißt **einfach** oder **Treppenfunktion**, falls es endlich viele $\lambda_i \in \mathbb{R}$ und messbare Menge A_i gibt, sodass

$$\varphi = \sum_{i=1}^{N} \lambda_i \, \chi_{A_i},$$

wobei χ_{A_i} die charakteristische Funktion bezeichnet. Sei $S(X, \mu)$ der Vektorraum der einfachen Funktionen. Das Integral für eine einfache, nicht negative Funktion ist durch

$$\int_A \varphi \, \mathrm{d}\mu := \sum_{i=1}^{N} \lambda_i \, \mu(A_i \cap A)$$

für $A \in \Sigma$ definiert, wobei hier $0 \cdot \infty = 0$ vereinbart wird.

Offensichtlich ist das definierte Integral monoton und linear. Die entsprechende Erweiterung auf allgemeine messbare Funktionen liefert

Definition 2.35 (Integral, Ana 3)**.** Das Integral einer messbaren, nicht negativen Funktion ist definiert als das Supremum der Integrale von einfachen Funktionen, die durch f beschränkt werden, d.h. ist $f : X \to \mathbb{R}^+$ messbar und $A \in \Sigma$, so ist

$$\int_A f \, \mathrm{d}\mu := \sup \left\{ \int_A \varphi \, \mathrm{d}\mu \;\middle|\; \varphi \in S(X, \mu), 0 \le \varphi \le f \right\}.$$

f heißt **integrierbar über** A (kurz $f \in \mathcal{L}^1(A, \mu)$), falls

$$\int_A f \, \mathrm{d}\mu < \infty,$$

und **integrierbar**, falls f über ganz X integrierbar ist.
Für messbare Funktionen $f : X \to \mathbb{R}$ setze man

$$f_+ := \max\{0, f\}, \quad f_- := \max\{0, -f\} \quad \text{und} \quad f = f_+ - f_-.$$

Falls f_+, f_- über $A \in \Sigma$ integrierbar sind, so sagt man, dass f integrierbar ist und es ist

$$\int_A f \, \mathrm{d}\mu := \int_A f_+ \, \mathrm{d}\mu - \int_A f_- \, \mathrm{d}\mu.$$

Eine messbare Funktion $f : X \to \mathbb{C}$ ist genau dann über $A \in \Sigma$ integrierbar, falls $\mathrm{Re}(f)$ und $\mathrm{Im}(f)$ integrierbar sind. Es gilt dann

$$\int_A f \, \mathrm{d}\mu := \int_A (\mathrm{Re}\, f)_+ \, \mathrm{d}\mu - \int_A (\mathrm{Re}\, f)_- \, \mathrm{d}\mu + \mathrm{i} \int_A (\mathrm{Im}\, f)_+ \, \mathrm{d}\mu - \mathrm{i} \int_A (\mathrm{Im}\, f)_- \, \mathrm{d}\mu.$$

Aus dieser Definition übertragen sich die elementaren Eigenschaften des Integrals für einfache Funktionen wie beispielsweise Linearität oder Monotonie. Außerdem verifiziert man leicht die Dreiecksungleichung sowie

$$\left| \int_A f \, \mathrm{d}\mu \right| \leq \int_A |f| \, \mathrm{d}\mu \qquad \forall\, f \in \mathcal{L}^1(A, \mu).$$

Satz 2.36 (Monotone Konvergenz / Beppo Levi, Ana 3)**.** *Sei (X, Σ, μ) ein Maßraum und $(f_n)_{n\in\mathbb{N}}$ eine Folge integrierbarer Funktionen $f_n : X \to \mathbb{R}_+$, sodass $f_{n+1}(x) \geq f_n(x)$ für μ-fast alle $x \in X$. Der Limes*

$$f(x) := \lim_{n\to\infty} f_n(x)$$

existiert dann für μ-fast alle $x \in X$ und f ist messbar. Außerdem gilt für jedes $A \in \Sigma$

$$\lim_{n\to\infty} \int_A f_n \, \mathrm{d}\mu = \int_A f \, \mathrm{d}\mu.$$

Beweis. Siehe z.B. Jürgen Elstrodt: Maß- und Integrationstheorie, Satz IV.2.7. □

Es ist einfach zu zeigen, dass die Aussage ohne die Monotonie nicht gilt (punktweise Konvergenz impliziert nicht, dass Limes und Integral vertauscht werden können).
Für eine nicht notwendigerweise monotone Funktionenfolge lässt sich jedoch zeigen, dass das Integral des Limes durch den Limes inferior der Integrale beschränkt werden kann.

Lemma 2.37 (Fatou, Ana 3)**.** *Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von nicht-negativen, integrierbaren Funktionen auf dem Maßraum (X, Σ, μ). Sei*

$$f(x) := \liminf_{n\to\infty} f_n(x).$$

Dann gilt für beliebiges $A \in \Sigma$

$$\liminf_{n\to\infty} \int_A f_n \, \mathrm{d}\mu \geq \int_A f \, \mathrm{d}\mu.$$

Beweis. Siehe z.B. Jürgen Elstrodt: Maß- und Integrationstheorie, Lemma IV.5.1. □

Eine wichtige Anwendung des Lemmas von Fatou ist folgender

Satz 2.38 (Dominierte Konvergenz, Ana 3)**.** *Sei (X, Σ, μ) ein Maßraum und $(f_n)_{n\in\mathbb{N}}$ eine Folge messbarer Funktionen $f_n : X \to \mathbb{K}$, sodass $f_n(x) \to f(x)$ für $n \to \infty$ für μ-fast alle $x \in X$. Existiert eine nicht-negative, integrierbare Funktion $g : X \to \mathbb{R}_+$ mit $|f_n(x)| \leq g(x)$ für μ-fast alle $x \in X$. Dann ist $|f(x)| \leq g(x)$ für μ-fast alle $x \in X$ und für jedes $A \in \Sigma$ gilt $f_n, f \in \mathcal{L}^1(A, \mu)$ mit*

$$\lim_{n\to\infty} \int_A f_n \, \mathrm{d}\mu = \int_A f \, \mathrm{d}\mu \qquad \textit{sowie} \qquad \lim_{n\to\infty} \int_A |f_n - f| \, \mathrm{d}\mu = 0.$$

Beweis. Siehe z.B. Jürgen Elstrodt: Maß- und Integrationstheorie, Satz IV.5.2. □

Definition 2.39 (Ana 3)**.** Seien (X_1, Σ_1, μ_1) und (X_2, Σ_2, μ_2) zwei Maßräume. Auf der Menge $X_1 \times X_2$ definiert man die **Produkt-σ-Algebra** $\Sigma_1 \times \Sigma_2$ als diejenige σ-Algebra, die aus Mengen der Form $A \times B$ mit $A \in \Sigma_1$ und $B \in \Sigma_2$ erzeugt wird.
Auf dieser lässt sich ein Produkt-Maß $\mu_1 \times \mu_2$ definieren mit

$$(\mu_1 \times \mu_2)(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2) \qquad \forall\, A_1 \in \Sigma_1, A_2 \in \Sigma_2.$$

Bemerkung: Es ist $\mathcal{B}(\mathbb{R}^{n+m}) = \mathcal{B}(\mathbb{R}^n) \times \mathcal{B}(\mathbb{R}^m)$ für $n, m \in \mathbb{N}$.

Satz 2.40 (Fubini, Ana 3)**.** *Seien (X_1, Σ_1, μ_1) und (X_2, Σ_2, μ_2) zwei σ-finite Maßräume und $f : X_1 \times X_2 \to \mathbb{R}$ eine messbare Funktion (bzgl. $\Sigma_1 \times \Sigma_2$). Ist $f \geq 0$, so gilt*

$$\int_{X_1\times X_2} f \, \mathrm{d}(\mu_1 \times \mu_2) = \int_{X_1} \left(\int_{X_2} f(x, y) \, \mathrm{d}\mu_2(y) \right) \mathrm{d}\mu_1(x) = \int_{X_2} \left(\int_{X_1} f(x, y) \, \mathrm{d}\mu_1(x) \right) \mathrm{d}\mu_2(y).$$

Allgemein ist $f \in \mathcal{L}^1(X_1 \times X_2, \mu_1 \times \mu_2)$ äquivalent zu

$$\int_{X_1} |f(x, \cdot)| \, \mathrm{d}\mu_1(x) \in \mathcal{L}^1(X_2, \mu_2) \quad \textit{und} \quad \int_{X_2} |f(\cdot, y)| \, \mathrm{d}\mu_2(y) \in \mathcal{L}^1(X_1, \mu_1).$$

In diesem Fall gilt auch obige Identität.

Beweis. Siehe z.B. Jürgen Elstrodt: Maß- und Integrationstheorie, Satz V.2.1. □

2.2.2 L^p-Räume

Die wichtigsten Funktionenräume der Integrationstheorie sind die L^p-Räume.

Definition 2.41 (Ana 3)**.** Sei (Ω, Σ, μ) ein Maßraum und $1 \leq p < \infty$. Dann definiert man den Raum

$$\mathcal{L}^p(\Omega, \mu) := \left\{ f : \Omega \to \mathbb{K}^m \;\middle|\; f \text{ messbar und } \int_\Omega |f|^p \, \mathrm{d}\mu < \infty \right\}.$$

Aufgrund der Konvexität der Funktion $x \mapsto |x|^p$ mit dem reellen Betrag hat $\mathcal{L}^p(\Omega,\mu)$ offenbar die Struktur eines Vektorraums. Für $f \in \mathcal{L}^p(\Omega,\mu)$ definiert man

$$\|f\|_p := \left(\int\limits_\Omega |f|^p \,\mathrm{d}\mu\right)^{\frac{1}{p}}.$$

Dann gilt für alle $f, g \in \mathcal{L}^p(\Omega,\mu), \lambda \in \mathbb{K}$

(i) $\|\lambda f\|_p = |\lambda|\,\|f\|_p$.

(ii) $\|f\|_p = 0 \;\Leftrightarrow\; f(x) = 0$ für μ-fast alle $x \in \Omega$.

Es existieren also Funktionen $f \in \mathcal{L}^p(\Omega,\mu)$ mit $f \neq 0$ und $\|f\|_p = 0$, d.h. $\|\cdot\|_p$ definiert auf $\mathcal{L}^p(\Omega,\mu)$ keine Norm. Man definiere auf $\mathcal{L}^p(\Omega,\mu)$ die Äquivalenzrelation

$$f \sim g \;\Leftrightarrow\; f = g \quad \mu\text{-fast überall.}$$

Mit der Äquivalenzklasse

$$[f] = \{g \in \mathcal{L}^p(\Omega,\mu) \mid f \sim g\}$$

identifizieren wir alle Funktionen, die mit f μ-fast überall übereinstimmen. Dann definiert man den Raum

$$L^p(\Omega,\mu) := \mathcal{L}^p(\Omega,\mu)/_\sim = \{[f] \mid f \in \mathcal{L}^p(\Omega,\mu)\}.$$

mit Norm

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

Einerseits ist $\|\cdot\|_p$ wohldefiniert, da $\|f\|_{\mathcal{L}^p} = \|g\|_{\mathcal{L}^p}$ für $f \sim g$. Andererseits impliziert $\|[f]\|_p = 0$ auch $f = 0$ μ-fast überall, d.h. $[f] = [0]$. Dass die Dreiecksungleichung auch für $p > 1$ gilt, zeigen wir im Anschluss.
Analog definiert man für den Fall $p = \infty$

$$\mathcal{L}^\infty(\Omega,\mu) = \{f : \Omega \to \mathbb{K}^m \mid f \text{ messbar und } \exists\, k \in \mathbb{R}_+ : |f| \leq k \;\; \mu\text{-f.ü. in } \Omega\}$$

und für ein $f \in \mathcal{L}^\infty(\Omega,\mu)$ setzt man

$$\|f\|_\infty := \inf\{k > 0 \mid |f(x)| < k \text{ für } \mu\text{-fast alle } x \in \Omega\}.$$

$\|f\|_\infty$ wird als **wesentliches** bzw. **essentielles Supremum** von f auf Ω bezeichnet und ist wiederum keine Norm auf $\mathcal{L}^\infty(\Omega,\mu)$. Also definiert man

$$L^\infty(\Omega,\mu) := \mathcal{L}^\infty(\Omega,\mu)/_\sim = \{[f] \mid f \in \mathcal{L}^\infty(\Omega,\mu)\}$$

und erhält einen normierten Raum $(L^\infty(\Omega,\mu), \|\cdot\|_\infty)$.

Bemerkung: Aus diesen Definitionen folgt, dass die Elemente aus $L^\infty(\Omega,\mu)$, $1 \leq p \leq \infty$, keine Funktionen sind, sondern Äquivalenzklassen von Funktionen. Trotzdem wird im Folgenden von Funktionen $f \in L^\infty(\Omega,\mu)$ gesprochen, wobei dabei ein Repräsentant der Äquivalenzklasse $[f]$ in $L^\infty(\Omega,\mu)$ gemeint ist.

Es bleibt also der Beweis der Dreiecksungleichung $\|f+g\|_p \leq \|f\|_p + \|g\|_p$ zu zeigen. Dazu werden zunächst die elementaren Ungleichungen von J. L. Jensen und O. Hölder gezeigt.

Lemma 2.42 (Jensen-Ungleichung, Ana 3)**.** *Sei (Ω, Σ, μ) ein Maßraum mit $\mu(\Omega) = 1$, $I \subset \mathbb{R}$ ein Intervall, $f : \Omega \to I$ integrierbar und $\phi : I \to \mathbb{R}$ konvex. Dann ist $\int_\Omega f \,\mathrm{d}\mu \in I$, der negative Teil $(\phi \circ f)_-$ ist integrierbar und es gilt:*

$$\phi\Big(\int\limits_\Omega f \,\mathrm{d}\mu \Big) \leq \int\limits_\Omega \phi \circ f \,\mathrm{d}\mu,$$

wobei die rechte Seite auch unendlich sein kann.

Beweis. Siehe z.B. Jürgen Elstrodt: Maß- und Integrationstheorie, Satz VI.1.3. □

Hieraus lassen sich einige wichtige Folgerungen ableiten. Zunächst sind Produkte von L^p-Funktionen im Allgemeinen nicht Elemente aus L^p, jedoch gilt

Satz 2.43 (Hölder-Ungleichung, Ana 3)**.** *Sei (Ω, Σ, μ) ein Maßraum und seien p, q konjugierte Hölder-Exponenten mit $1 \leq p, q \leq \infty$ und $\frac{1}{p} + \frac{1}{q} = 1$. Für $f \in L^p(\Omega, \mu)$ und $g \in L^q(\Omega, \mu)$ ist $fg \in L^1(\Omega, \mu)$ und es gilt*

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

Beweis. Für $p = \infty$ oder $q = \infty$ folgt die Ungleichung aus den Eigenschaften des Integrals, deswegen nehme man $1 < p, q < \infty$ an. Sei zusätzlich $g \neq 0$ in L^q mit $\|g\|_q = 1$. Definiere auf (Ω, Σ) ein neues Maß $\mathrm{d}\nu = |g|^q \,\mathrm{d}\mu$ mit

$$\nu(B) = \int\limits_\Omega \chi_B \,\mathrm{d}\nu = \int\limits_\Omega \chi_B(x)\, |g(x)|^q \,\mathrm{d}\mu(x)$$

für alle messbaren Mengen $B \in \Sigma$. Insbesondere ist dann $\nu(\Omega) = \|g\|_q^q = 1$.
Sei $A := \{x \in \Omega \mid g(x) \neq 0\}$. Dann ist

$$\|fg\|_1^p = \left| \int\limits_A |fg| \,\mathrm{d}\mu \right|^p \leq \left| \int\limits_A |f||g|^{1-q} \,\mathrm{d}\nu \right|^p.$$

Da $\phi(t) = |t|^p$ konvex ist, folgt mit der Jensen-Ungleichung wegen $(1 - q)p = -q$

$$\|fg\|_1^p \leq \int\limits_A \left| |f||g|^{1-q} \right|^p \,\mathrm{d}\nu = \int\limits_A |f|^p \,\mathrm{d}\mu \leq \|f\|_p^p,$$

was zu zeigen war. □

Schließlich lässt sich damit die Dreiecksungleichung für die Norm $\|\cdot\|_p$ zeigen.

Satz 2.44 (Minkowski-Ungleichung, Ana 3)**.** *Sei (Ω, Σ, μ) ein Maßraum und $1 \leq p \leq \infty$. Dann gilt*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p \qquad \forall\, f, g \in L^p(\Omega, \mu)$$

und $L^p(\Omega, \mu)$ ist ein normierter $\mathbb{K}$-Vektorraum.

Beweis. Für $p \in \{1, \infty\}$ ist die Aussage klar. Für $p < \infty$ folgt wegen $\frac{1}{p} + \frac{p-1}{p} = 1$ mit der Hölder-Ungleichung

$$\begin{aligned}\|f+g\|_p^p &\leq \int_\Omega |f| \cdot |f+g|^{p-1} \,\mathrm{d}\mu + \int_\Omega |g| \cdot |f+g|^{p-1} \,\mathrm{d}\mu \\ &\leq \|f\|_p \||f+g|^{p-1}\|_p + \|g\|_p \||f+g|^{p-1}\|_p \\ &= (\|f\|_p + \|g\|_p) \cdot \|f+g\|_p^{p-1}.\end{aligned}$$

□

Für $p = 2$ ist $L^2(\Omega, \mu)$ offensichtlich ein Prähilbertraum mit Skalarprodukt

$$(f, g)_{L^2(\Omega,\mu)} := \int_\Omega \overline{f} g \,\mathrm{d}\mu.$$

Eine weitere Folgerung der Jensen-Ungleichung sind die stetigen Einbettungen der L^p-Räume im Fall $\mu(\Omega) < \infty$:

Korollar 2.45 (Ana 3)**.** *Sei* (Ω, Σ, μ) *ein Maßraum mit* $\mu(\Omega) < \infty$*. Dann gilt für alle* $1 \leq p \leq q \leq \infty$

$$\|f\|_p \leq \mu(\Omega)^{\frac{1}{p} - \frac{1}{q}} \cdot \|f\|_q$$

und folglich $L^q(\Omega, \mu) \subset L^p(\Omega, \mu)$*.*

Beweis. Dies wird für $p, q \in (0, \infty)$ gezeigt. Man definiere für alle messbaren Mengen $A \subset \Omega$ die relative σ-Algebra mit

$$\Sigma_\Omega := \{A \subset \Omega \mid A \in \Sigma\} \quad \text{und} \quad \tilde{\mu}(A) := \frac{1}{\mu(\Omega)}\, \mu(A).$$

Insbesondere ist $(\Omega, \Sigma_\Omega, \tilde{\mu})$ ein Maßraum mit $\tilde{\mu}(\Omega) = 1$. Die Anwendung der Jensen-Ungleichung mit $\phi(t) = |t|^{\frac{q}{p}}$ liefert

$$\left(\int_\Omega |f|^p \,\frac{\mathrm{d}\mu}{\mu(\Omega)}\right)^{\frac{q}{p}} = \left(\int_\Omega |f|^p \,\mathrm{d}\tilde{\mu}\right)^{\frac{q}{p}} \leq \int_\Omega |f|^q \,\frac{\mathrm{d}\mu}{\mu(\Omega)}$$

und man erhält die Abschätzung, sodass auch die Inklusion folgt. □

Beispiel 2.46. Im Fall $\mu(\Omega) = \infty$ gilt die Inklusion $L^q(\Omega, \mu) \subset L^p(\Omega, \mu)$ für $p < q$ im Allgemeinen nicht:

(a) Sei $\Omega = \mathbb{R}, \Sigma = \mathcal{B}(\mathbb{R})$ mit dem Lebesgue-Maß $\lambda^1 := \lambda$. So gilt für $f_\alpha(x) := x^{-\alpha}\chi_{[1,\infty)}$

$$\|f_\alpha\|_p^p = \int_1^\infty x^{-\alpha p} \,\mathrm{d}\lambda(x) = \frac{1}{\alpha p - 1} < \infty \qquad \Leftrightarrow \qquad \alpha p > 1$$

Zum Beispiel ist $x \mapsto x^{-1} \in L^2([1, \infty), \lambda) \setminus L^1([1, \infty), \lambda)$.

(b) Für $\Omega = \mathbb{N}, \Sigma = \mathrm{Pot}(\mathbb{N})$ mit dem Zählmaß μ aus Beispiel 2.28 erhält man die Folgenräume $L^p(\Omega, \mu) = \ell^p$. In diesem Fall gilt

$$\ell^p \subset \ell^q \subset c_0 \subset c \subset \ell^\infty \qquad \forall\, 1 \leq p < q < \infty,$$

wobei c_0 den Untervektorraum der Nullfolgen und c den Raum der konvergenten Folgen bezeichne.

Satz 2.47 (Fischer-Riesz, Ana 3). *Sei (Ω, Σ, μ) ein Maßraum und $1 \leq p \leq \infty$. Dann ist $L^p(\Omega, \mu)$ vollständig bzgl. $\|\cdot\|_p$, d.h. ein Banach-Raum.*

Beweis. Man betrachte zunächst den Fall $1 \leq p < \infty$. Sei $(f_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $L^p(\Omega)$. Es genügt zu zeigen, dass eine Teilfolge $(f_{n_j})_{j\in\mathbb{N}}$ und $f \in L^p(\Omega)$ existieren mit $f_{n_j}(x) \to f(x)$ für μ-fast alle $x \in \Omega$ für $j \to \infty$. Dann gibt es nämlich für $\varepsilon > 0$ ein $n_0(\varepsilon) > 0$ mit $\|f_n - f_m\|_p < \varepsilon$ für $n, m \geq n_0(\varepsilon)$ sowie nach dem Lemma von Fatou für $n \geq n_0(\varepsilon)$

$$\|f_n - f\|_p^p = \int_\Omega \liminf_{j\to\infty} |f_n - f_{n_j}|^p \, \mathrm{d}\mu \leq \liminf_{j\to\infty} \int_\Omega |f_n - f_{n_j}|^p \, \mathrm{d}\mu \leq \varepsilon^p.$$

Man konstruiere also eine geeignete Teilfolge und wähle dazu sukzessive für $j \in \mathbb{N}$ eine aufsteigende Folge $n_j \in \mathbb{N}$ mit

$$\|f_n - f_{n_j}\|_p \leq \frac{1}{2^j} \qquad \forall\, n \geq n_j.$$

Um die Konvergenz von $(f_{n_j})_{j\in\mathbb{N}}$ für $j \to \infty$ zu zeigen, definiert man für beliebige $m \in \mathbb{N}$

$$F_m(x) := |f_{n_1}(x)| + \sum_{k=1}^m |f_{n_{k+1}}(x) - f_{n_k}(x)|.$$

Nach dem Satz über monotone Konvergenz folgt aus $F_m \in L^p(\Omega, \mu)$ wegen

$$\|F_m\|_p \leq \|f_{n_1}\|_p + \sum_{k=1}^m \|f_{n_k} - f_{n_{k+1}}\|_p \leq \|f_{n_1}\|_p + 1$$

dass auch der Limes

$$F(x) := \lim_{m\to\infty} F_m(x)$$

für μ-fast alle $x \in \Omega$ existiert und $F \in L^p(\Omega, \mu)$ gilt. Deswegen muss $|F(x)| < \infty$ für μ-fast alle $x \in \Omega$ sein und aufgrund der Teleskopsumme in F_m konvergiert die Folge $(f_{n_1} - f_{n_j})_{j\in\mathbb{N}}$ nach Analysis 1 punktweise μ-f.ü. in Ω. Folglich gibt es eine messbare Funktion $f : \Omega \to \mathbb{K}$ mit

$$f(x) := \lim_{j\to\infty} f_{n_j}(x) \qquad \text{für } \mu\text{-fast alle } x \in \Omega.$$

Da $|f_{n_j}(x)| \leq F(x)$ für μ-fast alle $x \in \Omega$ und $F \in L^p(\Omega, \mu)$ gilt, folgt aus dem Satz von der dominierten Konvergenz auch $f \in L^p(\Omega, \mu)$ und $\|f_{n_j} - f\|_p \to 0$ für $j \to \infty$, was zu zeigen war.
Für den Fall $p = \infty$ ist für eine Cauchy-Folge $(f_n)_{n\in\mathbb{N}}$ in $L^\infty(\Omega, \mu)$ die Menge

$$N := \bigcup_{n\in\mathbb{N}} \{x \in \Omega \mid |f_n(x)| > \|f_n\|_\infty\} \cup \bigcup_{n,m\in\mathbb{N}} \{x \in \Omega \mid |(f_m - f_n)(x)| > \|f_m - f_n\|_\infty\}$$

eine μ-Nullmenge und für alle $x \in N^c$ gilt

$$|f_m(x) - f_n(x)| \leq \|f_m - f_n\|_\infty \to 0 \qquad (\text{für } m, n \to \infty).$$

Daher konvergiert $(f_n)_{n\in\mathbb{N}}$ auf N^c gleichmäßig gegen

$$f \in L^\infty(\Omega, \mu) \qquad \text{mit} \quad f(x) = \lim_{n\to\infty} f_n(x) \cdot \chi_{N^c}(x).$$

Analog zum Beweis von Lemma 2.2 folgt auch $\|f_n - f\|_\infty \to 0$ für $n \to \infty$. □

Nach dem eben bewiesenen Satz enthält eine in $L^p(\Omega, \mu)$ für $p < \infty$ konvergente Folge stets eine Teilfolge, welche μ-f.ü. punktweise in Ω (aufgrund der Eindeutigkeit des Grenzwerts gegen denselben Limes) konvergiert. Im Allgemeinen ist jedoch nicht die gesamte Folge μ-f.ü. punktweise konvergent wie folgendes Beispiel zeigt.

Beispiel 2.48. Im Folgenden seien zu $\Omega \subset \mathbb{R}$ zugehörige Borel-σ-Algebren und das Lebesgue-Maß λ betrachtet.

(a) Für $\Omega = [0, 1]$ betrachte die Intervallaufteilung für $n \in \mathbb{N}$ mit

$$\Omega = \bigcup_{k=1}^{n} I_{k,n} \qquad \text{für} \quad I_{k,n} = \left[\frac{k-1}{n}, \frac{k}{n}\right].$$

Dann bildet die Funktionenfolge mit Elementen $f_{k,n} := \chi_{I_{k,n}}$ für $1 \leq k \leq n$ und $n \in \mathbb{N}$ eine in keinem Punkt konvergente Folge mit $\|f_{k,n}\|_p = \frac{1}{n} \to 0$ für $n \to \infty$.

(b) Für $\Omega = [0, 1]$ ist $(f_n)_{n\in\mathbb{N}}$ mit $f_n := n^{2/p}\chi_{[0,1/n]}$ punktweise konvergent mit $f_n(x) \to 0$, jedoch ist $\|f_n\|_p = n^{1/p}$ sogar unbeschränkt.

(c) Für $\Omega = \mathbb{R}$ liefert $(f_n)_{n\in\mathbb{N}}$ mit $f_n = n^{-1}\chi_{[0,2^n]}$ eine gleichmäßig gegen 0 konvergente Folge, die nicht in $L^p(\mathbb{R}, \lambda)$ für $p < \infty$ konvergiert, denn $\|f_n\|_p^p = n^{-p}2^n$.

Unter welchen Voraussetzungen eine punktweise μ-f.ü. konvergente Folge von L^p-Funktionen auch in der Norm konvergiert, zeigt beispielsweise der Satz von Vitaly (siehe H. W. Alt: Lineare Funktionalanalysis, Satz 1.19), welchen wir aus Zeitgründen nicht behandeln.

Im Folgenden interessiert uns die Approximierbarkeit von L^p-Funktionen. Aus der Definition 2.35 des Integrals lassen sich nicht negative messbare Funktionen durch eine Folge einfacher Funktionen bzgl. der L^1-Norm approximieren. Dieses Argument lässt sich leicht auf beliebige integrierbare bzw. Funktionen in $L^p(\Omega, \mu)$ erweitern, wobei die Approximation bzgl. $\|\cdot\|_p$ für endliche $p \geq 1$ möglich ist. Daher genügt es im Wesentlichen charakteristische Funktionen über messbare Mengen zu approximieren.
Da dies stark von der Regularität des Maßes abhängt, wählen wir im restlichen Abschnitt das Lebesgue-Maß mit der zugehörige Borelschen σ-Algebra, welches nach Lemma 2.30 sowohl von außen als auch von innen regulär ist. Zur Vereinfachung sei $L^p(\Omega, \lambda^n) = L^p(\Omega)$ für $\Omega \subset \mathbb{R}^n$ und $n \in \mathbb{N}$.

Satz 2.49. *Sei $1 \leq p < \infty$. Dann ist $C_c(\mathbb{R}^n)$ dicht in $L^p(\mathbb{R}^n)$ bzgl. der $\|\cdot\|_p$-Norm.*

Beweis. Seien $f \in L^p(\mathbb{R}^n)$ und $\varepsilon > 0$ beliebig. Dann gibt ein $R_\varepsilon > 0$ mit

$$\|f - \chi_{B_{R_\varepsilon}(0)} f\|_p < \varepsilon.$$

Nun wählen wir auf der beschränkten Menge $B_{R_\varepsilon}(0) \subset \mathbb{R}^n$ eine $\chi_{B_{R_\varepsilon}(0)} f$ in der L^p-Norm approximierende Folge von einfachen Funktionen, wobei diese die Gestalt

$$\sum_{i=1}^m \alpha_i \chi_{A_i} \qquad \text{mit } \alpha_i \in \mathbb{K},\ A_i \subset \mathbb{R}^n \text{ beschränkt}$$

besitzen. Es bleibt daher zu zeigen, dass jede Funktion χ_A für beschränkte Mengen $A \subset \mathbb{R}^n$ bzgl. $\|\cdot\|_p$ durch Funktionen aus $C_c(\mathbb{R}^n)$ approximiert werden können.
Zu A gibt es nach Lemma 2.30 zu jedem $\delta > 0$ eine kompakte Menge $K \subset A$ und eine offene Menge $U \supset A$ mit $\lambda^n(U \setminus K) < \delta$. Dann liefert beispielsweise

$$f_\delta(x) := \frac{\operatorname{dist}(x, U^c)}{\operatorname{dist}(x, K) + \operatorname{dist}(x, U^c)}$$

wegen $\operatorname{dist}(x, B) = 0 \Leftrightarrow x \in \overline{B}$ eine Funktion $f_\delta \in C_c(\mathbb{R}^n, [0,1])$ mit

$$\int_\Omega |f_\delta - \chi_A|^p \,\mathrm{d}\lambda^n = \int_{U\setminus K} |f_\delta - \chi_A|^p \,\mathrm{d}\lambda^n \le \lambda^n(U \setminus K) < \delta.$$

□

Analog zur Höhere Analysis werden wir im Folgenden einige Eigenschaften von Faltungen und entsprechende Regularisierungen von L^p-Funktionen erarbeiten, welche im weiteren Verlauf der Vorlesung auch für Sobolev-Funktionen nützlich sind.

Definition 2.50 (Ana 3)**.** Seien $f, g : \mathbb{R}^n \to \mathbb{R}$ messbar und sei

$$N := \{x \in \mathbb{R}^n \mid y \mapsto f(y)\, g(x-y) \text{ ist nicht integrierbar}\}.$$

Die **Faltung** $f * g : \mathbb{R}^n \to \mathbb{R}$ ist definiert durch

$$(f * g)(x) := \begin{cases} \displaystyle\int_{\mathbb{R}^n} f(y)\, g(x-y) \,\mathrm{d}\lambda^n(y) & \text{falls } x \notin N, \\ 0 & \text{falls } x \in N. \end{cases}$$

In den folgenden Beweisen schreiben wir, sofern klar ist, dass es sich um das Lebesgue-Maß λ^n handelt, vereinfacht $\mathrm{d}y$ anstatt $\mathrm{d}\lambda^n(y)$.

Satz 2.51 (Ana 3)**.** *Sei $1 \le p \le \infty$ und $g \in L^1(\mathbb{R}^n)$, $f \in L^p(\mathbb{R}^n)$. Dann ist die Menge N aus Definition 2.50 eine λ^n-Nullmenge und es ist $f * g \in L^p(\mathbb{R}^n)$ mit*

$$\|f * g\|_{L^p(\mathbb{R}^n)} \le \|f\|_{L^p(\mathbb{R}^n)} \|g\|_{L^1(\mathbb{R}^n)}.$$

Beweis. Für $p = \infty$ ist die Abschätzung trivial. Für $p = 1$ gilt nach dem Satz von Fubini

$$\begin{aligned}\int_{\mathbb{R}^n} |f * g| \,\mathrm{d}\lambda^n \quad &\le \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |f(y)|\, |g(x-y)| \,\mathrm{d}y \,\mathrm{d}x \\ &\overset{\text{Fubini}}{=} \int_{\mathbb{R}^n} |f(y)| \int_{\mathbb{R}^n} |g(x-y)| \,\mathrm{d}x \,\mathrm{d}y = \|f\|_1 \|g\|_1.\end{aligned}$$

Für $1 < p < \infty$ wähle man q mit $\frac{1}{p} + \frac{1}{q} = 1$ und erhält nach der Hölderschen Ungleichung

$$\int_{\mathbb{R}^n} |f * g|^p \,\mathrm{d}\lambda^n \leq \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} |g(y)|^{\frac{1}{p}} \,|g(y)|^{\frac{1}{q}} \,|f(x-y)| \,\mathrm{d}y \Big)^p \mathrm{d}x$$
$$\leq \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} |g(y)| \,|f(x-y)|^p \,\mathrm{d}y \Big)^{\frac{p}{p}} \cdot \Big(\int_{\mathbb{R}^n} |g(y)| \,\mathrm{d}y \Big)^{\frac{p}{q}} \mathrm{d}x.$$

Eine Anwendung des Satzes von Fubini liefert schließlich

$$\|f * g\|_p^p \leq \int_{\mathbb{R}^n} |g(y)| \int_{\mathbb{R}^n} |f(x-y)|^p \,\mathrm{d}x \,\mathrm{d}y \cdot \|g\|_1^{\frac{p}{q}} = \|g\|_1 \|f\|_p^p \cdot \|g\|_1^{\frac{p}{q}} = \|g\|_1^p \cdot \|f\|_p^p.$$

Der Satz von Fubini bzw. Tonelli zeigt dabei in allen Fällen $\lambda^n(N) = 0$. □

Lemma 2.52 (Ana 3)**.** *Seien $f, g, h \in L^1(\mathbb{R}^n)$. Dann gilt:*

(i) $f * g = g * f$.

(ii) $(f * g) * h = f * (g * h)$.

(iii) Ist zusätzlich $h \in L^\infty(\mathbb{R}^n)$ und $(Sf)(x) := f(-x)$, dann gilt:

$$\int_{\mathbb{R}^n} (f * g)\, h \,\mathrm{d}\lambda^n = \int_{\mathbb{R}^n} f(Sg * h) \,\mathrm{d}\lambda^n$$
$$= \int_{\mathbb{R}^n} g(Sf * h) \,\mathrm{d}\lambda^n.$$

(iv) Ist $g \in C_c^k(\mathbb{R}^n)$, dann ist $f * g \in C^k(\mathbb{R}^n)$ und für $k \in \mathbb{N}_0$ gilt:

$$\partial^\alpha (f * g) = f * \partial^\alpha g$$

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

Beweis. (i) bis (iii) folgen direkt aus dem Satz von Fubini bzw. dem eben Satz 2.51. Es bleibt also (iv) zu zeigen, wobei wir nur die Fälle $k \leq 1$ zeigen ($k > 1$ geht analog). Für $g \in C_c^0(\mathbb{R}^n)$ ist g gleichmäßig stetig auf dem Träger und es folgt $f * g \in C(\mathbb{R}^n)$, denn

$$\left| \int_{\mathbb{R}^n} \big(g(x-y) - g(x_0 - y) \big) f(y) \,\mathrm{d}y \right| \leq \varepsilon \|f\|_1 \to 0 \quad \text{für } x \to x_0.$$

Man nehme nun $g \in C_c^1(\mathbb{R}^n)$ an. Insbesondere ist ∇g gleichmäßig stetig. Für alle $y \in \mathbb{R}^n$ und $h \in \mathbb{R}^n, |r| < 1$ folgt

$$g(x + h - y) - g(x - y) - h \cdot \nabla g(x - y) = \int_0^1 h \cdot \nabla g(x + sh - y) - h \cdot \nabla g(x - y) \,\mathrm{d}s,$$

sodass der Differenzenquotient gleichmäßig in $y \in \mathbb{R}^n$ gegen $\nabla g(x - \cdot)$ konvergiert. Dies impliziert aber die Aussage gemäß

$$\left| \frac{1}{h} \big((f * g)(x + h) - (f * g)(x) \big) - (\nabla g * f)(x) \right| \leq \varepsilon(|h|) \|f\|_1 \qquad \text{für } h \to 0.$$

□

Bereits in Satz 2.8 haben wir Abschneidefunktionen kennengelernt, etwa die C^∞-Funktion

$$\psi : \mathbb{R}_+ \to \mathbb{R}_+, \quad x \mapsto \begin{cases} \mathrm{e}^{-1/x} & x > 0, \\ 0 & x \leq 0 \end{cases}$$

oder die Funktion

$$\psi_k : \mathbb{R}^n \to [0,1], \quad x \mapsto \psi(r_k^2 - \|x - x_k\|^2)$$

mit kompakten Träger $\operatorname{supp}(\psi_k) = \overline{B_{r_k}(x_k)}$.

Definition 2.53 (Dirac-Folge, Ana 3)**.** Eine Folge $(\varphi_k)_{k\in\mathbb{N}}$ in $L^1(\mathbb{R}^n)$ heißt **Dirac-Folge**, falls folgende Bedingungen erfüllt sind:

(i) Für jedes $k \in \mathbb{N}$ ist $\varphi_k \geq 0$.

(ii) Für jedes $k \in \mathbb{N}$ ist $\|\varphi_k\|_1 = 1$.

(iii) Für jedes $\varrho > 0$ ist $\displaystyle\int\limits_{\mathbb{R}^n \setminus B_\varrho(0)} \varphi_k \to 0$ für $k \to \infty$.

Beispiel 2.54.

(a) Die Folge $(\psi_k)_{k\in\mathbb{N}}$ erfüllt für $x_k = 0$ und $(r_k)_{k\in\mathbb{N}}$ als Nullfolge die Bedingungen (i) und (iii), jedoch ist ψ_k nicht normiert. Für die Normierung wähle ψ_1 für $x_k = 0, r_k = 1$, eine positive Nullfolge $(\rho_k)_{k\in\mathbb{N}}$ und definiere die glatte Dirac-Folge $(\eta_k)_{k\in\mathbb{N}}$ mit

$$\eta_k(x) := C\rho_k^{-n}\psi_1(x/\rho_k) \qquad \text{mit} \quad C := \|\psi_1\|_1^{-1} = \|\rho_k^{-n}\psi_1(\cdot/\rho_k)\|_1^{-1}.$$

Beachte, dass $\operatorname{supp}(\eta_k) = \overline{B_{\rho_k}(0)}$ und damit (iii) erfüllt ist.
Allgemein wird eine Dirac-Folge $(\varphi_\varepsilon)_{\varepsilon>0}$ glatter Funktionen mit $\operatorname{supp}(\varphi_\varepsilon) \subset B_\varepsilon(0)$ eine **Standard-Dirac-Folge** genannt.

(b) Zu $\varphi \in L^1(\mathbb{R}^n)$ mit $\varphi \geq 0$ lässt sich analog für $\varepsilon > 0$

$$\varphi_\varepsilon(x) := \|\varphi\|_1^{-1} \cdot \varepsilon^{-n}\varphi\left(\frac{x}{\varepsilon}\right)$$

definieren mit

$$\int\limits_{\mathbb{R}^n} \varphi_\varepsilon = 1 \quad \text{und} \quad \int\limits_{\mathbb{R}^n \setminus B_\varrho(0)} \varphi_\varepsilon \to 0 \text{ für } \varepsilon \to 0.$$

Daraus folgt, dass für jede Nullfolge $(\varepsilon_k)_{k\in\mathbb{N}}$ die Folge $(\varphi_{\varepsilon_k})_{k\in\mathbb{N}}$ eine (allgemeine) Dirac-Folge ist.

Satz 2.55 (Ana 3)**.** *Sei $(\varphi_\varepsilon)_{\varepsilon>0}$ eine Standard-Dirac-Folge. Dann gelten*

(i) Sei $f \in C(\mathbb{R}^n)$, dann gilt $\varphi_\varepsilon * f \to f$ für $\varepsilon \to 0$ gleichmäßig auf kompakten Mengen in $\mathbb{R}^n$.

(ii) Sei $f \in L^p(\mathbb{R}^n)$ für $1 \leq p < \infty$, dann konvergiert $\varphi_\varepsilon * f \to f$ für $\varepsilon \to 0$ in $L^p(\mathbb{R}^n)$.

Beweis. (i) Sei $K \subset \mathbb{R}^n$ eine kompakte Menge. Dann ist $f_{|K}$ gleichmäßig stetig, d.h. für jedes $\overline{\varepsilon} > 0$ gibt es ein $\delta > 0$ mit

$$|f(x-y) - f(x)| < \overline{\varepsilon} \qquad \forall\, x \in K,\ y \in B_\delta(0).$$

Damit folgt für hinreichend kleines $\varepsilon \leq \delta$ und alle $x \in K$

$$|(\varphi_\varepsilon * f)(x) - f(x)| \leq \int\limits_{B_\varepsilon(0)} |f(x-y) - f(x)|\, \varphi_\varepsilon(y)\, \mathrm{d}y \leq \overline{\varepsilon}\|\varphi_\varepsilon\|_1 = \overline{\varepsilon}.$$

(ii) Wähle zu $\overline{\varepsilon} > 0$ nach Satz 2.49 eine Funktion $g \in C_c(\mathbb{R}^n)$ mit $\|f-g\|_p < \overline{\varepsilon}/3$. Wegen

$$\operatorname{supp}(\varphi_\varepsilon * g) \subset \overline{B_\varepsilon(0)} + \operatorname{supp}(g) \subset \overline{B_1(0)} + \operatorname{supp}(g) =: K$$

konvergiert die stetige Funktion $\varphi_\varepsilon * g$ gleichmäßig auf K gegen g, insbesondere also auch in $L^p(\mathbb{R}^n)$ mit

$$\|\varphi_\varepsilon * g - g\|_p < \frac{\overline{\varepsilon}}{3} \qquad \text{für } \varepsilon \leq \varepsilon_0(\overline{\varepsilon}).$$

Schätzt man nun ab gemäß

$$\begin{aligned}\|\varphi_\varepsilon * f - f\|_p &\leq \|\varphi_\varepsilon * (f-g)\|_p + \|\varphi_\varepsilon * g - g\|_p + \|g - f\|_p \\ &\leq 2\|f-g\|_p + \|\varphi_\varepsilon * g - g\|_p < \overline{\varepsilon},\end{aligned}$$

so folgt die Konvergenz $\|\varphi_\varepsilon * f - f\|_p \to 0$ für $\varepsilon \to 0$.

□

Korollar 2.56 (Ana 3)**.** *Sei $\Omega \subset \mathbb{R}^n$ offen und $1 \leq p < \infty$. Dann liegt $C_c^\infty(\Omega)$ dicht in $L^p(\Omega)$.*

Beweis. Zunächst setzt man f auf $\mathbb{R}^n$ in L^p durch Null fort, indem man $\overline{f} \in L^p(\mathbb{R}^n)$ definiert durch

$$\overline{f}(x) = \begin{cases} f(x) & \text{falls } x \in \Omega, \\ 0 & \text{falls } x \in \mathbb{R}^n \setminus \Omega. \end{cases}$$

Dann ist $\|\overline{f}\|_{L^p(\mathbb{R}^n)} \leq \|f\|_{L^p(\Omega)}$. Wähle außerdem zur offenen Menge Ω eine kompakte Ausschöpfung $(K_k)_{k\in\mathbb{N}}$ wie in Beispiel 2.6 mit

$$\Omega = \bigcup_{k\in\mathbb{N}} K_k \qquad \text{und} \quad \operatorname{dist}(K_k, \mathbb{R}^n \setminus \Omega) \geq 2/k \quad \text{falls} \quad \Omega \subsetneq \mathbb{R}^n.$$

Definiere $g_k := \chi_{K_k}\overline{f} \in L^p(\mathbb{R}^n)$ und $f_k := \eta_k * g_k \in C^\infty(\mathbb{R}^n)$, dann ist $f_k \in C_c^\infty(\Omega)$, da

$$\operatorname{supp}(f_k) \subset \overline{B_{1/k}(0)} + K_k \subset \Omega.$$

Nun ist mit der Youngschen Ungleichung aus Satz 2.51

$$\begin{aligned}\|f - f_k\|_{L^p(\Omega)} &= \|\overline{f} - f_k\|_{L^p(\mathbb{R}^n)} \\ &\leq \|\overline{f} - \eta_k * \overline{f}\|_{L^p(\mathbb{R}^n)} + \|\eta_k * (\overline{f} - g_k)\|_{L^p(\mathbb{R}^n)} \\ &\leq \|\overline{f} - \eta_k * \overline{f}\|_{L^p(\mathbb{R}^n)} + \|\overline{f} - g_k\|_{L^p(\mathbb{R}^n)}.\end{aligned}$$

Nach dem eben bewiesenen Satz 2.55 genügt es dann $\|\overline{f} - g_k\|_{L^p(\mathbb{R}^n)} \to 0$ zu zeigen. Offensichtlich konvergiert aber $g_k \to \overline{f}$ punktweise und ist durch $|\overline{f}| \in L^p(\mathbb{R}^n)$ beschränkt, d.h. majorisierte Konvergenz liefert die Aussage. □

Korollar 2.57. *Sei $\Omega \subset \mathbb{R}^n$ offen und $1 \leq p < \infty$. Dann ist $L^p(\Omega)$ separabel.*

Beweis. Nach Korollar 2.56 gibt es zu $f \in L^p(\Omega)$ ein $g \in C_c^\infty(\Omega)$ mit $\|f-g\|_p < \varepsilon/2$. Wählt man wiederum eine kompakte Ausschöpfung $\Omega = \bigcup_{m\in\mathbb{N}} K_m$, so gibt es ein $m \in \mathbb{N}$ mit $\operatorname{supp} g \subset K_m$. Nach Satz 2.15 von Stone-Weierstraß lässt sich g auf K_m gleichmäßig durch Polynome mit rationalen Koeffizienten approximieren, d.h. es gibt ein solches Polynom p mit

$$\|g - p\chi_{K_m}\|_{L^\infty(K_m)} < \frac{\varepsilon}{2}\lambda^n(K_m)^{-1/p}.$$

Zusammen liefert dies

$$\|f - p\chi_{K_m}\|_{L^p(\Omega)} \leq \|f-g\|_{L^p(\Omega)} + \|g - p\chi_{K_m}\|_{L^p(K_m)} < \varepsilon.$$

Nun liefert die abzählbare Menge

$$\bigcup_{m\in\mathbb{N}} P_m \qquad \text{mit} \quad P_m := \{p\chi_{K_m} \mid p \text{ ist Polynom mit rationalen Koeffizienten}\}$$

die Behauptung. □

Wie man in den Übungen sieht, gelten diese Aussagen im Allgemeinen nicht für $p = \infty$.

Mit diesen Approximationen lässt sich ein wichtiges Lemma der Variationsrechnung beweisen, welches sogar für lokal integrierbare Funktionen gültig ist.

Definition 2.58 (L^p_{loc}-Räume)**.** Sei $1 \leq p \leq \infty$, $\Omega \subset \mathbb{R}^n$ offen und definiere

$$L^p_{\text{loc}}(\Omega) := \{f : \Omega \to \mathbb{R} \text{ messbar} \mid f \in L^p(K) \text{ für alle kompakten Teilmengen } K \subset \Omega\}.$$

Sei $(f_k)_{k\in\mathbb{N}}$ eine Folge in $L^p_{\text{loc}}(\Omega)$. Man sagt, dass $f_k \to f$ in $L^p_{\text{loc}}(\Omega)$, wenn $f \in L^p_{\text{loc}}(\Omega)$ und $\|(f_k - f)|_K\|_p \to 0$ für alle kompakten Teilmengen $K \subset \Omega$ und $k \to \infty$.

Falls $K \subset \mathbb{R}^n$ kompakt in Ω enthalten ist, d.h. $K \subset \Omega$, schreibt man auch $K \subset\subset \Omega$. Offensichtlich gilt $L^p(\Omega) \subset L^p_{\text{loc}}(\Omega) \subset L^q_{\text{loc}}(\Omega)$ für alle $p \geq q$.

Lemma 2.59. *Sei $\Omega \subset \mathbb{R}^n$ offen und $f \in L^1_{loc}(\Omega)$ sowie $(\varphi_\varepsilon)_{\varepsilon>0}$ eine Standard-Dirac-Folge. Dann gelten*

(i) $\varphi_\varepsilon * f \to f$ für $\varepsilon \to 0$ in $L^1_{loc}(\Omega)$. Genauer gibt es für jedes $K \subset\subset \Omega$ ein $\varepsilon_0(K) > 0$, sodass die Funktion $y \to \varphi_\varepsilon(y)\, f(x-y)$ für $\varepsilon \leq \varepsilon_0(K), x \in K$ integrierbar ist und es gilt $\varphi_\varepsilon * f \to f$ in $L^1(K)$.

(ii) Gilt $\displaystyle\int_\Omega f\psi \,\mathrm{d}\lambda^n = 0$ für alle $\psi \in C_c^\infty(\Omega)$, dann ist $f = 0$ λ^n-fast überall in Ω.

Beweis. (i) Sei $K \subset\subset \Omega$ und $\delta := \operatorname{dist}(K, \mathbb{R}^n \setminus \Omega)$. Dann ist $\delta > 0$, da K kompakt ist, und man definiere die kompakte Umgebung K' von K gemäß

$$K' := \{x \in \mathbb{R}^n \mid \operatorname{dist}(x, K) \leq \delta/2\} \subset \Omega.$$

Setze f auf $\mathbb{R}^n \setminus K'$ durch 0 fort zu $\overline{f} \in L^1(\mathbb{R}^n)$ definiert durch

$$\overline{f}(x) = \begin{cases} f(x) & \text{falls } x \in K', \\ 0 & \text{falls } x \in \mathbb{R}^n \setminus K'. \end{cases}$$

Für hinreichend kleines $\varepsilon \leq \delta/2$ ist dann

$$(\varphi_\varepsilon * f)(x) = (\varphi_\varepsilon * \overline{f})(x) \qquad \forall\, x \in K$$

und die Aussage folgt aus Satz 2.55 (ii).

(ii) Wähle zu Ω eine kompakte Ausschöpfung $(K_m)_{m\in\mathbb{N}}$ mit

$$\Omega = \bigcup_{m\in\mathbb{N}} K_m \qquad \text{mit} \quad \operatorname{dist}(K_m, \mathbb{R}^n \setminus \Omega) \geq 2/m.$$

Dann hat für $x \in K_m$ die Funktion $\eta_k(x - \cdot)$ kompakten Träger in Ω für hinreichend großes $k \geq m$ und durch testen der Variationsgleichung folgt $(\eta_k * f)(x) = 0$. Nach (i) ist dann $f = 0$ λ^n-f.ü. in K_m und die Aussage folgt.

□

2.3 Sobolev-Räume

Neben der Integration, für welche die L^p-Räume besonders geeignet sind, ist es oft wichtig die Ableitungen einer Funktion zu betrachten. Im klassischen Sinne differenzierbar sind dabei nicht alle L^p-Funktionen. Da der Begriff der klassischen Ableitung zu restriktiv ist, führen wir die schwachen Ableitungen ein und betrachten die sogenannten Sobolev-Räume, die aus L^p-Funktionen bestehen, deren schwache Ableitungen wiederum in L^p liegen. Analog lässt sich dieses Problem durch die (abstrakte) Vervollständigung lösen wie wir sehen werden.

Als Beispiel soll hier das Skalarprodukt

$$(f, g) := \int_0^1 f(x)\, g(x) \, + f'(x)\, g'(x) \,\mathrm{d}x \qquad \forall\, f, g \in C^1([0,1])$$

dienen. Da $(C([0,1]), \|\cdot\|_2)$ mit der L^2-Norm bereits unvollständig ist, kann auch $C^1([0,1])$ mit der von obigem Skalarprodukt induzierten Norm nicht vollständig sein. Der durch Vervollständigung des Prähilbertraums $(C^1([0,1]), (\cdot,\cdot))$ entstehende Quotientenraum lässt sich wiederum durch Auswahl von Repräsentanten (Funktionen) als Funktionenraum auffassen, welcher mit $H^{1,2}(\Omega)$ bezeichnet wird. Dies lässt sich leicht verallgemeinern zu

Definition 2.60. Sei $\Omega \subset \mathbb{R}^n$ offen mit der zugehörigen Borelschen σ-Algebra sowie dem Lebesgue-Maß λ^n und $m \in \mathbb{N}, 1 \leq p \leq \infty$. Dann definiert $\|\cdot\|_{m,p}$ eine Norm auf $X := \{f \in C^m(\Omega) \mid \|f\|_{m,p} < \infty\}$ mit

$$\|f\|_{m,p} := \left(\sum_{\alpha\in\mathbb{N}_0^n, |\alpha|\leq m} \|\partial^\alpha f\|^p_{L^p(\Omega)} \right)^{1/p} \qquad \text{für} \quad p < \infty$$

und für $p = \infty$

$$\|f\|_{m,\infty} := \sum_{\alpha \in \mathbb{N}_0^n, |\alpha| \le m} \|\partial^\alpha f\|_{L^\infty(\Omega)}.$$

Der Sobolev-Raum $H^{m,p}(\Omega)$ ist dann definiert als die Vervollständigung von X bezüglich der Norm $\|\cdot\|_{m,p}$.

Führt man sich den Vervollständigungsprozess von Satz 1.29 nochmals vor Augen, so wirkt dieser sehr abstrakt und wir benötigen ein handlicheres Kriterium.
Ist nämlich $f \in H^{m,p}(\Omega)$, so gibt es eine zugehörige Cauchy-Folge $(f_k)_{k\in\mathbb{N}}$ in X und für jeden Multiindex $\alpha \in \mathbb{N}_0^n$ mit $|\alpha| \le m$ ist $(\partial^\alpha f_k)_{k\in\mathbb{N}}$ eine Cauchy-Folge in $L^p(\Omega)$, denn

$$\|\partial^\alpha f_k - \partial^\alpha f_l\|_p \le \|f_k - f_l\|_{m,p} \to 0 \qquad \text{für } k, l \to \infty.$$

Da $L^p(\Omega)$ vollständig ist, gibt es Funktionen $g_\alpha \in L^p(\Omega)$ mit $\partial^\alpha f_k \to g_\alpha$ für $k \to \infty$. Für beliebige $\psi \in C_c^\infty(\Omega)$ folgt damit nach Höherer Analysis auch

$$\begin{aligned}\int_\Omega \partial^\alpha \psi f \,\mathrm{d}\lambda^n &= \lim_{k\to\infty} \int_\Omega \partial^\alpha \psi f_k \,\mathrm{d}\lambda^n = (-1)^{|\alpha|} \lim_{k\to\infty} \int_\Omega \psi \partial^\alpha f_k \,\mathrm{d}\lambda^n \\ &= (-1)^{|\alpha|} \int_\Omega \psi g_\alpha \,\mathrm{d}\lambda^n.\end{aligned}$$

Dies ergibt eine Beziehung ähnlich zur partiellen Integration zwischen g_α und $g_0 = f$ für jedes $\alpha \in \mathbb{N}_0^n$ mit $|\alpha| \le m$. Die Norm von $f = [(f_k)_{k\in\mathbb{N}}]$ ist damit gegeben durch

$$\|f\|_{m,p} = \lim_{k\to\infty} \|f_k\|_X = \left(\sum_{\alpha \in \mathbb{N}_0^n, |\alpha| \le m} \|g_\alpha\|_p^p \right)^{1/p},$$

falls $p < \infty$ ist bzw. für $p = \infty$ folgt nach Korollar 1.34

$$\|f\|_{m,p} = \lim_{k\to\infty} \|f_k\|_X = \sum_{\alpha \in \mathbb{N}_0^n, |\alpha| \le m} \|g_\alpha\|_\infty.$$

Die Eigenschaft partiell integrieren zu können, liefert uns den zweiten alternativen Ansatz für Sobolev-Räume.

Definition 2.61. Sei $\Omega \subset \mathbb{R}^n$ offen. Dann heißt $f \in L^1_{\text{loc}}(\Omega)$ **schwach differenzierbar**, falls es Funktionen $g_1, \dots, g_n \in L^1_{\text{loc}}(\Omega)$ gibt mit

$$\int_\Omega \partial_i \psi f \,\mathrm{d}\lambda^n = -\int_\Omega \psi g_i \,\mathrm{d}\lambda^n \qquad \forall\, \psi \in C_c^\infty(\Omega)$$

für jedes $i = 1, \dots, n$. Analog heißt f **m-mal schwach differenzierbar**, falls es für jedes $\alpha \in \mathbb{N}_0^n, |\alpha| \le m$ Funktionen $g_\alpha \in L^1_{\text{loc}}(\Omega)$ gibt mit

$$\int_\Omega \partial^\alpha \psi f \,\mathrm{d}\lambda^n = (-1)^{|\alpha|} \int_\Omega \psi g_\alpha \,\mathrm{d}\lambda^n \qquad \forall\, \psi \in C_c^\infty(\Omega).$$

Die Funktionen g_i bzw. g_α werden als **schwache Ableitungen** von f bezeichnet und man schreibt in diesem Fall $g_i = \partial_i f, g_\alpha = \partial^\alpha f$.

Bemerkung: Da die linken Integrale invariant sind unter einer Abänderung von f auf einer Lebesgue-Nullmenge, hängt die schwache Ableitung nur von der Äquivalenzklasse von f ab. Offensichtlich ist jede klassische Ableitung auch eine schwache Ableitung und nach Lemma 2.59 stimmen sie λ^n-f.ü. überein, d.h. es lässt sich ein klassisch differenzierbarer Vertreter auswählen.

Beispiel 2.62.

(a) Sei $\Omega = (-1, 1)$ und $f(x) = |x|$. Dann ist f schwach differenzierbar und besitzt die schwache Ableitung $f'(x) = \operatorname{sgn}(x)$.

(b) Sei $\Omega = (-1, 1)$ und $f(x) = \operatorname{sgn}(x)$. Dann ist f nicht schwach differenzierbar, denn

$$\int_\Omega f\,\varphi'\,\mathrm{d}\lambda = -2\,\varphi(0) \quad \forall\,\varphi \in C_c^\infty(\Omega).$$

Jedoch existiert kein $g \in L^1_{\mathrm{loc}}(\Omega)$, sodass $\int_\Omega g\,\varphi\,\mathrm{d}x = \varphi(0)$ für alle $\varphi \in C_c^\infty(\Omega)$ gilt.

(c) Sei $\Omega = B_1(0) \subset \mathbb{R}^n$ für $n \geq 2$, $\alpha \in \mathbb{R} \setminus \{0\}$ und $f(x) = |x|^\alpha$ für $x \neq 0$ und $f(0) = 0$. Dann ist f genau dann schwach differenzierbar, wenn $\alpha > -(n-1)$.

Definition 2.63 (Sobolev-Raum)**.** Sei $\Omega \subset \mathbb{R}^n$ offen mit der zugehörigen Borelschen σ-Algebra sowie dem Lebesgue-Maß λ^n und $m \in \mathbb{N}, 1 \leq p \leq \infty$. Dann definiere man den Funktionenraum

$$W^{m,p}(\Omega) := \left\{ f \in L^p(\Omega) \mid \text{es gibt schwache Ableitungen } \partial^\alpha f \in L^p(\Omega)\ \forall\,\alpha \in \mathbb{N}_0^n, |\alpha| \leq m \right\}$$

mit der Norm

$$\|f\|_{m,p} := \left(\sum_{\alpha \in \mathbb{N}_0^n, |\alpha| \leq m} \|\partial^\alpha f\|^p_{L^p(\Omega)} \right)^{1/p} \qquad \text{für} \quad p < \infty$$

und

$$\|f\|_{m,\infty} := \sum_{\alpha \in \mathbb{N}_0^n, |\alpha| \leq m} \|\partial^\alpha f\|_{L^\infty(\Omega)} \qquad \text{für} \quad p = \infty.$$

Satz 2.64. *Für $m \in \mathbb{N}$, $1 \leq p \leq \infty$ und $\Omega \subset \mathbb{R}^n$ offen ist $W^{m,p}(\Omega)$ ein Banachraum. Im Fall $p = 2$ ist $W^{m,2}(\Omega)$ ein Hilbertraum mit Skalarprodukt*

$$(f, g)_{m,2} := \sum_{\alpha \in \mathbb{N}_0^n, |\alpha| \leq m} \int_\Omega \partial^\alpha f\,\partial^\alpha g\,\mathrm{d}\lambda^n \qquad \forall\,f, g \in W^{m,2}(\Omega).$$

Beweis. Siehe Übungen. □

Bisher wurde die Beziehung $H^{m,p}(\Omega) \subset W^{m,p}(\Omega)$ zwischen den Sobolev-Räume bereits gezeigt. Wir werden im Folgenden zeigen, dass die Räume für $p < \infty$ identisch sind. Im Fall $p = \infty$ kann dies nicht gelten wie Beispiel 2.62 (a) zeigt. Zunächst benötigen wir jedoch Aussagen über die Approximierbarkeit von Sobolev-Funktionen aus $W^{m,p}(\Omega)$.

Lemma 2.65. *Sei $\Omega \subset \mathbb{R}^n$ offen und $f \in W^{m,p}(\Omega)$, $1 \leq p < \infty, m \in \mathbb{N}$ sowie $(\eta_\varepsilon)_{\varepsilon>0}$ eine Standard-Dirac-Folge. Sei $\overline{f}$ durch Null auf ganz $\mathbb{R}^n$ fortgesetzt und*

$$f_\varepsilon(x) := (\eta_\varepsilon * \overline{f})(x) = \int_{\mathbb{R}^n} \eta_\varepsilon(x-y)\,\overline{f}(y)\,\mathrm{d}y,$$

dann gelten

(i) Ist $\Omega = \mathbb{R}^n$, so ist $f_\varepsilon \in C^\infty(\mathbb{R}^n) \cap W^{m,p}(\mathbb{R}^n)$ für jedes $\alpha \in \mathbb{N}_0^n, |\alpha| \leq m$ und es gilt $\partial^\alpha f_\varepsilon = (\partial^\alpha f)_\varepsilon$ sowie $f_\varepsilon \to f$ in $W^{m,p}(\mathbb{R}^n)$ für $\varepsilon \to 0$.

(ii) Ist $\Omega \subsetneq \mathbb{R}^n$, dann gilt für alle offenen Mengen $U \subset \Omega$ mit $\delta = \operatorname{dist}(U, \partial\Omega) > 0$ $f_\varepsilon \in C^\infty(U) \cap W^{m,p}(U)$ für alle $0 < \varepsilon < \delta$ und $f_\varepsilon \to f$ in $W^{m,p}(U)$ für $\varepsilon \to 0$.

Beweis. Nach Lemma 2.52 ist $f_\varepsilon \in C^\infty(\mathbb{R}^n)$ bzw. $C^\infty(U)$, da $\eta_\varepsilon \in C_c^\infty(B_\varepsilon(0))$, und

$$(\partial^\alpha f_\varepsilon)(x) = (\overline{f} * \partial^\alpha \eta_\varepsilon)(x) = \int_\Omega \partial_x^\alpha \eta_\varepsilon(x-y)\,f(y)\,\mathrm{d}y.$$

Nach der Kettenregel ist dann

$$(\partial^\alpha f_\varepsilon)(x) = (-1)^{|\alpha|} \int_\Omega \partial_y^\alpha \big(\eta_\varepsilon(x-y)\big) f(y)\,\mathrm{d}y.$$

(i) Für $\Omega = \mathbb{R}^n$ folgt dann bereits aus $\eta_\varepsilon \in C_c^\infty(\mathbb{R}^n)$ und der Definition der schwachen Ableitung

$$(\partial^\alpha f_\varepsilon)(x) = \int_\Omega \eta_\varepsilon(x-y)\,\partial^\alpha f(y)\,\mathrm{d}y = (\partial^\alpha f)_\varepsilon(x).$$

Wegen $\partial^\alpha f \in L^p(\mathbb{R}^n)$ für alle $\alpha \in \mathbb{N}_0^n, |\alpha| \leq m$ ist nach Satz 2.51 $(\partial^\alpha f)_\varepsilon \in L^p(\mathbb{R}^n)$ und damit $f_\varepsilon \in W^{m,p}(\mathbb{R}^n)$. Schließlich folgt nach Satz 2.55

$$\begin{aligned}\|f_\varepsilon - f\|_{m,p}^p &= \sum_{\alpha\in\mathbb{N}_0^n, |\alpha|\leq m} \|\partial^\alpha f_\varepsilon - \partial^\alpha f\|_{L^p(\mathbb{R}^n)}^p \\ &= \sum_{\alpha\in\mathbb{N}_0^n, |\alpha|\leq m} \|(\partial^\alpha f)_\varepsilon - \partial^\alpha f\|_{L^p(\mathbb{R}^n)}^p \to 0 \qquad \text{für} \quad \varepsilon \to 0.\end{aligned}$$

(ii) Für $\Omega \subsetneq \mathbb{R}^n$ wähle $x \in \Omega$ mit $\operatorname{dist}(x, \partial\Omega) > \varepsilon$, dann liegt die Funktion $y \mapsto \eta_\varepsilon(x-y)$ in $C_c^\infty(\Omega)$. Wegen $f \in W^{m,p}(\Omega)$ erhält man per Definition der schwachen Ableitung wiederum $(\partial^\alpha f_\varepsilon)(x) = (\partial^\alpha f)_\varepsilon(x)$. Also stimmen $\partial^\alpha f_\varepsilon$ und $(\partial^\alpha f)_\varepsilon$ auf U überein und es folgt mit Satz 2.51 $f_\varepsilon \in W^{m,p}(U)$. Analog zum Fall $\Omega = \mathbb{R}^n$ folgt die Konvergenz auf U, denn Satz 2.55 (ii) gilt aufgrund der Dichtheit von $C_c^\infty(U)$ in $L^p(U)$ ebenso auf U statt $\mathbb{R}^n$.

□

Lemma 2.66 (Produktregel)**.** *Sei $\Omega \subset \mathbb{R}^n$ offen, $1 \leq p, q \leq \infty$ mit $\frac{1}{p} + \frac{1}{q} = 1$. Seien $f \in W^{1,p}(\Omega)$, $g \in W^{1,q}(\Omega)$. Dann ist $f \cdot g \in W^{1,1}(\Omega)$ und es gilt*

$$\partial_i(f \cdot g) = f \cdot \partial_i g + \partial_i f \cdot g \qquad \text{für} \quad i = 1, \ldots, n.$$

Beweis. Aufgrund der Symmetrie betrachten wir nur den Fall $p < \infty$. Durch Fortsetzung von f, g mit Null auf ganz $\mathbb{R}^n$ betrachte man die Glättung $f_\varepsilon = \eta_\varepsilon * \overline{f}$ wie im vorherigen Lemma. Dann gilt für alle $\psi \in C_c^\infty(\Omega)$

$$\int_\Omega (\psi f_\varepsilon)\, \partial_i g \,\mathrm{d}\lambda^n = -\int_\Omega (\psi \partial_i f_\varepsilon + f_\varepsilon \partial_i \psi)\, g \,\mathrm{d}\lambda^n.$$

Benutzt man $\|f_\varepsilon - f\|_{1,p} \to 0$ für $\varepsilon \to 0$ nach Lemma 2.65 (a), so folgt mit der Hölderschen Ungleichung für $\varepsilon \to 0$

$$\int_\Omega \psi(f\,\partial_i g + \partial_i f\, g) \,\mathrm{d}\lambda^n = -\int_\Omega f g \partial_i \psi \,\mathrm{d}\lambda^n \qquad \forall\, \psi \in C_c^\infty(\Omega).$$

Nach der Hölder-Ungleichung ist ebenso $f\,\partial_i g + \partial_i f\, g \in L^1(\Omega)$ und wir erhalten die schwache Ableitung $\partial_i(f \cdot g) \in L^1(\Omega)$. □

Erst im Jahr 1964 wurde die Gleichheit von $H^{m,p}$ und $W^{m,p}$ für $p < \infty$ bewiesen:

Satz 2.67 (Meyers-Serrin). *Sei $\Omega \subset \mathbb{R}^n$ offen und $f \in W^{m,p}(\Omega)$ für $1 \le p < \infty$, $m \in \mathbb{N}$. Dann existiert eine Folge $(f_k)_{k\in\mathbb{N}}$ in $W^{m,p}(\Omega) \cap C^\infty(\Omega)$ mit $f_k \to f$ in $W^{m,p}(\Omega)$ für $k \to \infty$. Damit folgt die Identität $H^{m,p}(\Omega) = W^{m,p}(\Omega)$.*

Beweis. Wegen $H^{m,p}(\Omega) \subset W^{m,p}(\Omega)$ folgt die Identität der Sobolev-Räume aus der ersten Aussage. Wähle zu $\Omega \subsetneq \mathbb{R}^n$ die offene Ausschöpfung

$$\Omega_k := \{x \in \Omega \mid |x| < k, \operatorname{dist}(x, \mathbb{R}^n \setminus \Omega) > 1/k\}$$

für $k \in \mathbb{N}$ und setze $\Omega_0 = \Omega_{-1} := \emptyset$. (Falls $\Omega = \mathbb{R}^n$ ist, wähle man eine Familie von offenen Quadern.) Dann liefert

$$\Omega = \bigcup_{k\in\mathbb{N}} U_k \qquad \text{mit} \quad U_k := \Omega_{k+1} \setminus \overline{\Omega_{k-1}} \subset \Omega$$

eine offene Überdeckung von Ω. Nach Satz 2.8 gibt es eine untergeordnete Zerlegung der Eins zu $(U_k)_{k\in\mathbb{N}}$, d.h. es existieren glatte Funktionen, deren kompakter Träger in einer der Mengen U_k liegt. Bezeichne für $k \in \mathbb{N}$ η_k die Summe der endlich vielen glatten Funktionen dieser Partition mit Träger in U_k. Dann erfüllt $\psi_k \in C_c^\infty(U_k)$ weiterhin für alle $x \in \Omega$ die endliche Summe

$$\sum_{k\in\mathbb{N}} \psi_k(x) = 1.$$

Durch Iteration der Produktregel folgt dann $\psi_k f \in W^{m,p}(\Omega)$ für $f \in W^{m,p}(\Omega)$. Sei $(\eta_\delta)_{\delta>0}$ eine Standard-Diracfolge. Dann hat $(\psi_k f)_\delta := \eta_\delta * (\psi_k f)$ für $0 < \delta < \frac{1}{(k+1)(k+2)}$ kompakten Träger in $V_k := \Omega_{k+2} \setminus \Omega_{k-2}$. Da $\overline{V_k} \subset \Omega$ kompakt enthalten ist, lässt sich nach Lemma 2.65 zu $\varepsilon > 0$ ein $0 < \delta_k < \frac{1}{(k+1)(k+2)}$ wählen mit

$$\|(\psi_k f)_{\delta_k} - \psi_k f\|_{W^{m,p}(\Omega)} = \|(\psi_k f)_{\delta_k} - \psi_k f\|_{W^{m,p}(V_k)} < \frac{\varepsilon}{2^k}.$$

Man definiere

$$g := \sum_{k\in\mathbb{N}} (\psi_k f)_{\delta_k}.$$

Da für jedes offene $\Omega' \subset \overline{\Omega'} \subset \Omega$ nur endlich viele Terme der Summe nicht verschwinden, folgt $g \in C^\infty(\Omega)$. Für $x \in \Omega_k$ gilt speziell

$$f(x) = \sum_{j=1}^{k} \psi_j(x) f(x) \qquad \text{sowie} \qquad g(x) = \sum_{j=1}^{k} (\psi_j f)_{\delta_j}(x).$$

Dies liefert die Abschätzung

$$\|g - f\|_{W^{m,p}(\Omega_k)} \le \sum_{j=1}^{k} \|(\psi_j f)_{\delta_j} - \psi_j f\|_{W^{m,p}(\Omega)} < \varepsilon$$

und die Ausschöpfung durch Ω_k mit dem Satz über monotone Konvergenz impliziert $g - f \in W^{m,p}(\Omega)$ mit $\|g - f\|_{W^{m,p}(\Omega)} < \varepsilon$. □

Bemerkung: (i) Für $p = \infty$ ist $H^{m,\infty}(\Omega)$ eine echte Teilmenge von $W^{m,\infty}(\Omega)$, denn zum Beispiel gehört die Funktion $f(x) = |x|$ zu $W^{1,\infty}((-1,1))$, jedoch gibt es für $\varepsilon \in (0, 1/2)$ keine Funktion $\phi \in C^1((-1,1))$ mit $\|\phi' - f'\|_\infty < \varepsilon$.

(ii) Im Allgemeinen ist $C_c^\infty(\Omega)$ nicht dicht in $W^{m,p}(\Omega)$ bzgl. der $W^{m,p}$-Norm für $p < \infty$, abgesehen von dem Fall $\Omega = \mathbb{R}^n$.

In manchen Situationen ist es sinnvoll den Raum der Sobolev-Funktionen mit Nullrandwerten zu betrachten, nämlich den Abschluss von $C_c^\infty(\Omega)$ bzgl. der $W^{m,p}$-Norm für $1 \le p < \infty$:

$$W_0^{m,p}(\Omega) := \{f \in W^{m,p}(\Omega) \mid \exists\, (f_k)_{k\in\mathbb{N}} \subset C_c^\infty(\Omega) \text{ mit } \|f - f_k\|_{m,p} \to 0 \text{ für } k \to \infty\}$$

$W_0^{m,p}(\Omega)$ ist per Definition ein abgeschlossener Unterraum von $W^{m,p}(\Omega)$. Ist $\Omega \ne \mathbb{R}^n$, so ist $W_0^{m,p}(\Omega)$ eine echte Teilmenge von $W^{m,p}(\Omega)$. Für $\Omega = \mathbb{R}^n$ erhält man

Korollar 2.68. *Es gilt $W_0^{m,p}(\mathbb{R}^n) = W^{m,p}(\mathbb{R}^n)$ für $m \in \mathbb{N}, 1 \le p < \infty$, d.h. $C_c^\infty(\mathbb{R}^n)$ liegt dicht in $W^{m,p}(\mathbb{R}^n)$ bzgl. der $W^{m,p}$-Norm.*

Beweis. Sei $f \in W^{m,p}(\mathbb{R}^n)$ für $m = 1$, der Fall $m > 1$ folgt induktiv. Wir konstruieren eine Folge $(f_k)_{k\in\mathbb{N}}$ in $C_c^\infty(\mathbb{R}^n)$. Wähle dazu nach Korollar 2.9 eine Abschneidefunktion $\psi \in C_c(B_2(0), [0,1])$ mit $\psi|_{\overline{B_1(0)}} = 1$ und setze $\psi_\varepsilon(x) = \psi(\varepsilon x)$. Für eine Nullfolge $(\varepsilon_k)_{k\in\mathbb{N}}$ wählen wir

$$f_k := \psi_{\varepsilon_k} \cdot f_{\varepsilon_k} = \psi_{\varepsilon_k} \cdot (\eta_{\varepsilon_k} * f) \in C_c^\infty(\mathbb{R}^n)$$

für eine Standard-Diracfolge $(\eta_\varepsilon)_{\varepsilon>0}$. Nach dem Satz über majorisierte Konvergenz folgt zunächst $\psi_\varepsilon f \to f$ in $L^p(\mathbb{R}^n)$ für $\varepsilon \to 0$ und somit wegen $|\psi_\varepsilon| \le 1$

$$\|f_k - f\|_p \le \|f_{\varepsilon_k} - f\|_p + \|\psi_{\varepsilon_k} f - f\|_p \to 0 \qquad (\text{für } k \to \infty)$$

nach Lemma 2.65. Für die schwache Ableitung verwenden wir die Differentiationsregeln und erhalten

$$\partial_i f_k = \partial_i \psi_{\varepsilon_k} \cdot f_{\varepsilon_k} + \psi_{\varepsilon_k} \cdot \partial_i f_{\varepsilon_k} = \partial_i \psi_{\varepsilon_k} \cdot f_{\varepsilon_k} + \psi_{\varepsilon_k} \cdot (\partial_i f)_{\varepsilon_k}.$$

Es folgt

$$\|\partial_i f_k - \partial_i f\|_p \le \|\partial_i \psi_{\varepsilon_k} \cdot f_{\varepsilon_k}\|_p + \|\psi_{\varepsilon_k} \cdot ((\partial_i f)_{\varepsilon_k} - \partial_i f)\|_p + \|\psi_{\varepsilon_k} \partial_i f - \partial_i f\|_p,$$

wobei die letzten beiden Terme gegen Null konvergieren für $k \to \infty$. Offensichtlich ist $\nabla\psi$ beschränkt und der erste Term lässt sich weiter abschätzen durch

$$\|\partial_i \psi_{\varepsilon_k} \cdot f_{\varepsilon_k}\|_p \leq \varepsilon_k \|\nabla\psi\|_\infty \|f_{\varepsilon_k}\|_p \leq \varepsilon_k \|\nabla\psi\|_\infty \|f\|_p \to 0 \qquad (\text{für } k \to \infty)$$

nach der Youngschen Ungleichung und die Aussage folgt. □

Bezüglich der L^p-Norm besitzen sowohl $C_c^\infty(\Omega)$ als auch $C^\infty(\Omega) \cap L^p(\Omega)$ dieselben Abschlüsse, nämlich $L^p(\Omega) = W_0^{0,p}(\Omega)$. Dies ist aber für Sobolev-Funktionen nicht mehr der Fall, da Ableitungen der approximierenden Folgen am Rand von Ω sehr groß werden können. Diese Probleme sind auch der Grund, warum der Beweis von Satz 2.67 wesentlich aufwendiger ist als der von Korollar 2.56.

Als weitere elementare Regeln für Sobolev-Funktionen haben wir

Lemma 2.69 (Kettenregel)**.** *Seien $\Omega \subset \mathbb{R}^n$ offen, $f \in W^{1,p}(\Omega)$ für $1 \leq p < \infty$ und $g \in C^1(\mathbb{R})$ mit beschränkter Ableitung $\|g'\|_\infty \leq C$. Dann ist $g \circ f \in L^p_{loc}(\Omega)$ schwach differenzierbar mit $\nabla(g \circ f) \in L^p(\Omega)^n$ und es gilt die Kettenregel*

$$\nabla(g \circ f) = (g' \circ f) \cdot \nabla f.$$

Ist zusätzlich $g(0) = 0$ oder $\lambda^n(\Omega) < \infty$, so folgt $g \circ f \in W^{1,p}(\Omega)$.

Beweis. Nach Analysis 1 ist g Lipschitz-stetig und es gilt $|g(x)| \leq |g(0)| + C|x|$ für alle $x \in \mathbb{R}$. Dies impliziert sofort $g \circ f \in L^p_{\text{loc}}(\Omega)$ bzw. zusätzlich $g \circ f \in L^p(\Omega)$, auch im Fall $p = \infty$. Da g' beschränkt ist, ist ebenso $g' \circ f \in L^\infty(\Omega)$, also $(g' \circ f) \cdot \nabla f \in L^p(\Omega)$.
Für die schwache Differenzierbarkeit sei nun $\psi \in C_c^\infty(\Omega)$ beliebig, dann gibt es eine offene Menge $U \supset \operatorname{supp}(\psi)$ mit kompaktem $\overline{U} \subset \Omega$, für welche nach Lemma 2.65 die Glättung $f_\varepsilon \in W^{1,p}(U) \cap C^\infty(U)$ für hinreichend kleines $\varepsilon > 0$ gegen f konvergiert mit

$$\|f_\varepsilon - f\|_{W^{1,p}(U)} \to 0 \qquad \text{für } \varepsilon \to 0.$$

Dann folgt mit der klassischen Kettenregel für $i = 1, \dots, n$

$$\int_\Omega (g \circ f_\varepsilon) \cdot \partial_i \varphi \, \mathrm{d}\lambda^n = -\int_\Omega (g' \circ f_\varepsilon) \partial_i f_\varepsilon \cdot \varphi \, \mathrm{d}\lambda^n.$$

Wegen $|g(f_\varepsilon) - g(f)| \leq C|f_\varepsilon - f|$ konvergiert die linke Seite gemäß

$$\|(g \circ f_\varepsilon) \cdot \partial_i \varphi - (g \circ f) \cdot \partial_i \varphi\|_{L^1(U)} \leq C \|\nabla\varphi\|_\infty \|f_\varepsilon - f\|_{L^p(U)},$$

denn $L^p(U) \subset L^1(U)$ mit $\lambda^n(U) < \infty$. Zur Herleitung des Grenzwerts der rechten Seite schätzen wir wie folgt ab

$$\begin{aligned}\|[(g' \circ f_\varepsilon)\partial_i f_\varepsilon - (g' \circ f)\partial_i f] \cdot \varphi\|_{L^1(U)} &\leq \|\varphi\|_\infty \|(g' \circ f_\varepsilon)\left(\partial_i f_\varepsilon - \partial_i f\right)\|_{L^1(U)} \\ &\quad + \|\varphi\|_\infty \|(g' \circ f_\varepsilon - g' \circ f)\partial_i f\|_{L^1(U)} \\ &\leq C\|\varphi\|_\infty \|\partial_i f_\varepsilon - \partial_i f\|_{L^p(U)} \\ &\quad + \|\varphi\|_\infty \|(g' \circ f_\varepsilon - g' \circ f)\partial_i f\|_{L^1(U)}.\end{aligned}$$

Da der erste Term gegen Null konvergiert, genügt es den zweiten zu betrachten. Da $f_\varepsilon \to f$ in $L^p(U)$ konvergiert, gibt es eine Teilfolge $(f_{\varepsilon_k})_{k\in\mathbb{N}}$, die punktweise λ^n-f.ü. in U konvergiert. Die Stetigkeit von g' impliziert die punktweise λ^n-f.ü. Konvergenz des Integranden gegen Null. Majorisierte Konvergenz liefert dann die Aussage, denn der Integrand ist beschränkt durch $2C|\partial_i f| \in L^1(U)$. □

Bemerkung: Falls $\lambda^n(\Omega) < \infty$ ist, gilt die Aussage auch für $p = \infty$, denn für alle $p < \infty$ ist $L^\infty(\Omega) \subset L^p(\Omega)$ und der Grenzübergang $p \to \infty$ liefert die Behauptung.

Wie auch für Lebesgue-integrierbare Funktionen gilt

Lemma 2.70. *Sei $\Omega \subset \mathbb{R}^n$ offen und $f \in W^{1,p}(\Omega)$ für $1 \leq p < \infty$. Dann sind auch $|f|$, $f^+ := \max\{f, 0\}$ und $f^- := \min\{f, 0\}$ in $W^{1,p}(\Omega)$ und es gilt*

$$\nabla f^+ = \begin{cases} \nabla f & \text{für} \quad f > 0, \\ 0 & \text{für} \quad f \leq 0 \end{cases} \qquad \text{und} \qquad \nabla f^- = \begin{cases} \nabla f & \text{für} \quad f < 0, \\ 0 & \text{für} \quad f \geq 0 \end{cases}$$

sowie

$$\nabla |f| = \begin{cases} \nabla f & \text{für} \quad f > 0, \\ 0 & \text{für} \quad f = 0, \\ -\nabla f & \text{für} \quad f < 0. \end{cases}$$

Beweis. Wegen $f^- = -(-f)^+$ und $|f| = f^+ - f^-$ genügt es die Aussage für f^+ zu zeigen. Wir wählen eine approximierende Folge $g_\varepsilon \in C^1(\mathbb{R})$ mit

$$g_\varepsilon(t) = \begin{cases} \sqrt{t^2 + \varepsilon^2} - \varepsilon, & t > 0, \\ 0 & t \leq 0 \end{cases}$$

für $\varepsilon > 0$. Dann ist $\|g_\varepsilon'\|_\infty \leq 1$ und nach der Kettenregel folgt $g_\varepsilon \circ f \in W^{1,p}(\Omega)$ mit

$$\nabla(g_\varepsilon \circ f) = g_\varepsilon'(f) \nabla f = \begin{cases} \frac{f \nabla f}{\sqrt{f^2 + \varepsilon^2}} & \text{für} \quad f > 0, \\ 0 & \text{für} \quad f \leq 0. \end{cases}$$

Da $g_\varepsilon(t) \to t^+$ für $\varepsilon \to 0$ und $t \in \mathbb{R}$ mit $g_\varepsilon(t) \leq t^+$ liefert majorisierte Konvergenz für jedes $\psi \in C_c^\infty(\Omega), i = 1, \ldots, n$

$$\int_\Omega \partial_i (g_\varepsilon \circ f) \psi \, \mathrm{d}\lambda^n = - \int_\Omega (g_\varepsilon \circ f) \partial_i \psi \, \mathrm{d}\lambda^n \to - \int_\Omega f^+ \partial_i \psi \, \mathrm{d}\lambda^n.$$

Analog folgt durch eine integrierbare Schranke von $\nabla(g_\varepsilon \circ f)$ für $\varepsilon \to 0$

$$\int_\Omega \partial_i (g_\varepsilon \circ f) \psi \, \mathrm{d}\lambda^n = \int_\Omega \partial_i \frac{f \partial_i f}{\sqrt{f^2 + \varepsilon^2}} \chi_{\{f > 0\}} \psi \, \mathrm{d}\lambda^n \to \int_\Omega \partial_i f \chi_{\{f>0\}} \psi \, \mathrm{d}\lambda^n.$$

□

Abschließend zeigen wir, dass auch der Hauptsatz der Differential- und Integralrechnung für Sobolev-Funktionen auf offenen, beschränkten Intervallen $I = (a, b) \subsetneq \mathbb{R}$ gültig bleibt und wir damit einen stetigen Repräsentanten auswählen können.

Satz 2.71. *Für $1 \leq p \leq \infty$ sei $f \in W^{1,p}(I)$ für ein offenes, beschränktes Intervall $I = (a, b) \subset \mathbb{R}$. Dann gibt es eine Funktion $g \in C(\overline{I})$ mit $g = f$ λ^n-f.ü. in I und*

$$g(x) - g(y) = \int_y^x f'(t) \, \mathrm{d}t \qquad \forall\, x, y \in \overline{I}.$$

Beweis. Für festes $x_0 \in I$ setze

$$h(x) := \int_{x_0}^{x} f'(t)\,\mathrm{d}t \qquad x \in I.$$

Dann lässt sich analog zur Höheren Analysis mit einigem Aufwand zeigen, dass $h \in C(I)$ gleichmäßig stetig ist und die Übungen liefern die schwache Differenzierbarkeit mit $h' = f'$. Insbesondere lässt sich damit h stetig auf $\overline{I}$ fortsetzen. Man erhält für alle $\psi \in C_c^\infty(I)$

$$\int_I h\psi'\,\mathrm{d}\lambda = -\int_I f'\psi\,\mathrm{d}\lambda = \int_I f\psi'\,\mathrm{d}\lambda.$$

Nach den Übungen ist dann aber bereits $f - h = C$ λ-f.ü. konstant in I. Dann liefert etwa $g = h + C$ die Aussage. $\square$

3 Lineare Operatoren

In vielen Anwendungen wie z.B. zur Lösung linearer partieller Differentialgleichungen tauchen linearen Abbildungen zwischen normierten Räumen auf, auch lineare Operatoren (manchmal auch nur Operatoren) genannt. Ist der Bildbereich ein skalarer Körper wie etwa $\mathbb{K}$, so spricht man von Funktional statt Operator.

3.1 Stetige Operatoren

Im Folgenden seien $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ stets zwei normierte $\mathbb{K}$-Vektorräume. Eine wichtige Klasse linearer Operatoren sind die stetigen bzw. beschränkten Operatoren.

Definition 3.1. Seien X, Y zwei normierte Räume. Ein Operator $T : X \to Y$ heißt **linear**, falls T eine lineare Abbildung zwischen diesen normierten Räumen ist. Ein **stetiger linearer Operator** $T : X \to Y$ ist eine lineare Abbildung, die zusätzlich stetig ist. Ein Operator heißt **beschränkt**, falls es eine Konstante $C > 0$ gibt mit

$$\|T(x)\|_Y \leq C\|x\|_X \qquad \forall\, x \in X.$$

Für lineare Operatoren schreibt man vereinfacht $T(x) = Tx$ für $x \in X$.

Satz 3.2. *Seien X, Y normierte Räume und $T : X \to Y$ ein linearer Operator. Dann sind die folgenden Aussagen äquivalent:*

(i) T ist stetig.

(ii) T ist stetig in $x = x_0$.

(iii) $\sup_{\|x\|_X \leq 1} \|Tx\|_Y < \infty$.

(iv) T ist beschränkt.

Beweis. Es ist klar, dass (ii) aus (i) folgt. Die Linearität liefert (iii) $\Rightarrow$ (iv).

(ii)$\Rightarrow$(iii): Für festes $\varepsilon > 0$ gibt es ein $\delta > 0$ mit $T(\overline{B_\delta(x_0)}) \subset \overline{B_\varepsilon(Tx_0)}$. Sei nun $x \in \overline{B_1(0)}$ beliebig, dann ist $x_0 + \delta x \in \overline{B_\delta(x_0)}$. Also ist

$$Tx_0 + \delta Tx = T(x_0 + \delta x) \in \overline{B_\varepsilon(Tx_0)}$$

und somit ist $Tx \in \overline{B_{\varepsilon/\delta}(0)}$ gleichmäßig beschränkt.

(iv)$\Rightarrow$(i): Wähle ein beliebiges $x_1 \in X$. Dann folgt

$$\|Tx - Tx_1\|_Y = \|T(x - x_1)\|_Y \leq C\|x - x_1\|_X \to 0 \qquad (\text{für } x \to x_1).$$

$\square$

Bemerkung: Ist X endlich dimensional, so ist jede lineare Abbildung $T : X \to Y$ stetig, denn wählt man $\{e_1, \ldots, e_n\}$ als Basis von X, so folgt aus der Darstellung $x = \sum_{i=1}^n x_i e_i$

$$\|Tx\|_Y = \left\|\sum_{i=1}^n x_i Te_i\right\|_Y \leq \left(\sum_{i=1}^n \|Te_i\|_Y\right) \max_{i=1,\ldots,n} |x_i| \leq C\|x\|_X,$$

wobei $\|\cdot\|_\infty$ mit $\|x\|_\infty = \max_{i=1,\ldots,n} |x_i|$ eine zu $\|\cdot\|_X$ äquivalente Norm auf X ist.

Definition 3.3. Seien X, Y normierte Räume. Dann bezeichnet $\mathscr{L}(X,Y)$ den Raum aller stetigen linearen Operatoren von X nach Y. $\mathscr{L}(X,Y)$ besitzt die Struktur eines $\mathbb{K}$-Vektorraums mit den Operationen:

$$\begin{aligned} (T+S)(x) &:= T(x) + S(x) \qquad && \forall\, T, S \in \mathscr{L}(X,Y),\ \forall\, x \in X, \\ (\lambda T) &:= \lambda T(x) && \forall\, T \in \mathscr{L}(X,Y),\ \forall\, \lambda \in \mathbb{K}. \end{aligned}$$

Für $T \in \mathscr{L}(X,Y)$ definiert man die nach Satz 3.2 endliche **Operatornorm**

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

Falls klar ist, welche Norm gemeint ist, schreibt man auch $\|T\| = \|T\|_{\mathscr{L}(X,Y)}$. Man macht sich leicht klar, dass $(\mathscr{L}(X,Y), \|\cdot\|)$ ein normierter Raum ist. Falls $X = Y$ ist, schreibt man vereinfacht $\mathscr{L}(X) \equiv \mathscr{L}(X,X)$ und kann man das Produkt für $T, S : X \to X$ durch $TS(x) := (T \circ S)(x) = T(S(x))$ für $x \in X$ definieren. Offensichtlich ist $\|ST\| \leq \|S\|\|T\|$.

Satz 3.4. *Es seien X ein normierter Raum und Y ein Banachraum. Dann ist $\mathscr{L}(X,Y)$ ein Banachraum.*

Beweis. Sei $(T_k)_{k\in\mathbb{N}}$ eine Cauchy-Folge in $\mathscr{L}(X,Y)$ und $x \in X$. Dann ist wegen

$$\|T_k x - T_l x\|_Y \leq \|T_k - T_l\| \cdot \|x\|_X$$

$(T_k x)_{k\in\mathbb{N}}$ eine Cauchy-Folge in Y. Da Y vollständig ist, existiert $Tx := \lim_{k\to\infty} T_k x$ und die Eigenschaften des Grenzwerts liefern die Linearität von $T : X \to Y$. Außerdem ist $T \in \mathscr{L}(X,Y)$, da für $x \in X$ gilt

$$\|Tx - T_k x\|_Y = \lim_{l\to\infty} \|T_k x - T_l x\|_Y \leq \liminf_{l\to\infty} \|T_k - T_l\| \cdot \|x\|_X,$$

d.h. $T - T_k \in \mathscr{L}(X,Y)$ für $k \in \mathbb{N}$. Schließlich ist

$$\|T - T_k\| \leq \liminf_{l\to\infty} \|T_k - T_l\| \to 0 \qquad (\text{für } k \to \infty),$$

weil $(T_k)_{k\in\mathbb{N}}$ eine Cauchy-Folge ist. □

Definition 3.5. Seien X, Y normierte Räume.

(i) Man definiert den (topologischen) **Dualraum** X^* von X durch $X^* := \mathscr{L}(X,\mathbb{K})$. Elemente aus X^* heißen stetige lineare **Funktionale** auf X. Da $(\mathbb{K}, |\cdot|)$ vollständig ist, ist auch X^* ein Banachraum.

(ii) Für $T \in \mathscr{L}(X,Y)$ heißt

$$\ker(T) := \{x \in X \mid Tx = 0\}$$

der **Kern** oder **Nullraum** von T. Wegen der Stetigkeit von T ist $\ker(T)$ abgeschlossen in X. Mit

$$\operatorname{ran}(T) := \{Tx \mid x \in X\}$$

bezeichnet man das **Bild** von T (*range*). $\operatorname{ran}(T)$ ist im Allgemeinen nicht abgeschlossen.

(iii) $T \in \mathscr{L}(X,Y)$ wird als eine **(stetige) Einbettung** von X nach Y bezeichnet, falls T injektiv ist, d.h. wenn $\ker(T) = \{0\}$ ist. Gilt zusätzlich $\|Tx\|_Y = \|x\|_X$ für alle $x \in X$, so heißt T **Isometrie**.

Beispiel 3.6.

(a) Sei $X = C([0,1])$, $Y = C^1([0,1])$. Für $f \in X, x \in [0,1]$ definiert

$$(Tf)(x) := \int_0^x f(y)\,\mathrm{d}y$$

einen linearen Operator $T \in \mathscr{L}(X,Y)$ mit abgeschlossenem Bild

$$\operatorname{ran}(T) = \{f \in Y \mid f(0) = 0\}.$$

Wird T als Operator in $\mathscr{L}(X)$ aufgefasst, so ist das Bild $\{f \in Y \mid f(0) = 0\}$ nicht abgeschlossen in X, da sich z.B. die Betragsfunktion gleichmäßig approximieren lässt mit Funktionen aus $\operatorname{ran}(T)$.

(b) Sei (Ω, Σ, μ) ein Maßraum und p, q konjugierte Hölder-Exponenten mit $1 \leq p, q \leq \infty$ und $\frac{1}{p} + \frac{1}{q} = 1$. Für $g \in L^q(\Omega, \mu)$ ist nach der Hölder-Ungleichung aus Satz 2.43

$$T_q f := \int_\Omega \overline{g} f\,\mathrm{d}\mu \qquad \forall\, f \in L^p(\Omega, \mu)$$

ein lineares Funktional $T_q \in (L^p(\Omega,\mu))^*$. Dies definiert insbesondere eine Einbettung $\phi : L^q(\Omega,\mu) \to (L^p(\Omega,\mu))^*$.

(c) Seien $\Omega \subset \mathbb{R}^n$ offen und für jeden Multiindex $\alpha \in \mathbb{N}_0^n$ mit $|\alpha| \leq m$ seien $a_\alpha : \Omega \to \mathbb{R}$ Funktionen. Dann definiert

$$(Tf)(x) := \sum_{|\alpha| \leq m} a_\alpha(x) \partial^\alpha f(x)$$

einen linearen **partiellen Differentialoperator** der Ordnung m. Es gilt

- $T \in \mathscr{L}(C^m(\overline{\Omega}), C(\overline{\Omega}))$, falls Ω beschränkt ist und $a_\alpha \in C(\overline{\Omega})$ für $|\alpha| \leq m$.
- $T \in \mathscr{L}(C^{m,\beta}(\overline{\Omega}), C^{0,\beta}(\overline{\Omega}))$ mit $\beta \in (0,1]$, falls Ω beschränkt ist mit Koeffizienten $a_\alpha \in C^{0,\beta}(\overline{\Omega})$ für $|\alpha| \leq m$.
- $T \in \mathscr{L}(W^{m,p}(\Omega), L^p(\Omega))$ mit $1 \leq p \leq \infty$, falls $a_\alpha \in L^\infty(\Omega)$ für $|\alpha| \leq m$.

Definition 3.7.

(i) Ein stetiger linearer Operator $P : X \to X$ heißt **Projektion**, wenn $P^2 := P \circ P = P$ gilt. Die Menge der Projektionen ist definiert durch

$$P(X) := \{P \in \mathscr{L}(X) \mid P^2 = P\}.$$

(ii) Ist $T \in \mathscr{L}(X,Y)$ bijektiv und gilt $T^{-1} \in \mathscr{L}(Y,X)$, dann heißt T **invertierbar** oder (stetiger linearer) **Isomorphismus**.

(iii) Für $T \in \mathscr{L}(X, Y)$ definiert man eine lineare Abbildung

$$T^* : Y^* \to X^* \qquad \text{mit} \qquad (T^* y^*)(x) := y^*(Tx) \qquad \forall\, x \in X, y^* \in Y^*.$$

T^* wird als **adjungierter (dualer) Operator** zu T bezeichnet. Wegen

$$|(T^* y^*)(x)| \leq \|y^*\|_{Y^*} \|Tx\|_Y \leq \|y^*\|_{Y^*} \|T\|\, \|x\|_X$$

ist $T^* \in \mathscr{L}(Y^*, X^*)$ mit $\|T^*\| \leq \|T\|$, denn

$$\|T^* y^*\|_{X^*} = \sup_{\|x\|_X \leq 1} |(T^* y^*)(x)| \leq \|y^*\|_{Y^*} \|T\|.$$

Projektionen waren bereits in Hilberträumen wie etwa in Korollar 1.48 präsent. Dass P in diesem Fall auch beschränkt ist, folgt aus der Lipschitz-Stetigkeit der Distanzfunktion. Im Folgenden wollen wir abschließend ein Resultat zur Invertierbarkeit von Operatoren geben, das es uns in manchen Fällen erlaubt explizit die Inverse zu bestimmen.

Lemma 3.8 (Neumann-Reihe)**.** *Sei X ein Banach-Raum und $T \in \mathscr{L}(X)$ mit*

$$\limsup_{m \to \infty} \|T^m\|^{1/m} < 1.$$

Dann ist $I - T$ bijektiv mit $(I - T)^{-1} \in \mathscr{L}(X)$ und es gilt

$$(I - T)^{-1} = \sum_{n=0}^{\infty} T^n,$$

wobei $T^0 := I$ die Identität auf X sei.

Beweis. Wähle $m \in \mathbb{N}$ und $\theta \in (0, 1)$ mit $\|T^n\| \leq \theta^n$ für alle $n \geq m$. Betrachtet man die Partialsummen $S_l = \sum_{n=0}^{l} T^n$. Dann gilt für $m \leq k < l$

$$\|S_l - S_k\| \leq \sum_{n=k+1}^{l} \|T^n\| \leq \sum_{n=k+1}^{\infty} \theta^n \to 0 \quad \text{für } k, l \to \infty.$$

Somit ist $(S_l)_{l \in \mathbb{N}}$ eine Cauchy-Folge und deswegen konvergent. Sei $S := \lim_{l \to \infty} S_l$. Dann ist wegen $I - T \in \mathscr{L}(X)$

$$(I - T)\, S = \lim_{l \to \infty} (I - T)\, S_l = \lim_{l \to \infty} (I - T^{l+1}) = I,$$

denn $\|T^{l+1}\| \leq \theta^{l+1} \to 0$ für $l \to \infty$. □

3.2 Dualräume

Ein wesentliches Konzept der Funktionalanalysis ist die Betrachtung von Dualräumen, um Eigenschaften der zugrundeliegenden normierten Räume zu erhalten. Wie bereits in Satz 3.4 festgestellt wurde, ist zu einem normierten $\mathbb{K}$-Vektorraum X der Dualraum X^* versehen mit der Norm

$$\|x^*\| = \sup_{\|x\|_X \leq 1} |x^*(x)| \qquad \forall\, x^* \in X^* = \mathscr{L}(X, \mathbb{K})$$

ein Banachraum. Nach Beispiel 3.6 ist auch für reellwertiges $g \in L^q(\Omega, \mu)$

$$T_q f := \int_\Omega g f \, \mathrm{d}\mu \qquad \forall \, f \in L^p(\Omega, \mu)$$

ein lineares Funktional $T_q \in (L^p(\Omega, \mu))^*$ (wobei im Folgenden stets reellwertige Funktionen betrachtet seien) und g lässt sich als Element von $(L^p(\Omega, \mu))^*$ auffassen. In Analogie zu Skalarprodukten bzw. Bilinearformen schreiben wir auch für die Anwendung des Funktionals

$$\langle T_q, f \rangle := T_q f = \int_\Omega g f \, \mathrm{d}\mu = \langle g, f \rangle$$

und bezeichnen dies als **duale Paarung** von $(f, T_q) \in (L^p(\Omega, \mu), (L^p(\Omega, \mu))^*)$. Aus der Hölder-Ungleichung folgt dann

$$|\langle T_q, f \rangle| \leq \|f\|_p \|g\|_q$$

bzw. $\|T_q\| \leq \|g\|_q$. Genauer lässt sich zeigen

Lemma 3.9. *Sei (Ω, Σ, μ) ein Maßraum und p, q konjugierte Hölder-Exponenten mit $1 \leq p, q \leq \infty$ und $\frac{1}{p} + \frac{1}{q} = 1$. Für $g \in L^q(\Omega, \mu)$ ist $\|T_q\| = \|g\|_q$, wobei T_q wie in Beispiel 3.6 definiert sei.*

Beweis. Für $q = 1$ und $g \in L^1(\Omega, \mu)$ definiere

$$f := \begin{cases} \frac{g}{|g|}, & g \neq 0, \\ 0, & g = 0. \end{cases}$$

Dann ist $f \in L^\infty(\Omega, \mu)$ mit $\|f\|_\infty = 1$ und

$$\langle T_1 g, f \rangle = \int_\Omega g f \, \mathrm{d}\mu = \int_\Omega |g| \, \mathrm{d}\mu = \|g\|_1,$$

sodass das Supremum angenommen wird durch $\|T_1\| = \|g\|_1$.
Die Fälle $q > 1$ werden in den Übungen bewiesen.

□

Um zu zeigen, dass T_q tatsächlich einen isometrischen Isomorphismus darstellt und wir $(L^p(\Omega, \mu))^* \cong L^q(\Omega, \mu)$ identifizieren können, genügt es daher die Surjektivität zu zeigen. Dies ist jedoch sehr aufwendig, sodass wir einige Lemmata benötigen. In Analogie zur Parallelogramm-Identität folgt

Lemma 3.10 (Hanner Ungleichungen)**.** *Sei (Ω, Σ, μ) ein Maßraum und $f, g \in L^p(\Omega, \mu)$.*

(i) Für $1 \leq p \leq 2$ gilt

$$\big|\|f\|_p + \|g\|_p\big|^p + \big|\|f\|_p - \|g\|_p\big|^p \leq \|f + g\|_p^p + \|f - g\|_p^p \tag{3.1}$$

und

$$\big|\|f + g\|_p + \|f - g\|_p\big|^p + \big|\|f + g\|_p - \|f - g\|_p\big|^p \leq 2^p \left(\|f\|_p^p + \|g\|_p^p\right). \tag{3.2}$$

(ii) *Für $2 \leq p < \infty$ gelten die umgekehrten Ungleichungen.*

Beweis. Offenbar folgt (3.2) aus Ungleichung (3.1), indem man f, g durch $f+g$ bzw. $f-g$ ersetzt. Die Fälle $p=1$ und $p=2$ sind offensichtlich durch Dreiecksungleichung und Parallelogramm-Identität. Man betrachte den Fall $1 < p < 2$ und wähle ohne Einschränkung $f, g \neq 0$ und

$$0 < r := \frac{\|g\|_p}{\|f\|_p} \leq 1,$$

da man sonst gemäß Symmetrie der Ungleichung auch $\|f\|_p \leq \|g\|_p$ wählen kann. Für $0 < r \leq 1$ setzt man

$$\begin{aligned} \alpha(r) &:= (1+r)^{p-1} + (1-r)^{p-1} \\ \beta(r) &:= r^{1-p}\left((1+r)^{p-1} - (1-r)^{p-1}\right). \end{aligned}$$

Aus den Übungen folgt die Abschätzung

$$|a|^p \alpha(r) + |b|^p \beta(r) \leq |a+b|^p + |a-b|^p \qquad \forall\, a, b \in \mathbb{R}$$

und die Wahl $a = f(x), b = g(x)$ sowie Integration über Ω liefert

$$\|f\|_p^p \,\alpha(r) + \|g\|_p^p \,\beta(r) \leq \|f+g\|_p^p + \|f-g\|_p^p.$$

Schließlich zeigt sich per Definition von r

$$(\|f\| + \|g\|)^p + (\|f\| - \|g\|)^p \leq \|f+g\|^p + \|f-g\|^p.$$

Im Fall $p > 2$ gelten die umgekehrten Ungleichungen

$$|a|^p \alpha(r) + |b|^p \beta(r) \geq |a+b|^p + |a-b|^p \qquad \forall\, a, b \in \mathbb{R}$$

und die Aussage folgt analog.

□

Aus diesen Ungleichungen folgen einige Eigenschaften der L^p-Räume. Unter anderem sind die L^p-Räume für $1 < p < \infty$ gleichmäßig konvex.

Definition 3.11. Ein normierter Raum $(X, \|\cdot\|)$ heißt **gleichmäßig konvex**, falls es für jedes $\varepsilon \in (0,2)$ ein $\delta > 0$ gibt, sodass

$$\|x\| = \|y\| = 1 \text{ und } \left\|\frac{x+y}{2}\right\| > 1 - \delta \quad \Rightarrow \quad \|x-y\| < \varepsilon.$$

Beispiel 3.12. Die Dreiecksungleichung impliziert, dass die abgeschlossene Einheitskugel $\{x \in X \mid \|x\|_X \leq 1\}$ in einem normierten Raum stets konvex ist.

(a) Prähilberträume sind aufgrund der Parallelogramm-Identität gleichmäßig konvex.

(b) Sei $\mathbb{R}^2$ mit $\|(x_1, x_2)\| = \max(|x_1|, |x_2|)$ versehen. Dann ist die Einheitskugel konvex, aber nicht gleichmäßig konvex.

(c) Die Räume $L^1(\Omega)$ und $L^\infty(\Omega)$ sind nicht gleichmäßig konvex.

Korollar 3.13. *Sei (Ω, Σ, μ) ein Maßraum. Dann ist $L^p(\Omega, \mu)$ für $1 < p < \infty$ gleichmäßig konvex.*

Beweis. In den Übungen zeigt man die folgenden Ungleichungen von Clarkson, wobei $q = p/(p-1)$ der duale Exponent sei. Für $p \in (1, 2)$ gilt

$$\|f + g\|_p^q + \|f - g\|_p^q \le 2\left(\|f\|_p^p + \|g\|_p^p\right)^{q-1}$$

und für $p \ge 2$ hat man

$$\|f + g\|_p^p + \|f - g\|_p^p \le 2^{p-1}\left(\|f\|_p^p + \|g\|_p^p\right).$$

Für $1 < p \le 2$ folgt damit für $\varepsilon > 0$ und $f, g \in L^p(\Omega, \mu)$ mit $\|f\|_p = \|g\|_p = 1$ und $\|f + g\| > 2(1 - \delta)$

$$\|f - g\|_p < \left(2^q - (2(1-\delta))^q\right)^{1/q} = 2\left(1 - (1-\delta)^q\right)^{1/q} = \varepsilon$$

für $\delta = 1 - (1 - (\varepsilon/2)^q)^{1/q} > 0$.
Die zweite Clarkson Ungleichung liefert analog für $\delta = 1 - (1 - (\varepsilon/2)^p)^{1/p} > 0$ die Ungleichung

$$\|f - g\|_p < \left(2^p - (2(1-\delta))^p\right)^{1/p} = 2\left(1 - (1-\delta)^p\right)^{1/p} = \varepsilon.$$

□

Ähnlich wie in Hilberträumen lassen sich Minimierungsaufgaben eindeutig lösen.

Satz 3.14. *Sei $(X, \|\cdot\|)$ ein gleichmäßig konvexer Banachraum und $A \subset X$ abgeschlossen, konvex, nicht-leer und $x \in X$. Dann existiert ein eindeutiges $x_0 \in A$, sodass*

$$\|x - x_0\| = \operatorname{dist}(x, A) = \inf_{a \in A} \|x - a\|.$$

Beweis. Siehe Übungen. □

Satz 3.15. *Seien $1 < p, q < \infty$ mit $\frac{1}{p} + \frac{1}{q} = 1$ dual und (Ω, Σ, μ) ein Maßraum. Dann ist die Abbildung*

$$\begin{aligned} \phi : L^q(\Omega, \mu) &\to \left(L^p(\Omega, \mu)\right)^* \\ g &\mapsto \phi_g \end{aligned} \quad \textit{mit} \quad \phi_g(f) = \int_\Omega gf \,\mathrm{d}\mu \quad \forall\, f \in L^p(\Omega, \mu)$$

ein linearer isometrischer Isomorphismus.

Beweis. Nach Lemma 3.9 bleibt zu zeigen, dass ϕ surjektiv ist. Sei $L \in \left(L^p(\Omega, \mu)\right)^*$ und

$$K := \{g \in L^p(\Omega, \mu) \mid Lg = 0\} = \ker(L).$$

K ist ein linearer und abgeschlossener Unterraum von $L^p(\Omega, \mu)$. O.B.d.A. nehme man $L \neq 0$ an, dann findet man ein $h \in L^p(\Omega, \mu) \setminus K$. Nach Satz 3.14 gibt es dann genau ein $f \in K$ mit

$$\|f - h\|_p = \inf_{g \in K} \|g - h\| > 0.$$

Also besitzt die reellwertige Funktion $t \mapsto \|f + tg - h\|_p^p$ für ein beliebiges $g \in K$ ein Minimum bei $t = 0$. Diese Funktion ist differenzierbar nach den Differentiationsregeln für Parameterintegrale mittels majorisierter Konvergenz und es ist $p > 1$

$$\frac{\mathrm{d}}{\mathrm{d}t} \int |f + tg - h|^p \, \mathrm{d}\mu = p \int_\Omega |f + tg - h|^{p-2} \cdot (f + tg - h) \cdot g \, \mathrm{d}\mu.$$

Also erhält man mit $t = 0$ und $u := |f - h|^{p-2}(f - h)$ gerade

$$0 = \int_\Omega gu \, \mathrm{d}\mu \qquad \forall\, g \in K.$$

Offensichtlich ist $u \in L^q(\Omega, \mu)$ mit $\|u\|_q^q = \|f - h\|_p^p > 0$. Für ein beliebiges $g \in L^p(\Omega, \mu)$ schreibe man

$$g = \frac{L(g)}{L(f - h)}(f - h) + \tilde{g}.$$

mit $L\tilde{g} = 0$, d.h. $\tilde{g} \in K$. Also folgt

$$\int_\Omega gu \, \mathrm{d}\mu = \int_\Omega \frac{L(g)}{L(f - h)}(f - h)u \, \mathrm{d}\mu = \|f - h\|_p^p \frac{L(g)}{L(f - h)} \qquad \forall\, g \in L^p(\Omega, \mu).$$

Das bedeutet, dass $L = \phi_\nu$ mit

$$\nu = \frac{L(f - h)}{\|f - h\|_{L^p}^p} u$$

ist. Dies zeigt die Surjektivität von ϕ. □

Satz 3.15 bedeutet, dass jedes stetige lineare Funktional L auf $L^p(\Omega, \mu)$ für $1 < p < \infty$ als Integraloperator geschrieben werden kann. Im Fall $p = 1$ ist obiger Beweis nicht durchführbar. Außerdem benötigt man einen σ-finiten Maßraum.

Satz 3.16. *Sei (Ω, Σ, μ) ein σ-finiter Maßraum. Dann ist die Abbildung*

$$\begin{aligned} \phi : L^\infty(\Omega, \mu) &\to \big(L^1(\Omega, \mu)\big)^* \\ g &\mapsto \phi_g \end{aligned} \quad \textit{mit} \quad \phi_g(f) = \int_\Omega gf \, \mathrm{d}\mu \quad \forall\, f \in L^1(\Omega, \mu)$$

ein linearer isometrischer Isomorphismus.

Beweis. Nach Lemma 3.9 genügt es wieder die Surjektivität nachzuweisen, d.h. dass jedes $L \in (L^1(\Omega, \mu))^*$ von der Form ϕ_g ist für ein $g \in L^\infty(\Omega, \mu)$.
Sei zunächst $\mu(\Omega) < \infty$, so folgt mit der stetigen Einbettung $L^2(\Omega, \mu) \to L^1(\Omega, \mu)$ aus Korollar 2.45 auch

$$\big(L^1(\Omega, \mu)\big)^* \subset \big(L^2(\Omega, \mu)\big)^* \cong L^2(\Omega, \mu)$$

nach eben Bewiesenem. Somit gibt es zu jedem $L \in (L^1(\Omega, \mu))^*$ eine Funktion $g \in L^2(\Omega, \mu)$ mit

$$\langle L, f \rangle = \int_\Omega gf \, \mathrm{d}\mu \qquad \forall\, f \in L^2(\Omega, \mu).$$

Wir zeigen, dass bereits $g \in L^\infty(\Omega, \mu)$ gilt:
Zu beliebigem $A \in \Sigma$ betrachte man die Funktion

$$g_A := \begin{cases} \frac{g}{|g|}\chi_A, & g \neq 0, \\ 0, & g = 0. \end{cases}$$

Dann ist $g_A \in L^\infty(\Omega, \mu)$ mit

$$\int\limits_A |g| \,\mathrm{d}\mu = \int\limits_\Omega g g_A \,\mathrm{d}\mu = \langle L, g_A \rangle \leq \|L\| \|g_A\|_1 = \|L\| \mu(A).$$

Nach den Übungen folgt dann bereits $\|g\|_\infty \leq \|L\|$, da $A \in \Sigma$ beliebig war. Es bleibt also die Fortsetzung auf $L^1(\Omega, \mu)$ zu zeigen:
Zu $f \in L^1(\Omega, \mu)$ wähle man die punktweise fast überall approximierende Folge

$$f_k(x) := \begin{cases} f(x), & \text{falls} \quad |f(x)| \leq k, \\ 0, & \text{sonst.} \end{cases}$$

Dann ist $f_k \in L^\infty(\Omega, \mu)$ und dominierte Konvergenz liefert $f_k \to f$ sowie $gf_k \to gf$ in $L^1(\Omega, \mu)$, d.h.

$$\langle L, f \rangle = \lim_{k \to \infty} \langle L, f_k \rangle = \lim_{k \to \infty} \int\limits_\Omega g f_k \,\mathrm{d}\mu = \int\limits_\Omega g f \,\mathrm{d}\mu.$$

Der Fall $\mu(\Omega) = \infty$ kann durch die Annahme, dass der Maßraum σ-endlich ist, behandelt werden. Das heißt zu Ω gibt es eine abzählbare Ausschöpfung durch Teilmengen Ω_k endlichen Maßes mit

$$\Omega_k \subset \Omega_{k+1} \subset \Omega, \quad \Omega = \bigcup_{k \in \mathbb{N}} \Omega_k, \quad \mu(\Omega_k) < \infty.$$

Der Raum $L^1(\Omega_k, \mu)$ kann auf natürliche Weise in $L^1(\Omega, \mu)$ (normerhaltend) eingebettet werden, indem man die Funktion außerhalb Ω_k durch Null fortsetzt. Damit folgt $(L^1(\Omega, \mu))^* \subset (L^1(\Omega_k, \mu))^*$ für jedes $k \in \mathbb{N}$. Das heißt zu jedem $L \in (L^1(\Omega, \mu))^*$ gibt es wegen $\mu(\Omega_k) < \infty$ genau eine Funktion $g_k \in L^\infty(\Omega_k, \mu)$ mit

$$\langle L, f \rangle = \int\limits_\Omega g_k f \,\mathrm{d}\mu \qquad \forall\, f \in L^1(\Omega_k, \mu)$$

sowie

$$\|g_k\|_{L^\infty(\Omega_k, \mu)} = \|L\|_{(L^1(\Omega_k, \mu))^*} \leq \|L\|.$$

Durch die Fortsetzung mit Null haben wir bereits als Integrationsgebiet ganz Ω gewählt. Außerdem folgt damit für $k \leq \ell$, dass g_ℓ eingeschränkt auf Ω_k mit g_k übereinstimmt. Das heißt $(g_k)_{k \in \mathbb{N}}$ ist eine monoton wachsende Folge, deren punktweise Grenzwert g existiert und wiederum messbar ist. Die Abschätzung $|g_k| \leq |g_\ell|$ impliziert dann $|g_k| \leq |g|$ sowie

$$\|g\|_{L^\infty(\Omega, \mu)} \leq \|L\|,$$

d.h. $g \in L^\infty(\Omega, \mu)$. Nun bleibt die Fortsetzbarkeit auf $L^1(\Omega, \mu)$ zu zeigen:
Zu $f \in L^1(\Omega, \mu)$ wähle man die approximierende Folge $f_k := f\chi_{\Omega_k}$, sodass nach dem bisherigen

$$\langle L, f_k \rangle = \int\limits_\Omega g_k f_k \,\mathrm{d}\mu = \int\limits_\Omega g_k f \,\mathrm{d}\mu$$

gilt. Dominierte Konvergenz liefert schließlich wieder die Behauptung

$$\langle L, f\rangle = \int_\Omega g f \, \mathrm{d}\mu \qquad \forall\, f \in L^1(\Omega,\mu).$$

□

Bemerkung: Nach Beispiel 3.6 und Lemma 3.9 ist zwar $\phi : L^1(\Omega,\mu) \to (L^\infty(\Omega,\mu))^*$ eine Isometrie, jedoch im Allgemeinen nicht surjektiv.

Analog zu den L^p-Räumen lassen sich weitere Dualräume mit oft erheblichem Aufwand charakterisieren. Beispielsweise lässt sich für eine kompakte Teilmenge $K \subset \mathbb{K}^n$ der Dualraum $(C_{\mathbb{K}}(K))^*$ durch reguläre Borel-Maße beschreiben. Einen Beweis findet man beispielsweise im Buch von H.W. Alt, Satz 4.22 (Satz von Riesz-Radon). Grundlegend hierfür ist der Satz von Hahn-Banach, welcher die Fortsetzbarkeit stetiger linearer Funktionale unter Normerhaltung garantiert.

Für den Beweis dieses Satzes ist das Lemma von Zorn nötig, weshalb wir kurz grundlegende Begriffe einer (teilweise) geordneten Menge wiederholen.

Definition 3.17. Eine **partielle Ordnung** auf einer Menge P ist eine Relation $\leq$ mit folgenden Eigenschaften für alle Elemente $a, b, c \in P$:

(P1) Aus $a \leq b$ und $b \leq a$ folgt $a = b$. (Antisymmetrie)

(P2) Es gilt $a \leq a$. (Reflexivität)

(P3) Aus $a \leq b$ und $b \leq c$ folgt $a \leq c$. (Transitivität)

Eine Teilmenge $M \subset P$ heißt **total geordnet**, falls für $a, b \in M$, $a \neq b$, entweder $a \leq b$ oder $b \leq a$ gilt. Ein Element $b \in P$ heißt **obere Schranke** für $M \subset P$, falls $a \leq b$ für alle $a \in M$ gilt. Ein Element $b \in P$ heißt **maximales Element** in P, falls kein $a \in P$ existiert mit $a \neq b$ und $b \leq a$.

Wir erinnern an das zum Auswahlaxiom äquivalente

Lemma 3.18 (Zorn). *Sei $P \neq \emptyset$ eine partiell geordnete Menge, sodass jede total geordnete Teilmenge von P eine obere Schranke besitzt. Dann enthält P mindestens ein maximales Element.*

Definition 3.19. Sei X ein $\mathbb{R}$-Vektorraum. Eine Abbildung $p : X \to \mathbb{R}$ heißt **sublineares Funktional**, falls für jedes $x, y \in X$ und alle $\lambda \in \mathbb{R}_+$ gilt:

(i) $p(\lambda x) = \lambda p(x)$.

(ii) $p(x + y) \leq p(x) + p(y)$.

Lemma 3.20. *Sei X ein $\mathbb{R}$-Vektorraum und $M \subsetneq X$ ein linearer Unterraum. Seien $p : X \to \mathbb{R}$ ein sublineares Funktional, $f : M \to \mathbb{R}$ linear und gelte $f(x) \leq p(x)$ für*

$x \in M$. Definiere für jedes $x_0 \in X \setminus M$ die Summe $\widetilde{M} := M \oplus \operatorname{span}\{x_0\}$, dann gibt es eine lineare Fortsetzung $F : \widetilde{M} \to \mathbb{R}$ mit

$$F(x) \leq p(x) \quad \forall\, x \in \widetilde{M} \qquad \textit{und} \qquad F|_M = f.$$

Beweis. Elemente in $\widetilde{M}$ besitzen die Gestalt $m + tx_0$ für $m \in M, t \in \mathbb{R}$. Für $c_0 \in \mathbb{R}$ machen wir den Ansatz $F : \widetilde{M} \to \mathbb{R}$ definiert durch

$$F(m + t\,x_0) := f(m) + t\,c_0 \qquad \forall\, m \in M, t \in \mathbb{R}.$$

Dann ist F linear, $F|_M = f$ und es bleibt p als Schranke von F zu identifizieren.
Für $y', y'' \in M$ gilt

$$\begin{aligned} f(y') - f(y'') = f(y' - y'') &\leq p(y' + x_0 - x_0 - y'') \\ &\leq p(y' + x_0) + p(-x_0 - y''). \end{aligned}$$

Setze

$$s_1 := \sup_{y'' \in M} \Big(- f(y'') - p(-x_0 - y'') \Big) \qquad \text{und} \qquad s_2 := \inf_{y' \in M} \Big(- f(y') + p(y' + x_0) \Big),$$

und wähle ein $c_0 \in [s_1, s_2]$. Damit folgt für $t > 0$

$$\begin{aligned} F(m + t\,x_0) = t \cdot F\left(\frac{m}{t} + x_0\right) &= t \cdot \left(f\left(\frac{m}{t}\right) + c_0\right) \\ &\leq t \cdot p\left(\frac{m}{t} + x_0\right) = p(m + t\,x_0) \end{aligned}$$

wegen $c_0 \leq s_2$. Andererseits erhält man für $t < 0$

$$\begin{aligned} F(m + t\,x_0) = -t \cdot F\left(-\frac{m}{t} - x_0\right) &= -t \cdot \left(-f\left(\frac{m}{t}\right) - c_0\right) \\ &\leq -t \cdot p\left(-\frac{m}{t} - x_0\right) = p(m + t\,x_0) \end{aligned}$$

wegen $c_0 \geq s_1$. Hieraus folgt $F(m + t\,x_0) \leq p(m + t\,x_0)$ für alle $m \in M, t \in \mathbb{R}$. □

Satz 3.21 (Hahn-Banach (reell)). *Sei X ein $\mathbb{R}$-Vektorraum, $M \subsetneq X$ ein linearer Unterraum und $f : M \to \mathbb{R}$ linear. Sei $p : X \to \mathbb{R}$ ein sublineares Funktional mit*

$$f(x) \leq p(x) \qquad \forall\, x \in M.$$

Dann existiert ein lineares Funktional $F : X \to \mathbb{R}$ mit

$$F|_M = f \qquad \textit{und} \qquad F(x) \leq p(x) \quad \forall\, x \in X.$$

Beweis. Im endlich-dimensionalen Fall folgt dies induktiv aus Lemma 3.20. Ansonsten betrachte man die Klasse der Fortsetzungen

$$\mathcal{F} := \left\{ (\widetilde{M}, \tilde{f}) \;\middle|\; \begin{array}{l} \widetilde{M} \subset X \text{ linearer Unterraum mit } M \subset \widetilde{M}, \\ \tilde{f} \text{ lineares Funktional auf } \widetilde{M} \text{ mit } \tilde{f}|_M = f \text{ und } f \leq p \text{ auf } \widetilde{M} \end{array} \right\}.$$

$\mathcal{F}$ ist nicht-leer wegen $(M, f) \in \mathcal{F}$. Auf $\mathcal{F}$ definiere man eine Halbordnung durch

$$(\widetilde{M}, \tilde{f}) \leq (\widetilde{N}, \tilde{g}) \;:\Leftrightarrow\; \quad \widetilde{M} \subset \widetilde{N} \quad \text{und} \quad \tilde{g}|_{\widetilde{M}} = \tilde{f}.$$

Wir zeigen nun, dass jede total geordnete Teilmenge von $\mathcal{F}$ eine obere Schranke besitzt. Ist $\mathcal{G} \subset \mathcal{F}$ eine total geordnete Menge von $\mathcal{F}$, so lässt sich die folgende Fortsetzung

$$\widetilde{M} := \bigcup_{(N,g)\in\mathcal{G}} N, \qquad \tilde{f}(x) := g(x), \text{ falls } x \in N \text{ mit } (N,g) \in \mathcal{G}$$

betrachten. Dabei ist $\tilde{f}$ wohldefiniert, denn für $x \in N_1 \cap N_2$, wobei $(N_1, g_1), (N_2, g_2) \in \widetilde{M}$ seien, folgt aus der totalen Ordnung bereits $g_1(x) = g_2(x)$. Ebenso übertragen sich die Eigenschaften $\tilde{f}|_M = f$ sowie $\tilde{f} \leq p$ auf $\widetilde{M}$. Es bleibt also die Linearität von $\widetilde{M}$ sowie $\tilde{f}$ zu zeigen:
Für gegebene $x, y \in \widetilde{M}$ und $\lambda \in \mathbb{R}$ existieren wegen der totalen Ordnung

$$(N_x, g_x), (N_y, g_y) \in \mathcal{G} \quad \text{mit} \quad x \in N_x, y \in N_y.$$

mit entweder $(N_x, g_x) \leq (N_y, g_y)$ oder $(N_y, g_y) \leq (N_x, g_x)$. Folglich ist entweder $x, y \in N_y$ oder $x, y \in N_x$, sodass auch $x + \lambda y \in \widetilde{M}$ gilt. Außerdem ist mit entweder $\xi = y$ oder $\xi = x$

$$\tilde{f}(x + \lambda\, y) = g_\xi(x + \lambda y) = g_\xi(x) + \lambda g_\xi(y) = \tilde{f}(x) + \lambda\, \tilde{f}(y).$$

Dann ist $(\widetilde{M}, \tilde{g}) \in \mathcal{F}$ eine obere Schranke für $\mathcal{G}$ und das Lemma von Zorn impliziert nun, dass ein maximales Element (N, F) in $\mathcal{F}$ existiert. Falls $N \neq X$ gilt, so wähle man ein $x_0 \in X \setminus N$ und wende Lemma 3.20 an. Das heißt es existiert ein H mit $(N + x_0\,\mathbb{R}, H) \in \mathcal{F}$, also ist $(N, F) \leq (N + x_0\,\mathbb{R}, H)$ und $(N, F) \neq (N + x_0\,\mathbb{R}, H)$. Dies steht im Widerspruch zur Maximalität von (N, F). Also ist $N = X$ und $F : X \to \mathbb{R}$ ist linear mit $F|_M = f$ und $F(x) \leq p(x)$ für alle $x \in X$. □

Besitzt p sogar die Eigenschaften einer Norm, erhält man mit einer kleinen Modifikation für den komplexen Fall

Satz 3.22 (Hahn-Banach für Funktionale)**.** *Sei $(X, \|\cdot\|)$ ein normierter $\mathbb{K}$-Vektorraum, $M \subset X$ ein linearer Unterraum und $f \in M^*$. Dann gibt es eine Fortsetzung*

$$F \in X^* \qquad \textit{mit} \qquad F|_M = f \quad \textit{und} \quad \|F\|_{X^*} = \|f\|_{M^*}.$$

Beweis. Sei zunächst $\mathbb{K} = \mathbb{R}$. Man definiere $p(x) := \|f\|_{M^*}\|x\|$. Dann ist p offenbar sublinear mit

$$f(x) \leq |f(x)| \leq \|f\|_{M^*}\|x\| = p(x) \qquad \forall\, x \in M.$$

Nach Satz 3.21 gibt es daher eine lineare Fortsetzung $F : X \to \mathbb{R}$ mit $F|_M = f$ und $F \leq p$ auf X. Damit ist auch

$$\pm F(x) = F(\pm x) \leq p(\pm x) = \|f\|_{M^*}\|x\| \qquad \forall\, x \in X,$$

d.h. $\|F\|_{X^*} \leq \|f\|_{M^*}$. Andererseits gilt

$$\|f\|_{M^*} = \sup_{\substack{x\in M,\\ \|x\|\leq 1}} |f(x)| = \sup_{\substack{x\in M,\\ \|x\|\leq 1}} |F(x)| \leq \sup_{\substack{x\in X,\\ \|x\|\leq 1}} |F(x)| = \|F\|_{X^*}.$$

Also folgt insgesamt $\|F\|_{X^*} = \|f\|_{M^*}$.
Im Fall $\mathbb{K} = \mathbb{C}$ fasse man X, M als $\mathbb{R}$-Vektorräume $X_\mathbb{R}, M_\mathbb{R}$ auf und betrachte $f_\mathbb{R} \in M_\mathbb{R}^*$ mit

$$f_\mathbb{R}(x) := \operatorname{Re}(f(x)) \qquad \forall\, x \in M_\mathbb{R}.$$

In der Tat ist $f_\mathbb{R}$ ein beschränktes lineares Funktional mit $\|f_\mathbb{R}\|_{M^*_\mathbb{R}} \leq \|f\|_{M^*}$. Die Linearität in $\mathbb{C}$ von f liefert

$$\operatorname{Im}(f(x)) = \operatorname{Im}(f(-\mathrm{i}^2 x)) = \operatorname{Im}(-\mathrm{i} f(\mathrm{i}x)) = -\operatorname{Re}(f(\mathrm{i}x)) = -f_\mathbb{R}(\mathrm{i}x)$$

und damit

$$f(x) = \operatorname{Re}(f(x)) + \operatorname{Im}(f(x)) = f_\mathbb{R} - \mathrm{i} f_\mathbb{R}(\mathrm{i}x).$$

Sei nun $F_\mathbb{R}$ die Fortsetzung von $f_\mathbb{R}$ auf $X_\mathbb{R}$ mit $\|F_\mathbb{R}\|_{X^*_\mathbb{R}} = \|f_\mathbb{R}\|_{M^*_\mathbb{R}}$. Definiere

$$F(x) := F_\mathbb{R}(x) - \mathrm{i} F_\mathbb{R}(\mathrm{i}x) \qquad \forall\, x \in X_\mathbb{R}.$$

Dann ist $F = f$ auf M und $F : X \to \mathbb{C}$ ist $\mathbb{C}$-linear, denn F ist weiterhin $\mathbb{R}$-linear und es gilt

$$F(\mathrm{i}x) = F_\mathbb{R}(\mathrm{i}x) - \mathrm{i} F_\mathbb{R}(-x) = \mathrm{i} F(x).$$

Die Fortsetzung liefert wiederum $\|f\|_{M^*} \leq \|F\|_{X^*}$. Für die umgekehrte Ungleichung wähle man zu $x \in X$ mit $F(x) \in \mathbb{C}$ Argument und Radius, sodass $F(x) = r\mathrm{e}^{\mathrm{i}\theta}$ ist. Dann folgt

$$|F(x)| = r = \operatorname{Re}(\mathrm{e}^{-\mathrm{i}\theta} F(x)) = \operatorname{Re}(F(\mathrm{e}^{-\mathrm{i}\theta} x)) = F_\mathbb{R}(\mathrm{e}^{-\mathrm{i}\theta} x) \leq \|F_\mathbb{R}\|_{X^*_\mathbb{R}} \|x\|$$

und die Aussage wegen $\|F_\mathbb{R}\|_{X^*_\mathbb{R}} = \|f_\mathbb{R}\|_{M^*_\mathbb{R}} \leq \|f\|_{M^*}$. □

Korollar 3.23. *Sei $(X, \|\cdot\|_X)$ ein normierter $\mathbb{K}$-Vektorraum, $Z \subset X$ ein linearer Unterraum und $y \in X \setminus Z$. Sei $d := \operatorname{dist}(y, Z) > 0$, dann gibt es ein $F \in X^*$ mit $\|F\| = 1$, $F|_Z = 0$ und $F(y) = d$.*

Beweis. Setze $M := Z \oplus y\,\mathbb{K}$ und definiere die lineare Funktion

$$f : M \to \mathbb{K} \qquad \text{mit} \quad f(z + \alpha\, y) := \alpha\, d \quad \forall\, \alpha \in \mathbb{K}.$$

Nun ist $f|_Z = 0$ und $f(y) = d$. Es bleibt $\|f\|_{M^*} = 1$ zu zeigen, da die Behauptung dann aus Satz 3.22 folgt.
Für $\alpha \in \mathbb{K} \setminus \{0\}$ und $z \in Z$ findet man

$$|f(z + \alpha\, y)| = |\alpha|\, d \leq |\alpha| \left\| y - \left(-\frac{z}{\alpha}\right) \right\|_X = \|\alpha\, y + z\|_X,$$

d.h. $\|f\|_{M^*} \leq 1$. Andererseits, sei $(z_n)_{n\in\mathbb{N}}$ eine Minimalfolge in Z mit $\|y - z_n\|_X \to d$ für $n \to \infty$. Dann gilt abschließend $\|f\|_{M^*} \geq 1$ gemäß

$$d = f(y - z_n) \leq \|f\|_{M^*} \|y - z_n\|_X \to \|f\|_{M^*} d \qquad (\text{für} \quad n \to \infty).$$

□

Korollar 3.24. *Sei $(X, \|\cdot\|_X)$ ein normierter $\mathbb{K}$-Vektorraum. Dann gelten*

(i) Für $x \in X \setminus \{0\}$ gibt es ein $f \in X^*$ mit $\|f\| = 1$ und $f(x) = \|x\|_X$.

(ii) X^* trennt die Punkte von X, d.h. für alle $x_1, x_2 \in X$ mit $x_1 \neq x_2$ gibt es ein $f \in X^*$, sodass $f(x_1) \neq f(x_2)$ gilt.

(iii) Ist $f(x) = 0$ für alle $f \in X^*$, so folgt $x = 0$.

(iv) Für alle $x \in X$ gilt

$$\|x\|_X = \sup_{\substack{f \in X^*, \\ \|f\|=1}} \langle f, x \rangle.$$

Beweis. Siehe Übungen. □

Als weitere Folgerung erhält man für den adjungierten Operator

Korollar 3.25. *Seien X, Y normierte Räume und $T \in \mathscr{L}(X,Y)$. Dann erfüllt der adjungierte Operator $T^* \in \mathscr{L}(Y^*, X^*)$ die Identität $\|T^*\| = \|T\|$.*

Beweis. Siehe Übungen. □

3.3 Lösungsmethoden in Hilberträumen

Bereits für den speziellen Hilbertraum $H = L^2(\Omega, \mu)$ haben wir gesehen, dass sich H^* mit H selbst identifizieren lässt. Diese Eigenschaft gilt generell in Hilberträumen wie der Darstellungssatz von Fréchet-Riesz zeigt:

Satz 3.26 (Riesz'scher Darstellungssatz)**.** *Sei $(H, (\cdot,\cdot))$ ein Hilbertraum. Die Abbildung $\phi : H \to H^*$ definiert durch*

$$(\phi(u))\,(v) := (u, v)$$

ist ein antilinearer (d.h. $\phi(\alpha x + y) = \overline{\alpha}\phi(x) + \phi(y)$) isometrischer Isomorphismus.

Beweis. Die Linearität des Skalarprodukts im 2. Argument sowie die Ungleichung von Cauchy-Schwarz implizieren $\phi(u) \in H^*$ für jedes $u \in H$ und $\phi(u)(u) = \|u\|^2$ liefert die Isometrie $\|\phi(u)\|_{H^*} = \|u\|_H$. Die Antilinearität folgt ebenfalls aus der Definition des Skalarprodukts. Folglich bleibt die Surjektivität zu zeigen.
Sei o.E. $L \in H^* \setminus \{0\}$ und definiere den linearen Unterraum

$$M = \ker L = \{u \in H \mid L(u) = 0\} \subset H.$$

Da L stetig ist, ist M abgeschlossen und man erhält mit Korollar 1.48 genau eine Projektion $P : H \to M$ mit $(x - P(x), y) = 0$ für alle $y \in M, x \in H$. Zu $L \neq 0$ wähle man ein $v_0 \in H$ mit $L(v_0) \neq 0$, o.E. sei $L(v_0) = 1$. Definiert man $v = v_0 - P(v_0)$, so ist $(v, y) = 0$ für alle $y \in M$ und wegen $P(v_0) \in M$ gilt $L(v) = L(v_0) = 1$. Nun schreiben wir für beliebiges $u \in H = M \oplus M^\perp$

$$u = u - L(u)v + L(u)v \qquad \text{mit} \quad u - L(u)v \in M$$

und es folgt

$$(v, u) = (v, u - L(u)v) + (v, L(u)v) = L(u)\|v\|^2$$

bzw. $L = \phi(v/\|v\|^2)$. □

Bemerkung: Aus dem Satz 3.26 folgt, dass H^* die Struktur eines Hilbertraums besitzt. Für $f, g \in H^*$ definiert man

$$(f, g)_{H^*} := (\phi^{-1}g, \phi^{-1}f)$$

und es ist $\|f\|_{H^*}^2 = \|\phi^{-1}f\|_H^2 = (f, f)_{H^*}$.

Als Folgerung bzw. Erweiterung auf allgemeinere Sesquilinearformen erhält den wichtigen

Satz 3.27 (Lax-Milgram). *Sei $(H, (\cdot,\cdot)_H)$ ein Hilbertraum über $\mathbb{K}$ und $a : H \times H \to \mathbb{K}$ eine Sesquilinearform (d.h. linear im zweiten und antilinear im ersten Argument). Es gebe Konstanten $0 < c_0 \le C_0 < \infty$ mit*

$$\begin{aligned} |a(x,y)| &\le C_0 \,\|x\|_H \,\|y\|_H \qquad && \forall\, x,y \in H, \qquad && \textbf{\textit{(Stetigkeit)}} \\ \operatorname{Re}(a(x,x)) &\ge c_0 \,\|x\|_H^2 && \forall\, x \in H. && \textbf{\textit{(Koerzivität)}} \end{aligned}$$

Dann existiert genau ein linearer Operator $A : H \to H$ mit

$$a(x,y) = (Ax,y)_H \qquad \forall\, x,y \in H.$$

Ferner ist $A \in \mathscr{L}(H)$ invertierbar mit $\|A\| \le C_0$ und $\|A^{-1}\| \le c_0^{-1}$.

Beweis. Für $x \in H$ ist die Abbildung $y \mapsto a(x,y)$ in H^* mit $\|a(x,\cdot)\| \le C_0\,\|x\|_H$. Nach Satz 3.26 existiert für ein beliebiges $x \in H$ genau ein $A(x) = \phi^{-1}a(x,\cdot) \in H$ mit

$$a(x,y) = (A(x),y) \qquad \forall\, y \in H \quad \text{und} \quad \|A(x)\|_H = \|a(x,\cdot)\| \le C_0\|x\|_H.$$

Weiterhin gilt für beliebige $y \in H$

$$\begin{aligned} (A(x_1+\lambda\, x_2),y) &= a(x_1+\lambda\, x_2,y) = a(x_1,y) + \overline{\lambda}\, a(x_2,y) \\ &= (A(x_1),y) + \overline{\lambda}\ (A(x_2),y) = (A(x_1) + \lambda\, A(x_2),y)\,, \end{aligned}$$

d.h. $A \in \mathscr{L}(H)$ mit $\|A\| \le C_0$.
Die Koerzivität impliziert die Injektivität von A, denn

$$c_0\|x\|^2 \le \operatorname{Re}(a(x,x)) = \operatorname{Re}(Ax,x) \le |(Ax,x)| \le \|Ax\|\,\|x\|.$$

Wir bemerken zunächst, dass $\operatorname{ran}(A) = \{Ax \mid x \in H\}$ ist, denn für ein Folge $(Ax_n)_{n\in\mathbb{N}}$ in A mit $Ax_n \to y$ in H ist $(Ax_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge und wegen $\|Ax\|_H \ge c_0\|x\|_H$ ist auch $(x_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in H. Somit gibt es ein $x \in H$, sodass $x_n \to x$ und aufgrund der Stetigkeit von A folgt $y = Ax$ und $y \in \operatorname{ran}(A)$.
Aus Satz 1.50 folgt

$$H = (\operatorname{ran}(A)) \oplus (\operatorname{ran}(A))^\perp\,.$$

Es bleibt der Fall $(\operatorname{ran}(A))^\perp \neq 0$ für die Surjektivität zu betrachten. Dann gibt es ein $x_0 \in (\operatorname{ran}(A))^\perp$, $x_0 \neq 0$, mit $(Ax_0,x_0) = 0$ im Widerspruch zur Koerzivität. Damit ist $\operatorname{ran}(A) = H$ und A ist bijektiv.
Somit ist A^{-1} wohldefiniert und ebenfalls linear. Weiterhin folgt aus $\|Ax\|_H \ge c_0\|x\|_H$ die Abschätzung $\|A^{-1}\| \le c_0^{-1}$, denn

$$c_0\|A^{-1}x\|_H \le \|A(A^{-1}x)\|_H = \|x\|_H \qquad \forall\, x \in H.$$

□

Eine Variante des eben bewiesenen Satzes von Lax und Milgram ist folgendes für die Anwendung wichtige

Korollar 3.28. *Sei $(H, (\cdot,\cdot)_H)$ ein Hilbertraum über $\mathbb{K}$ und $a : H \times H \to \mathbb{K}$ eine stetige, koerzive Sesquilinearform und $L \in H^*$ gegeben. Dann existiert genau eine Lösung $x \in H$ des Variationsproblems*

$$a(x,y) = L(y) \qquad \forall\, y \in H.$$

Außerdem gilt $\|x\|_H \leq c_0^{-1}\|L\|$. Falls a zusätzlich antisymmetrisch ist, d.h.

$$a(x,y) = \overline{a(y,x)} \qquad \forall\, x,y \in H,$$

so ist die Lösung des Variationsproblems auch das absolute Minimum des Funktionals

$$F(y) = \frac{1}{2}\, a(y,y) - \operatorname{Re} L(y).$$

Beweis. Mit den Notationen aus den vorherigen Sätzen ist

$$a(x,y) = (Ax,y)_H = (\phi(Ax))(y) \qquad \forall\, x,y \in H$$

und $\phi A : H \to H^*$ bijektiv. Daher ist $x := (\phi A)^{-1}L = A^{-1}\phi^{-1}L$ die eindeutige Lösung des Variationsproblems und es gilt

$$\|x\|_H = \|A^{-1}\phi^{-1}L\|_H \leq \frac{1}{c_0}\|\phi^{-1}L\|_H = \frac{1}{c_0}\|L\|.$$

Ist a ein Skalarprodukt, so folgt mit $\operatorname{Re} L(y-x) = \operatorname{Re}(a(x,y)) - a(x,x)$

$$F(y) - F(x) = \frac{1}{2}\,(a(y,y) + a(x,x)) - \operatorname{Re} a(x,y) = \frac{1}{2}\, a(x-y,x-y) \geq \frac{c_0}{2}\,\|x-y\|^2.$$

□

Als Anwendung des Satzes von Lax-Milgram betrachten wir sogenannte elliptische partielle Differentialgleichungen. Dazu sei $\Omega \subset \mathbb{R}^n$ offen und beschränkt und wir suchen reellwertige Funktionen $u \in C^2(\Omega)$ als Lösung der elliptischen Differentialgleichung

$$-\sum_{i=1}^{n} \partial_i \sum_{j=1}^{n} a_{ij}\partial_j u + bu = f \qquad \text{in} \quad \Omega,$$

wobei $b, f \in C(\Omega)$ und $a_{ij} \in C^1(\Omega)$ reellwertig vorgegeben seien. Des weiteren sei der obige Differentialoperator **gleichmäßig elliptisch**, d.h. es gebe ein $\gamma > 0$ mit

$$\sum_{i,j=1}^{n} a_{ij}(x) z_i z_j \geq \gamma\|z\|^2 \qquad \forall\, z \in \mathbb{R}^n, x \in \Omega. \qquad \text{(Positivitätsbedingung)}$$

Ein Beispiel hierfür ist die **Poisson-Gleichung**

$$-\Delta u = f$$

oder mit $f = 0$ auch **Laplace-Gleichung** genannt. Damit diese partielle Differentialgleichung eindeutig lösbar ist, benötigt man weitere Bedingungen, sogenannte **Randbedingungen**. Wir betrachten hier nur die gebräuchlichsten Varianten von **Randwertproblemen (RWP)**:

(i) **Dirichlet-RWP**: Man suche $u \in C^2(\Omega) \cap C(\overline{\Omega})$, sodass die Gleichung erfüllt ist und $u = g$ auf $\partial\Omega$ für ein gegebenes $g \in C(\partial\Omega)$ gilt.

(ii) **Neumann-RWP**: Man suche $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$, sodass die Gleichung erfüllt ist und

$$-\sum_{i=1}^{n} \nu_i \sum_{j=1}^{n} a_{ij} \cdot \partial_j u = g \quad \text{auf } \partial\Omega,$$

wobei $\nu = (\nu_1, \ldots, \nu_n)$ als äußere Normale auf $\partial\Omega$ und $g \in C(\partial\Omega)$ gegeben sind.

Lösungen $u \in C^2(\Omega) \cap C(\overline{\Omega})$ bzw. $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$ heißen auch **klassische Lösungen**, da sie im klassischen Sinne differenzierbar sind. Jedoch wird bereits anhand der Poisson-Gleichung klar, dass für eine solche Lösung f als stetig vorausgesetzt werden muss. In vielen Fällen in der Praxis gibt es jedoch auch unstetige Daten in einer partiellen Differentialgleichung und man muss den Lösungsbegriff abschwächen.

Sei hierzu $u \in C^2(\Omega) \cap C(\overline{\Omega})$ eine klassische Lösung der obigen elliptischen Differentialgleichung zu **homogenen Dirichlet-Randbedingungen**, d.h. sei $g = 0$, dann multiplizieren wir die Gleichung mit einer Testfunktion $\xi \in C_c^\infty(\Omega)$, integrieren über Ω und erhalten

$$\int_\Omega \left[-\sum_{i=1}^{n} \partial_i \sum_{j=1}^{n} a_{ij} \partial_j u + bu - f \right] \xi \, \mathrm{d}\lambda^n = 0 \qquad \forall\, \xi \in C_c^\infty(\Omega).$$

Hat Ω einen stückweise glatten Rand $\partial\Omega$, so folgt mit partieller Integration unter Beachtung von $\xi|_{\partial\Omega} = 0$

$$\int_\Omega \sum_{i=1}^{n} \partial_i \xi \sum_{j=1}^{n} a_{ij} \partial_j u + (bu - f)\xi \, \mathrm{d}\lambda^n = 0 \qquad \forall\, \xi \in C_c^\infty(\Omega).$$

In dieser variationellen Formulierung der Differentialgleichung lässt sich eine Lösung bereits im Sobolev-Raum $W^{1,1}(\Omega)$ definieren, sofern $a_{ij}, b \in L^\infty(\Omega)$ und $f \in L^1_{\mathrm{loc}}(\Omega)$ erfüllen. Dass eine Lösung $u \in W^{1,1}(\Omega)$ auch zumindest fast überall definierte Randwerte besitzt, liegt an dem Spuroperator $S : W^{1,1}(\Omega) \to L^1(\partial\Omega)$, welcher für stetige Funktionen $u \in W^{1,1}(\Omega) \cap C(\overline{\Omega})$ mit den tatsächlichen Werten von $u|_{\partial\Omega}$ übereinstimmt. Zur Realisierung der Nullrandwerte im Dirichlet-RWP verwendet man daher den Abschluss von $C_c^\infty(\Omega)$ bzgl. der $W^{1,1}$-Norm bzw. allgemeiner bezüglich der $W^{m,p}$-Norm (siehe Abschnitt 2.3)

$$W_0^{m,p}(\Omega) := \{u \in W^{m,p}(\Omega) \mid \exists\, (u_k)_{k\in\mathbb{N}} \subset C_c^\infty(\Omega) : \|u - u_k\|_{m,p} \to 0 \text{ für } k \to \infty\}$$

Unter gewissen Voraussetzungen zeigt sich, dass diese Funktionen tatsächlich Nullrandwerte im verallgemeinerten Sinn einer Spur besitzen. Da uns mit dem Satz von Lax-Milgram eine Methode in Hilberträumen zu Verfügung steht, betrachten wir $W_0^{1,2}(\Omega)$ und schreiben abkürzend in der üblicheren Schreibweise $H_0^1(\Omega)$.

Definition 3.29. Sei $\Omega \subset \mathbb{R}^n$ offen und beschränkt. Seien $a_{ij}, b \in L^\infty(\Omega)$ und $f \in L^2(\Omega)$, sodass die Positivitätsbedingung für a_{ij} λ^n-fast überall in Ω erfüllt ist. Eine Funktion $u \in H_0^1(\Omega)$ heißt **schwache Lösung** des homogenen Dirichlet-RWP, falls

$$\int_\Omega \sum_{i=1}^{n} \partial_i \xi \sum_{j=1}^{n} a_{ij} \partial_j u + bu\xi \, \mathrm{d}\lambda^n = \int_\Omega f\xi \, \mathrm{d}\lambda^n \qquad \forall\, \xi \in H_0^1(\Omega).$$

Aufgrund des Abschlusses genügt es, diese Gleichung für alle $\xi \in C_c^\infty(\Omega)$ zu überprüfen.

Existiert eine schwache Lösung $u \in C^2(\Omega) \cap C(\overline{\Omega})$, so lässt sich wiederum partiell integrieren für alle Testfunktionen $\xi \in C_c^\infty(\Omega)$ und man erhält die Differentialgleichung durch das Fundamentallemma 2.59 der Variationsrechnung. Stetige Nullrandwerte werden dann durch den Spuroperator übertragen und man erhält eine klassische Lösung.

Im Fall der homogenen Neumann-Randbedingung liefert partielle Integration ebenfalls eine Integralgleichung, wobei in diesem Fall mit Funktionen aus $C^\infty(\Omega)$ getestet wird. Die Randbedingung liefert dann

Definition 3.30. Sei $\Omega \subset \mathbb{R}^n$ offen und beschränkt. Seien $a_{ij}, b \in L^\infty(\Omega)$ und $f \in L^2(\Omega)$, sodass die Positivitätsbedingung für a_{ij} λ^n-fast überall in Ω erfüllt ist. Eine Funktion $u \in H^{1,2}(\Omega)$ heißt **schwache Lösung** des homogenen Neumann-RWP, falls

$$\int_\Omega \sum_{i=1}^n \partial_i \xi \sum_{j=1}^n a_{ij} \partial_j u + bu\xi \, \mathrm{d}\lambda^n = \int_\Omega f\xi \, \mathrm{d}\lambda^n \qquad \forall\, \xi \in H^{1,2}(\Omega).$$

Aufgrund des Abschlusses genügt es, diese Gleichung für alle $\xi \in C^\infty(\overline{\Omega})$ zu überprüfen (siehe Definition 2.60).

Man macht sich leicht klar, dass eine schwache Lösung $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$ des Neumann-RWP eine klassische Lösung ist, sofern das Normalenfeld ν auf $\partial\Omega$ hinreichend regulär ist, etwa C^1.

Um den Satz von Lax-Milgram für die Existenz und Eindeutigkeit von schwachen Lösungen zu benutzen, wird nun diese Integralformulierung verwendet. Man will die linke Seite als eine Bilinearform auf einem geeigneten Hilbtertraum betrachten und die rechte Seite als Linearform. Geeignete Hilberträume sind dabei $H_0^{1,2}(\Omega)$ bzw. $H^{1,2}(\Omega)$ und man erhält

Satz 3.31 (Existenz für schwaches Neumann-RWP)**.** *Sei $\Omega \subset \mathbb{R}^n$ offen und beschränkt. Seien $a_{ij} \in L^\infty(\Omega)$, sodass die Positivitätsbedingung für a_{ij} λ^n-fast überall in Ω erfüllt ist. Für $b \in L^\infty(\Omega)$ gebe es ein $b_0 > 0$ mit $b \geq b_0$ λ^n-f.ü. in Ω und sei $f \in L^2(\Omega)$.*
Dann existiert genau eine Lösung $u \in H^{1,2}(\Omega)$ des schwachen Neumann-Problems und es gibt ein von f, u unabhängiges $c > 0$ mit

$$\|u\|_{1,2} \leq c\,\|f\|_2.$$

Beweis. Für $u, v \in H^{1,2}(\Omega)$ definiere

$$a(u,v) := \sum_{i,j=1}^n \int_\Omega a_{ij}\, \partial_i u\, \partial_j v + b\, u\, v \, \mathrm{d}\lambda^n,$$

dann ist a ist linear und wegen $\mathbb{K} = \mathbb{R}$ sesquilinear. Nun ist a auch stetig, denn

$$\begin{aligned} |a(u,v)| &\leq \sum_{i,j=1}^n \|a_{ij}\|_\infty \|\partial_i u\|_2 \|\partial_j v\|_2 + \|b\|_\infty \|u\|_2 \|v\|_2 \\ &\leq \left(\sum_{i,j=1}^n \|a_{ij}\|_\infty + \|b\|_\infty \right) \cdot \|u\|_{1,2} \|v\|_{1,2}. \end{aligned}$$

a ist auch koerziv mit $c_0 = \min\{\gamma, b_0\}$:

$$a(u,u) \geq \gamma \int\limits_\Omega \sum_{j=1}^n |\partial_j u|^2 \, \mathrm{d}\lambda^n + b_0 \int\limits_\Omega |u|^2 \, \mathrm{d}\lambda^n \geq c_0 \cdot \|u\|_{1,2}^2.$$

Für $v \in H^{1,2}(\Omega)$ definiere man das lineare Funktional L mit

$$L(v) = \int\limits_\Omega f \, v \, \mathrm{d}\lambda^n.$$

Dann ist $L \in (H^{1,2}(\Omega))^*$, denn $\|L\| \leq \|f\|_2 < \infty$ wegen

$$|L(v)| \leq \|f\|_2 \|v\|_{1,2}.$$

Damit besitzt das Variationsproblem

$$a(u,v) = F(v) \qquad \forall\, v \in H^{1,2}(\Omega)$$

genau eine Lösung $u \in H^{1,2}(\Omega)$ mit

$$\|u\|_{1,2} \leq \frac{1}{c_0}\|L\| \leq \frac{1}{c_0}\|f\|_2.$$

□

Für das schwache Dirichlet-Problem erhält man dasselbe Resultat. Wobei sich das Resultat mit der in $H_0^1(\Omega)$ gültigen Poincaré-Ungleichung auf $b \geq 0$ verallgemeinern lässt (siehe Übungen).
Dass schwache Lösungen hinreichend regulär und damit wiederum klassische Lösungen sind, ist nicht leicht zu zeigen. Damit beschäftigt sich die Regularitätstheorie für elliptische Gleichungen, welche wir in den Vorlesungen über partielle Differentialgleichungen weiter verfolgen werden.

3.4 Hauptsätze für Operatoren

In diesem Abschnitt finden sich fundamentale Sätze der linearen Funktionalanalysis, unter anderem auch das Prinzip der gleichmäßigen Beschränktheit, welches aus der punktweisen Beschränktheit einer Familie von linearen Operatoren bereits Beschränktheit bzgl. der Operatornorm folgert.

Lemma 3.32. *Sei $(X, \mathcal{T})$ ein topologischer Raum. Folgende Aussagen sind äquivalent:*

(i) Es sei $(A_i)_{i\in\mathbb{N}}$ eine Folge abgeschlossener Mengen. Falls das Innere von A_i leer ist für alle $i \in \mathbb{N}$, so ist auch das Innere von $\bigcup\limits_{i\in\mathbb{N}} A_i$ leer.

(ii) Es sei $(B_i)_{i\in\mathbb{N}}$ eine Folge offener Mengen. Falls B_i dicht in X ist für jedes $i \in \mathbb{N}$, so ist auch $\bigcap\limits_{i\in\mathbb{N}} B_i$ dicht in X.

Beweis. Nach den Übungen gilt $X \setminus \overline{A} = (X \setminus A)^\circ$, sodass die Aussage direkt folgt, denn eine Menge ist dicht genau dann wenn das Komplement ein leeres Inneres besitzt. □

Definition 3.33. Ein topologischer Raum $(X, \mathcal{T})$ heißt **Baire'scher Raum**, falls eine der äquivalenten Bedingungen aus Lemma 3.32 erfüllt ist.

Satz 3.34 (Baire)**.** *Jeder vollständige metrische Raum ist ein Baire'scher Raum.*

Beweis. Sei (X, d) ein vollständiger metrischer Raum mit $X \neq \emptyset$. Sei $(B_i)_{i\in\mathbb{N}}$ eine Folge offener und dichter Mengen. Wir zeigen $\overline{B} = X$ für

$$B := \bigcap_{i\in\mathbb{N}} B_i.$$

Dazu zeige man, dass für eine beliebige nicht-leere offene Menge $G \subset X$ der Schnitt $G \cap B$ nicht-leer ist. Da B_1 offen und dicht in X ist, ist $G \cap B_1 \neq \emptyset$ und offen. Deswegen kann man ein $\varepsilon_1 \in (0, 1]$ und $x_1 \in X$ finden mit $\overline{B_{\varepsilon_1}(x_1)} \subset B_1 \cap G$.
Da B_2 offen und dicht in X ist, ist $B_{\varepsilon_1}(x_1) \cap B_2 \neq \emptyset$ und offen. Also findet man ein $\varepsilon_2 \in (0, \frac{1}{2}]$ und $x_2 \in X$ mit $\overline{B_{\varepsilon_2}(x_2)} \subset B_{\varepsilon_1}(x_1) \cap B_2$. Iterativ findet man $\varepsilon_n \in (0, \frac{1}{n}]$ und $x_n \in X$ mit

$$\overline{B_{\varepsilon_n}(x_n)} \subset B_{\varepsilon_{n-1}}(x_{n-1}) \cap B_n \qquad \forall\, n \in \mathbb{N}, n \geq 2.$$

Aufgrund der Schachtelung ist $(x_n)_{n\in\mathbb{N}}$ ist eine Cauchy-Folge in X, denn für alle $m \geq n$ ist

$$d(x_n, x_m) \leq \varepsilon_n \leq \frac{1}{n} \to 0 \qquad \text{für} \quad m, n \to \infty.$$

Da X vollständig ist, existiert der Limes $x = \lim_{n\to\infty} x_n$ und wegen $x_k \in \overline{B_{\varepsilon_n}(x_n)}$ für alle $k \geq n$ gilt dies auch für den Limes. Per Konstruktion ist folglich $x \in G \cap \bigcap_{i\in\mathbb{N}} B_i \neq \emptyset$. □

Satz 3.35 (Banach-Steinhaus)**.** *Sei $(X, \|\cdot\|_X)$ ein Banachraum, $(Y, \|\cdot\|_Y)$ ein normierter Raum und $\mathcal{F} \subset \mathscr{L}(X, Y)$, sodass für jedes $x \in X$ eine Schranke $C_x > 0$ existiert mit $\sup_{T\in\mathcal{F}} \|Tx\|_Y \leq C_x$. Dann gilt bereits $\sup_{T\in\mathcal{F}} \|T\| < \infty$.*

Beweis. Für $k \in \mathbb{N}$ setze man

$$A_k := \{x \in X \mid \|Tx\|_Y \leq k \;\; \forall\, T \in \mathcal{F}\}.$$

Dann ist nach Voraussetzung

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

Außerdem ist A_k für jedes $k \in \mathbb{N}$ abgeschlossen, denn für eine Folge $(x_j)_{j\in\mathbb{N}}$ in A_k mit $x_j \to x$ für $j \to \infty$ gilt für beliebiges $j \in \mathbb{N}$:

$$\|Tx\|_Y \leq \|T(x_j)\|_Y + \|T(x - x_j)\|_Y \leq k + \|T\| \cdot \|x - x_j\|_X \quad \forall\, T \in \mathcal{F}$$

Für ein festes $T \in \mathcal{F}$ und beliebiges $\varepsilon > 0$ wähle man also j so groß, dass $\|x - x_j\|_X \leq \frac{\varepsilon}{\|T\|}$. Dann ist $\|Tx\| \leq k + \varepsilon$ und für $\varepsilon \to 0$ erhält man $\|Tx\| \leq k$. Da die obige Abschätzung aber für beliebige $T \in \mathcal{F}$ gilt, ist auch $\sup_{T\in\mathcal{F}} \|Tx\| \leq k$, d.h. $x \in A_k$.
Nach Satz 3.34 gibt es wegen $\mathring{X} = X$ ein $k_0 \in \mathbb{N}$ mit $\mathring{A}_{k_0} \neq \emptyset$. Man findet also ein $x_0 \in X$ und $\varepsilon_0 > 0$ mit $\overline{B_{\varepsilon_0}(x_0)} \subset \mathring{A}_{k_0}$. Folglich hat man für beliebige $y \in X$ mit $\|y\| \leq \varepsilon_0$

$$\|Ty\|_Y = \|T(y + x_0) - Tx_0\|_Y \leq \|T(y + x_0)\|_Y + \|Tx_0\|_Y \leq k_0 + C_{x_0} \qquad \forall\, T \in \mathcal{F}$$

und schließlich
$$\sup_{\|y\|_X \le 1} \|Ty\|_Y \le \frac{k + C_{x_0}}{\varepsilon_0} \qquad \forall\, T \in \mathcal{F}.$$
□

Definition 3.36. Seien X, Y topologische Räume, so heißt eine Abbildung $f : X \to Y$ **offen**, falls gilt
$$U \text{ offen in } X \qquad \Rightarrow \qquad f(U) \text{ offen in } Y.$$

Satz 3.37 (Satz von der offenen Abbildung)**.** *Seien X, Y Banachräume. Dann gilt für $T \in \mathscr{L}(X, Y)$*
$$T \textit{ surjektiv} \qquad \Leftrightarrow \qquad T \textit{ offen.}$$

Beweis. Sei zunächst T offen, dann gibt es ein $\delta > 0$ mit $B_\delta^Y(0) \subset T(B_1^X(0))$ und die Linearität von T liefert für beliebiges $r > 0$ auch $B_r^Y(0) \subset T(B_{r/\delta}^X(0))$, sodass T surjektiv ist. Genauer gilt
$$T \text{ offen} \qquad \Leftrightarrow \qquad \exists\, \delta > 0 : \quad B_\delta^Y(0) \subset T(B_1^X(0)),$$
wie man sich leicht mit einem Skalierungsargument klar macht.
Ist nun T surjektiv, so genügt es ein $r > 0$ zu finden mit $B_r^Y(0) \subset T(B_1^X(0))$. Wegen
$$Y = T(X) = \bigcup_{n \in \mathbb{N}} \overline{T(B_n^X(0))}$$
existiert nach dem Baire'schen Kategoriensatz 3.34 ein $n \in \mathbb{N}$, sodass $\overline{T(B_n^X(0))}$ ein nichtleeres Inneres besitzt. Daher gibt es ein $\varepsilon > 0$ und $y_0 \in Y$ gibt mit
$$B_\varepsilon^Y(y_0) \subset \overline{T(B_n^X(0))}.$$
Damit gibt es zu jedem $y \in B_\varepsilon^Y(0)$ eine Folge $(x_i)_{i \in \mathbb{N}} \subset B_n^X(0)$ mit $Tx_i \to y + y_0$ für $i \to \infty$. Nun existiert ein $\delta > 0$ mit
$$B_\delta^Y(0) \subset \overline{T(B_1^X(0))}.$$
In der Tat, wählt man ein $x_0 \in X$ mit $Tx_0 = y_0$, so liefert die Linearität eine Folge in $B_1^X(0)$ mit
$$T\left(\frac{x_i - x_0}{n + \|x_0\|_X}\right) \to \frac{y}{n + \|x_0\|_X} \qquad \text{für } i \to \infty.$$
Nun lässt sich $\delta = \varepsilon/(n + \|x_0\|_X)$ wählen.
Wähle nun zu $y_0 \in B_\delta^Y(0)$ ein approximierendes $x_0 \in B_1^X(0)$, sodass $y_0 - Tx_0 \in B_{\delta/2}^Y(0)$. Damit ist $y_1 := 2(y_0 - Tx_0) \in B_\delta^Y(0)$. Fährt man analog mit y_1 fort, so erhält man induktiv Punkte $y_k \in B_\delta^Y(0)$ und $x_k \in B_1^X(0)$ mit
$$y_{k+1} = 2(y_k - Tx_k).$$
Zum einen ist $T(2^{-k}x_k) = 2^{-k}y_k - 2^{-(k+1)}y_{k+1}$ und eine Teleskopsumme liefert
$$T\left(\sum_{k=0}^{m} 2^{-k}x_k\right) = y_0 - 2^{-(m+1)}y_{m+1} \to y_0 \qquad (\text{für} \quad m \to \infty).$$

Andererseits sind die Partialsummen

$$\sum_{k=0}^{m} \|2^{-k} x_k\|_X \leq \sum_{k=0}^{\infty} 2^{-k} = 2$$

konvergent bzw. die Reihe

$$x := \sum_{k=0}^{\infty} 2^{-k} x_k \in X$$

absolut konvergent in X und, da X ein Banachraum ist, auch konvergent (siehe Übungen). Damit folgt $\|x\|_X \leq 2$ und die Stetigkeit von T impliziert $Tx = y_0$. Dies zeigt, dass $B_\delta^Y(0) \subset T(\overline{B_2^X(0)})$ bzw. $B_{\delta/3}^Y(0) \subset T(B_1^X(0))$ gilt. □

Satz 3.38 (Satz von der inversen Abbildung)**.** *Seien X, Y Banachräume und $T : X \to Y$ eine lineare, stetige Bijektion. Dann ist $T^{-1} \in \mathscr{L}(Y, X)$.*

Beweis. T^{-1} ist wohldefiniert und linear. Nach Satz 3.37 ist T offen, d.h. T^{-1} stetig und damit auch beschränkt. □

Korollar 3.39. *Sei X ein $\mathbb{K}$-Vektorraum und $(X, \|\cdot\|_1)$ und $(X, \|\cdot\|_2)$ Banachräume über X. Es existiere ein $c > 0$, sodass $\|x\|_1 \leq c\,\|x\|_2$ für alle $x \in X$ gilt. Dann gibt es ein $d > 0$ mit*

$$\|x\|_2 \leq d\,\|x\|_1 \qquad \forall\, x \in X,$$

d.h. die zwei Normen sind äquivalent.

Beweis. Aus der Annahme folgt, dass die Identität $\mathrm{id} : (X, \|\cdot\|_2) \to (X, \|\cdot\|_1)$ eine stetige bijektive Abbildung ist. Nach Satz 3.38 ist dann $\mathrm{id}^{-1} : (X, \|\cdot\|_1) \to (X, \|\cdot\|_2)$ stetig und beschränkt. □

Satz 3.40 (Satz vom abgeschlossenen Graphen)**.** *Seien X, Y Banachräume und ein linearer Operator $T : X \to Y$ gegeben. Dann sind folgende Aussagen äquivalent:*

(i) $\mathrm{Graph}\,(T) := \{(x, Tx) \in X \times Y \mid x \in X\}$ *ist abgeschlossen in $X \times Y$ bzgl. der Graphennorm $\|(x, y)\|_{X\times Y} = \|x\|_X + \|y\|_Y$.*

(ii) T ist stetig.

Beweis. Zunächst ist $(X \times Y, \|\cdot\|_{X\times Y})$ wiederum ein Banachraum, siehe Beispiel 1.31.

$\Leftarrow$: Sei $(x, y) \in \overline{\mathrm{Graph}\,(T)}$. Dann existiert eine Folge $(x_k)_{k\in\mathbb{N}}$ in X mit $(x_k, Tx_k) \to (x, y)$ für $k \to \infty$. Also konvergieren $x_k \to x$ in X und $Tx_k \to y$ in Y. Da T stetig ist, muss $y = Tx$ sein und $(x, y) \in \mathrm{Graph}\,(T)$. Damit ist $\mathrm{Graph}\,(T)$ abgeschlossen.

$\Rightarrow$: Man definiere die Abbildung

$$\phi : X \to X \times Y, \qquad \phi(x) := (x, T(x)).$$

Das Bild von ϕ ist $\mathrm{ran}\,(\phi) = \mathrm{Graph}\,(T)$ und als abgeschlossene Teilmenge von $X \times Y$ ein Banachraum bzgl. der Norm $\|\cdot\|_{X\times Y}$. Daher ist ϕ eine lineare Bijektion zwischen zwei Banachräumen und somit auch $\phi^{-1} : \mathrm{ran}\,(\phi) \to X$ mit $\phi^{-1}(x, Tx) = x$. Wegen

$$\|\phi^{-1}(x, Tx)\|_X = \|x\|_X \leq \|(x, Tx)\|_{X\times Y}$$

ist ϕ^{-1} eine stetige lineare Bijektion und Satz 3.38 liefert die Stetigkeit von ϕ. Da auch die Projektion $P : X \times Y \to Y$ definiert durch $P(x, y) = y$ stetig ist, folgt die Stetigkeit von $T = P \circ \phi$.

□

Definition 3.41 (Abgeschlossener Operator)**.** Seien X, Y normierte Räume. Eine lineare Abbildung $T : X \to Y$ heißt **abgeschlossen**, falls Graph (T) abgeschlossen in $X \times Y$ ist.

Beispiel 3.42. (a) Für offenes und beschränktes $\Omega \subset \mathbb{R}^n$, $1 \leq p < \infty$, sind auf $W_0^{1,p}(\Omega)$ nach der Poincaré-Ungleichung die Normen $\|\cdot\|_{1,p}$ und $\|\nabla \cdot\|_p$ äquivalent.

(b) Seien $X = C^1([0, 1])$ und $Y = C^0([0, 1])$ mit der Norm $\|\cdot\|_\infty$ ausgestattet und

$$T : X \to Y, \quad Tf = f'.$$

Dann ist der Ableitungsoperator T abgeschlossen, aber nicht stetig und $(X, \|\cdot\|_\infty)$ ist nicht vollständig.

(c) Betrachtet man den Ableitungsoperator aus (b) mit

$$T : (X, \|\cdot\|_\infty) \to (L^2([0, 1]), \|\cdot\|_2), \quad Tf = f',$$

so ist dieser nicht abgeschlossen. Verwendet man jedoch $(W^{1,2}([0, 1]), \|\cdot\|_{1,2})$ anstatt $(X, \|\cdot\|_\infty)$, ist T als schwacher Ableitungsoperator abgeschlossen.

Bemerkung: Oft betrachtet man auch Operatoren T, welche nur auf einem Untervektorraum $D \subset X$ definiert sind und schreibt dann für diesen Definitionsbereich $D = \operatorname{dom}(T)$ bzw. $T : \operatorname{dom}(T) \subset X \to Y$. Dann heißt der Operator abgeschlossen, falls gilt:
Konvergiert eine Folge $(x_n)_{n\in\mathbb{N}}$ für $x_n \in D$ gegen ein $x \in X$ bzgl. $\|\cdot\|_X$ und konvergiert die Folge $(Tx_n)_{n\in\mathbb{N}}$ bzgl. $\|\cdot\|_Y$, etwa gegen $y \in Y$, so folgt $x \in D$ und $Tx = y$.

4 Schwache Topologien

Seien X eine Menge und $(Y_i)_{i\in I}$ eine Familie von topologischen Räumen und für alle $i \in I$ Abbildungen $\phi_i : X \to Y_i$ gegeben. Wir suchen die gröbste Topologie $\mathcal{T}$ auf X, so dass alle ϕ_i stetig sind, d.h. die Topologie, die am wenigsten offene Mengen enthält. Seien $\Omega_i \in Y_i$ offene Mengen, dann sind notwendigerweise alle $\phi_i^{-1}(\Omega_i)$ Elemente von $\mathcal{T}$. Wir bezeichnen die Familie aller solcher Mengen mit $(U_\lambda)_{\lambda\in\Lambda}$ und suchen nun das kleinste Mengensystem $\mathcal{T}$, dass sowohl $(U_\lambda)_{\lambda\in\Lambda}$ enthält als auch abgeschlossen bzgl. endlicher Durchschnitte und beliebiger Vereinigungen ist.
Dazu bilden wir zuerst alle möglichen endlichen Durchschnitte von Mengen aus $(U_\lambda)_{\lambda\in\Lambda}$. Dieses System bezeichnen wir mit $\mathcal{S}$. Danach bilden wir beliebige Vereinigungen von Mengen aus $\mathcal{S}$ Dieses neue System ist dann $\mathcal{T}$, denn es ist klar, dass $\mathcal{T}$ abgeschlossen bzgl. beliebigen Vereinigungen ist. Dass $\mathcal{T}$ auch bzgl. endlicher Schnitte abgeschlossen ist, macht man sich mit etwas Mengentheorie leicht klar.

Lemma 4.1. *Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in X. Die Folge $(x_n)_{n\in\mathbb{N}}$ konvergiert gegen $x \in X$ bzgl. $\mathcal{T}$ genau dann wenn $\phi_i(x_n) \to \phi_i(x)$ für $n \to \infty$ für alle $i \in I$ konvergiert.*

Beweis. Da alle ϕ_i bzgl. $\mathcal{T}$ stetig sind, impliziert $x_n \to x$ bzgl. $\mathcal{T}$ sofort $\phi_i(x_n) \to \phi_i(x)$ für $n \to \infty$ und alle $i \in I$.
Ist umgekehrt U eine offene Umgebung von x. Da eine Umgebungsbasis von x bzgl. $\mathcal{T}$ durch endliche Durchschnitte der Form $\phi_i^{-1}(\Omega_i)$, wobei Ω_i offene Umgebung von $\phi_i(x)$ in Y_i ist, gegeben ist, können wir annehmen, dass U die Form $U = \bigcap_{i\in J} \phi_i^{-1}(\Omega_i)$, $J \subset I$ endlich, hat. Für alle $i \in J$ existiert ein N_i mit $\phi_i(x_n) \in \Omega_i$ für alle $n \geq N_i$. Setzt man $N := \max_{i\in J} N_i$, so ist $x_n \in U$ für alle $n \geq N$. $\square$

Lemma 4.2. *Sei $(Y, \mathcal{S})$ ein topologischer Raum und $\psi : Y \to X$ eine Abbildung. Dann ist ψ stetig genau dann wenn $\phi_i \circ \psi : Y \to Y_i$ für alle $i \in I$ stetige Abbildungen sind.*

Beweis. Die Hinrichtung ist klar, da Kompositionen stetiger Funktionen wiederum stetig sind. Sei umgekehrt $U \in \mathcal{T}$. Dann lässt sich die offene Menge U schreiben als

$$U = \bigcup_{\text{beliebig}} \bigcap_{\text{endlich}} \phi_i^{-1}(\Omega_i), \qquad \Omega_i \subset Y_i \quad \text{offen.}$$

Damit folgt mit den Rechenregeln für Urbilder aus Analysis 1

$$\psi^{-1}(U) = \bigcup_{\text{beliebig}} \bigcap_{\text{endlich}} \psi^{-1}\left(\phi_i^{-1}(\Omega_i)\right) = \bigcup_{\text{beliebig}} \bigcap_{\text{endlich}} (\phi_i \circ \psi)^{-1}(\Omega_i)$$

und $\psi^{-1}(U)$ ist ebenfalls offen. Daher ist ψ stetig. $\square$

Im Folgenden werden wir zwei wichtige Beispiele solcher Topologien kennen lernen: die schwache sowie die schwach-$*$ Topologie. Diese schwächen den Konvergenzbegriff bzgl. einer gegebenen Norm auf einem Banachraum in der Regel wesentlich ab, lassen jedoch hilfreiche schwächere Aussagen zu. Beispielsweise wird die Lösbarkeit der Minimierungsaufgaben wie in den Sätzen 1.46 oder 3.14 auf weitere Banachräume verallgemeinert. Dabei war die Schwierigkeit die Konvergenz einer Minimalfolge oder zumindest die einer Teilfolge zu zeigen. Da jedoch in unendlich dimensionalen Banachräumen die Einheitskugel nicht kompakt ist, ist das nicht möglich. Wie sich zeigen wird, kann aber in manchen Fällen die Folgenkompaktheit bzgl. einer der schwachen Topologien garantiert werden.

4.1 Die schwache Topologie

Wir wenden nun die obige Konstruktion auf einen normierten $\mathbb{K}$-Vektorraum X an sowie auf alle Funktionale $f \in X^*$ und erhalten

Definition 4.3 (Schwache Topologie)**.** Sei X ein normierter Raum, X^* der Dualraum. Die gröbste Topologie $\mathcal{T}_w$ auf X, sodass alle $f \in X^*$ stetig sind, heißt **schwache Topologie** auf X. Man sagt, eine Folge $(x_n)_{n\in\mathbb{N}}$ in X **konvergiert schwach** gegen $x \in X$, falls die Folge bzgl. $\mathcal{T}_w$ konvergiert. Dies ist äquivalent dazu, dass für alle $f \in X^*$ gilt

$$f(x_n) \to f(x) \qquad \text{bzw.} \qquad \langle f, x_n\rangle \to \langle f, x\rangle \qquad \text{für} \quad n \to \infty.$$

Wir schreiben dann auch $x_n \rightharpoonup x$ für $n \to \infty$.

Lemma 4.4. *Sei $(X, \|\cdot\|)$ ein normierter $\mathbb{K}$-Vektorraum. Dann ist $(X, \mathcal{T}_w)$ ein Hausdorff-Raum. Außerdem gilt $\mathcal{T}_w \subset \mathcal{T}_{\|\cdot\|}$ auf X für die von der Norm induzierten Topologie.*

Beweis. Seien $x \neq y$ in X gegeben. Dann findet man nach Korollar 3.24(ii) ein $f \in X^*$ mit $f(x) \neq f(y)$. Da $\mathbb{K}$ als metrischer Raum hausdorff'sch ist, findet man auch offene Umgebungen U_x bzw. U_y von $f(x)$ bzw. von $f(y)$ in $\mathbb{K}$ mit $U_x \cap U_y = \emptyset$. Dann sind $f^{-1}(U_x)$ und $f^{-1}(U_y)$ offene Umgebungen von x bzw. y bzgl. $\mathcal{T}_w$ mit $f^{-1}(U_x) \cap f^{-1}(U_y) = \emptyset$.
Offensichtlich ist $\mathcal{T}_{\|\cdot\|}$ feiner als $\mathcal{T}_w$, denn alle Funktionale $f \in X^*$ sind auch bezüglich $\mathcal{T}_{\|\cdot\|}$ stetig, und $\mathcal{T}_w$ ist die gröbste solche Topologie. □

Damit impliziert **starke Konvergenz**, d.h. Konvergenz bzgl. der Norm $\|\cdot\|$, stets auch schwache Konvergenz. Insbesondere sind im Hausdorff-Raum schwache Limiten eindeutig bestimmt. Unter welchen Bedingungen die schwache mit der starken Topologie übereinstimmt, zeigt

Satz 4.5. *Sei $(X, \|\cdot\|)$ ein normierter $\mathbb{K}$-Vektorraum. Dann gilt*

$$\mathcal{T}_w = \mathcal{T}_{\|\cdot\|} \quad \Leftrightarrow \quad \dim X < \infty.$$

Falls $\dim X = \infty$, so enthält jede $\mathcal{T}_w$ - offene Nullumgebung einen unendlich dimensionalen linearen Unterraum.

Beweis. Wir bemerken zunächst, dass für $x_0 \in X$ eine Umgebungsbasis bzgl. $\mathcal{T}_w$ durch offene Mengen der Form

$$V = \{x \in X \mid |f_i(x - x_0)| < \varepsilon \quad \forall\, i \in J\} = \bigcap_{i\in J} f_i^{-1}\left(B_\varepsilon(f_i(x_0))\right),$$

gegeben ist, wobei J endlich, $f_i \in X^*$ und $\varepsilon > 0$ sind.

$\Rightarrow$: Sei $W \subset X$ eine $\mathcal{T}_w$ - offene Nullumgebung und seien $f_1, \ldots, f_k \in X^*$ und $\varepsilon > 0$ mit

$$V := \{x \in X \mid |f_i(x)| < \varepsilon \quad \forall\, i = 1, \ldots, k\} \subset W.$$

Betrachte den linearen Unterraum

$$N := \{x \in X \mid f_i(x) = 0 \quad \forall\, i = 1, \ldots, k\} \subset W.$$

Falls $\dim X = \infty$, so ist auch N unendlich dimensionaler Untervektorraum, was sofort aus der Dimensionsformel für lineare Abbildungen folgt. Wäre nun die Normkugel $B_1(0) = \{x \in X \mid \|x\| < 1\} \in \mathcal{T}_w$, so gäbe es eine obige Nullumgebung $W \subset B_1(0)$, jedoch enthält $B_1(0)$ sicherlich keinen linearen Unterraum. Das zeigt $\mathcal{T}_w \neq \mathcal{T}_{\|\cdot\|}$.

$\Leftarrow$: O.B.d.A. sei $X = \mathbb{K}^n$ mit $\|x\|_\infty = \max_{i=1,\dots,n} |x_i|$. Man definiere $f_i \in X^*$ mit

$$f_i(x) := x_i \qquad \forall\, i = 1, \dots, n.$$

Ist nun $\Omega \in \mathcal{T}_{\|\cdot\|}$ offen, so gibt es zu jedem $x_0 \in \Omega$ eine Kugel $B_\varepsilon(x_0) \subset \Omega$ und man erhält mit f_i

$$B_\varepsilon(x_0) \supset \{x \in X \mid |f_i(x - x_0)| < \varepsilon \quad \forall\, i = 1, \dots, n\} =: U_{x_0}$$

und wegen $U_{x_0} \in \mathcal{T}_w$ ist auch

$$\Omega = \bigcup_{x_0 \in \Omega} U_{x_0} \in \mathcal{T}_w$$

offen. Damit ist insgesamt $\mathcal{T}_w = \mathcal{T}_{\|\cdot\|}$.

□

Im Allgemeinen ist der topologische Raum $(X, \mathcal{T}_w)$ jedoch nicht metrisierbar und die Begriffe Kompaktheit, Abgeschlossenheit oder Stetigkeit bzgl. der schwachen Topologie lassen sich nicht mit dem entsprechenden schwachen Konvergenzbegriff für Folgen charakterisieren (siehe Übungen).

Beispiel 4.6. (a) Sei $1 \leq p < \infty$, (Ω, Σ, μ) ein Maßraum, der für $p = 1$ σ-finit sei. Dann gilt nach den Sätzen 3.15, 3.16 (für reellwertige Funktionen) für $\frac{1}{p} + \frac{1}{q} = 1$

$$f_k \rightharpoonup f \quad \text{in} \quad L^p(\Omega, \mu) \qquad \Leftrightarrow \qquad \int_\Omega f_k g \,\mathrm{d}\mu \to \int_\Omega f g \,\mathrm{d}\mu \qquad \forall\, g \in L^q(\Omega, \mu).$$

(b) Es lässt sich mit einigem Aufwand zeigen, dass im Folgenraum ℓ^1 schwache Konvergenz äquivalent zur Normkonvergenz ist (Lemma von Schur), jedoch ist nach Satz 4.5 $\mathcal{T}_w \neq \mathcal{T}_{\|\cdot\|}$.

(c) Für $1 < p < \infty$ ist im Folgenraum ℓ^p eine Folge $(x^{(k)})_{k \in \mathbb{N}}$ genau dann schwach konvergent mit $x^{(k)} \rightharpoonup x$ für $k \to \infty$, wenn die Folge beschränkt ist und für jedes $i \in \mathbb{N}$ gilt $x_i^{(k)} \to x_i$ für $k \to \infty$.

Neben dem Dualraum X^* spielt oft auch der Bidualraum

$$X^{**} := (X^*)^* = \mathscr{L}(X^*, \mathbb{K})$$

eine wichtige Rolle. Für $x \in X$ definiert man das lineare Funktional $\tilde{x}$ auf X^* durch

$$\tilde{x}(f) := \langle f, x \rangle = f(x) \qquad \forall\, f \in X^*.$$

Die Abbildung

$$J_X : X \to X^{**}, \quad J_X(x) := \tilde{x} \qquad \forall\, x \in X$$

heißt die **natürliche** (oder **kanonische**) **Inklusion** und man erhält

Lemma 4.7. *Sei $(X, \|\cdot\|_X)$ ein normierter $\mathbb{K}$-Vektorraum. Dann ist die natürliche Inklusion $J_X : X \to X^{**}$ eine lineare Isometrie.*

Beweis. Aus der Abschätzung $|\tilde{x}(f)| \leq \|f\| \cdot \|x\|_X$ folgt die Wohldefiniertheit von J_X und J_X ist offenbar linear. Andererseits gilt nach Korollar 3.24

$$\|J_X(x)\|_{X^{**}} = \|\tilde{x}\|_{X^{**}} = \sup_{\substack{f \in X^*,\\ \|f\|=1}} |\tilde{x}(f)| = \sup_{\substack{f \in X^*,\\ \|f\|=1}} |f(x)| = \|x\|_X.$$

□

Daraus folgen weitere Eigenschaften einer schwach konvergenten Folge.

Lemma 4.8. *Seien $(X, \|\cdot\|_X)$ ein normierter $\mathbb{K}$-Vektorraum, $x \in X$ und $(x_k)_{k\in\mathbb{N}}$ eine schwach konvergente Folge in X mit $x_k \rightharpoonup x$ für $k \to \infty$. Dann gelten*

(i) $(x_k)_{k\in\mathbb{N}}$ ist beschränkt und erfüllt

$$\|x\|_X \leq \liminf_{k\to\infty} \|x_k\|_X.$$

(ii) Konvergiert $f_k \to f$ stark in X^* für $k \to \infty$, so konvergiert

$$\langle f_k, x_k \rangle \to \langle f, x \rangle \qquad \text{für} \quad k \to \infty.$$

Beweis. (i) Nach Voraussetzung gilt $f(x_k) \to f(x)$ für jedes $f \in X^*$. Betrachtet man die natürliche Inklusion J_X von X in seinen Bidualraum X^{**} und die Folge $(\tilde{x}_k)_{k\in\mathbb{N}}$ von linearen Operatoren mit $\tilde{x}_k := J_X(x_k) \in \mathscr{L}(X^*, \mathbb{K})$, so ist diese punktweise für $f \in X^*$ beschränkt gemäß $|\tilde{x}_k(f)| = |f(x_k)| \leq c_f$, denn $(f(x_k))_{k\in\mathbb{N}}$ konvergiert in $\mathbb{K}$. Nach Satz 3.35 ist die Folge damit bereits gleichmäßig beschränkt und es gibt ein $c > 0$ mit

$$\sup_{k\in\mathbb{N}} \|\tilde{x}_k\|_{X^{**}} \leq c.$$

Das impliziert nach Lemma 4.7 die Beschränktheit der Folge, d.h.

$$\|x_k\|_X = \|J_X(x_k)\|_{X^{**}} \leq c \qquad \forall\, k \in \mathbb{N}.$$

Aufgrund der schwachen Konvergenz ist dann auch

$$|f(x)| = \lim_{k\to\infty} |f(x_k)| \leq \liminf_{k\to\infty} \|f\|\|x_k\|_X$$

und Korollar 3.24 liefert

$$\|x\|_X = \sup_{\substack{f \in X^*,\\ \|f\|=1}} f(x) \leq \liminf_{k\to\infty} \|x_k\|_X.$$

(ii) Folgt aus der Abschätzung

$$|f_k(x_k) - f(x)| \leq \|f_k - f\|\|x_k\|_X + |f(x_k) - f(x)| \leq c\|f_k - f\| + |f(x_k) - f(x)|.$$

□

Beispiel 4.9. Im Folgenden sei $1 < p < \infty, n \in \mathbb{N}$.

(a) Betrachte die Einheitsvektoren $e_k = (0, \dots, 0, 1, 0, \dots) \in \ell^p$ mit der 1 an der k-ten Stelle. Dann ist $\|e_k\|_p = 1$ für alle $k \in \mathbb{N}$, aber $e_k \rightharpoonup 0$ für $k \to \infty$, d.h. $(e_k)_{k\in\mathbb{N}}$ kann nicht stark konvergieren.

(b) (Konzentration) Sei $\psi \in C_c(\mathbb{R}^n)$ mit $\operatorname{supp}\psi \subset B_1(0)$ und $\|\psi\|_{L^p(\mathbb{R}^n)} = 1$. Definiere

$$f_k : B_1(0) \to \mathbb{R}, \quad x \mapsto k^{n/p}\psi(kx).$$

Dann gilt $\|f_k\|_{L^p(B_1(0))} = \|\psi\|_{L^p(\mathbb{R}^n)} = 1$, $f_k(x) \to 0$ für $x \neq 0$ und $f_k \rightharpoonup 0$ in $L^p(B_1(0))$ für $k \to \infty$.

(c) (Oszillation) Sei $\Omega = (0, 1) \subset \mathbb{R}$ undi $h : \mathbb{R} \to \mathbb{R}$ p-periodisch, d.h. $h(x + p) = h(x)$ für alle $x \in \mathbb{R}$ mit $\|h\|_{L^p(\Omega)} < \infty$. Definiere $g_k(x) = h(kx)$, so ist $\|g_k\|_{L^p(\Omega)} = \|h\|_{L^p(\Omega)}$ und

$$g_k \rightharpoonup \frac{1}{p}\int_0^p h \,\mathrm{d}\lambda \qquad \text{in} \quad L^p(\Omega).$$

Wie bereits im Beweis von Satz 4.5 erwähnt, ist die bzgl. der Norm offene Kugel $B_1(0)$ im unendlich dimensionalen Fall nicht bezüglich $\mathcal{T}_w$ offen. Entsprechend ist $X \setminus B_1(0)$ nicht **schwach abgeschlossen**, d.h. nicht abgeschlossen bezüglich $\mathcal{T}_w$. Wir wollen nun zeigen, dass jedoch konvexe Mengen $M \subset X$, welche (stark) abgeschlossen sind bzgl. der Norm, bzgl. schwacher Konvergenz abgeschlossen sind. Dazu benötigen wir die folgende geometrische Variante des Satzes 3.21 von Hahn-Banach.

Satz 4.10 (Trennungssatz)**.** *Seien $(X, \|\cdot\|_X)$ ein normierter $\mathbb{K}$-Vektorraum, $A \subset X$ abgeschlossen und konvex. Dann gibt es zu jedem $x_0 \in X \setminus A$ ein $f \in X^*$ sowie ein $\alpha \in \mathbb{R}$ mit*

$$\operatorname{Re} f(x) \leq \alpha \quad \forall\, x \in A \qquad \textit{und} \qquad \operatorname{Re} f(x_0) > \alpha.$$

Beweis. Wir betrachten nur den Fall $\mathbb{K} = \mathbb{R}$. Im komplexen Fall gehe man wie im Beweis von Satz 3.22 vor und betrachte X als $\mathbb{R}$-Vektorraum. Ohne Einschränkung sei $0 \in \mathring{A}$ (durch Translation und gegebenenfalls Übergang zu $\overline{B_\varepsilon(A)}$ für $\varepsilon < \operatorname{dist}(x_0, A)$). Wir definieren das Minkowski-Funktional für $x \in X$

$$p(x) := \inf\left\{r > 0 \mid \frac{x}{r} \in A\right\}.$$

Wegen $0 \in \mathring{A}$ gibt es eine kleine Umgebung $B_\delta(0) \subset \mathring{A}$ und es folgt

$$0 \leq p(x) \leq \|x\|/\delta < \infty \qquad \forall\, x \in X.$$

Nach den Übungen ist p sublinear und es gilt

$$p(x) \leq 1 \quad \forall\, x \in A \qquad \text{und} \qquad p(x_0) > 1.$$

Wir definieren nun die lineare Abbildung $f : \operatorname{span}\{x_0\} \to \mathbb{R}$ durch

$$f(\lambda x_0) := \lambda p(x_0) \qquad \forall\, \lambda \in \mathbb{R}.$$

Eine Fallunterscheidung für $\lambda \in \mathbb{R}$ liefert $f \leq p$ auf $\operatorname{span}\{x_0\}$ und es lässt sich Satz 3.21 anwenden. Danach existiert eine lineare Fortsetzung F von f auf ganz X mit $F \leq p$. Wegen $0 \in \mathring{A}$ ist $\pm F(x) = F(\pm x) \leq \|x\|/\delta$ und damit $F \in X^*$. Es lässt sich nun $\alpha = 1$ wählen, denn $F \leq p \leq 1$ auf A und $F(x_0) = f(x_0) = p(x_0) > 1$. $\square$

Satz 4.11. *Seien $(X, \|\cdot\|_X)$ ein normierter $\mathbb{K}$-Vektorraum, $A \subset X$ abgeschlossen und konvex. Dann ist A* ***schwach folgenabgeschlossen****, d.h. für jede in X schwach konvergente Folge $x_k \rightharpoonup x$ mit $x_k \in A$ ist auch $x \in A$. Außerdem ist A schwach abgeschlossen, d.h. abgeschlossen bzgl. $\mathcal{T}_w$.*

Beweis. Wäre für den schwachen Limes $x \notin A$, so gäbe es nach dem Trennungssatz 4.10 ein $f \in X^*$ und ein $\alpha \in \mathbb{R}$ mit

$$\operatorname{Re} f(y) \leq \alpha \quad \forall\, y \in A \qquad \text{und} \qquad \operatorname{Re} f(x) > \alpha.$$

Die schwache Konvergenz impliziert aber $\operatorname{Re} f(x) \leq \alpha$, ein Widerspruch.
Für die Abgeschlossenheit bzgl. $\mathcal{T}_w$ zeigen wir die Offenheit von $X \setminus A$. Sei also $\xi \notin A$, dann gibt es nach obigem Satz durch Übergang zu $g = -f \in X^*$ ein $\beta \in \mathbb{R}$ (etwa der Mittelwert $\beta = \frac{1}{2}(-\alpha + \operatorname{Re} g(\xi))$), sodass

$$\operatorname{Re} g(\xi) < \beta < \operatorname{Re} g(y) \quad \forall\, y \in A$$

gilt. Nun ist

$$V := \{x \in X \mid \operatorname{Re} g(x) < \beta\} = g^{-1}(\{z \in \mathbb{K} \mid \operatorname{Re} z < \beta\})$$

eine offene Umgebung in $\mathcal{T}_w$ mit $\xi \in V \subset X \setminus A$ und das Komplement von A ist offen. □

4.2 Die schwach - $*$ Topologie

Viele Sätze der Funktionalanalysis beinhalten den Dualraum eines normierten Raums. Neben der starken Topologie $\mathcal{T}_{\|\cdot\|}$, gegeben durch die Norm auf X^*, lässt sich analog die schwache Topologie $\mathcal{T}_w$ wie in vorherigem Abschnitt auf X^* definieren, sodass alle Funktionale $f \in X^{**}$ bezüglich dieser stetig sind. Eine weitere etwas schwächere Möglichkeit bietet sich, indem man sich auf die Funktionale $f \in J_X(X) \subset X^{**}$ der natürlichen Inklusion einschränkt. Dabei war (siehe Lemma 4.7)

$$J_X : X \to X^{**}, \quad J_X(x) := \tilde{x} \qquad \forall\, x \in X$$

mit dem linearen Funktional $\tilde{x}$ auf X^* definiert durch

$$\tilde{x}(f) := \langle f, x\rangle = f(x) \qquad \forall\, f \in X^*.$$

Definition 4.12 (Schwach - $*$ Topologie)**.** Sei X ein normierter Raum. Auf X^* definiert man die **schwach - $*$ Topologie** (bezeichnet mit $\mathcal{T}_w^*$) als die gröbste Topologie bzgl. welcher alle $f \in J_X(X) \subset X^{**}$ stetig sind. Man sagt, eine Folge $(f_n)_{n\in\mathbb{N}}$ in X^* **konvergiert schwach - $*$** gegen $f \in X^*$, falls die Folge bzgl. $\mathcal{T}_w^*$ konvergiert. Dies ist äquivalent zur punktweise Konvergenz, d.h. dass für alle $x \in X$ gilt

$$f_n(x) \to f(x) \qquad \text{bzw.} \qquad \langle f_n, x\rangle \to \langle f, x\rangle \qquad \text{für} \quad n \to \infty.$$

Wir schreiben dann $f_n \overset{*}{\rightharpoonup} f$ für $n \to \infty$.

Lemma 4.13. *Sei $(X, \|\cdot\|)$ ein normierter $\mathbb{K}$-Vektorraum. Dann ist $(X^*, \mathcal{T}_w^*)$ hausdorff'sch und schwach - $*$ Limiten sind eindeutig. Zudem gilt $\mathcal{T}_w^* \subset \mathcal{T}_w \subset \mathcal{T}_{\|\cdot\|}$ auf X^*.*

Beweis. Seien $f, g \in X^*$ mit $f \neq g$ gegeben. Dann findet man ein $x \in X$ mit $f(x) \neq g(x)$ bzw. per Definition $J_X(x)(f) \neq J_X(x)(g)$. In $\mathbb{K}$ gibt es offene Umgebungen U_f, U_g von $f(x)$ bzw. $g(x)$ mit $U_f \cap U_g = \emptyset$. Betrachtet man die (in $\mathcal{T}_w^*$ offenen) Urbilder $J_X(x)^{-1}(U_f)$ bzw. $J_X(x)^{-1}(U_g)$, so erhält man wiederum die Trennungseigenschaft. Somit ist $(X^*, \mathcal{T}_w^*)$ ein Hausdorff-Raum. □

In endlich dimensionalen Räumen stimmen nach Linearer Algebra die Dimensionen von X, X^* sowie X^{**} überein und mit Satz 4.5 folglich auch die starke, die schwache und die schwach - $*$ Topologie. Im unendlich dimensionalen Fall ist dies im Allgemeinen nicht richtig wie wir im nächsten Abschnitt sehen werden. Wie man sich leicht klar macht, gilt

Lemma 4.14. *Seien $(X, \|\cdot\|_X)$ ein normierter $\mathbb{K}$-Vektorraum, $f \in X^*$ und $(f_k)_{k\in\mathbb{N}}$ eine schwach - $*$ konvergente Folge in X^* mit $f_k \overset{*}{\rightharpoonup} f$ für $k \to \infty$. Dann gelten*

(i) $(f_k)_{k\in\mathbb{N}}$ ist beschränkt in X^* und erfüllt

$$\|f\|_{X^*} \leq \liminf_{k\to\infty} \|f_k\|_{X^*}.$$

(ii) Konvergiert $x_k \to x$ stark in X für $k \to \infty$, so konvergiert

$$\langle f_k, x_k \rangle \to \langle f, x \rangle \qquad \textit{für} \quad k \to \infty.$$

Beweis. Analog zu Lemma 4.8. □

Bemerkung: Für einen beliebigen topologischen Vektorraum $(X, \mathcal{T})$ kann man den Dualraum $(X, \mathcal{T})^*$ definieren als denjenigen Raum aller $\mathcal{T}$-stetigen linearen Funktionale auf X. Dann folgt für den normierten Raum $(X, \|\cdot\|)$

$$(X, \mathcal{T}_w)^* = X^* \equiv \left(X, \mathcal{T}_{\|\cdot\|}\right)^*,$$

denn jedes $f \in \left(X, \mathcal{T}_{\|\cdot\|}\right)^*$ ist per Definition stetig bzgl. $\mathcal{T}_w$ und andererseits impliziert $\mathcal{T}_w \subset \mathcal{T}_{\|\cdot\|}$, dass jede $\mathcal{T}_w$ - stetige Funktion auch $\mathcal{T}_{\|\cdot\|}$ - stetig ist. Analog gilt

$$(X^*, \mathcal{T}_w)^* = X^{**} \equiv \left(X^*, \mathcal{T}_{\|\cdot\|_{X^*}}\right)^* \quad \text{und} \quad (X^*, \mathcal{T}_w^*)^* = J_X(X) \subset X^{**}.$$

Schwache Topologien spielen deswegen eine wichtige Rolle, weil sie die Kompaktheit (bzw. die nicht notwendigerweise äquivalente Folgenkompaktheit) der abgeschlossenen Einheitskugel unter gewissen Voraussetzungen wiederherstellen. In X^* sorgt dafür die schwach - $*$ Topologie.

Definition 4.15. Sei X ein normierter Raum. Eine Menge $M \subset X$ (bzw. X^*) heißt **schwach (bzw. schwach - $*$) folgenkompakt**, falls jede Folge in M eine schwach (bzw. schwach - $*$) konvergente Teilfolge besitzt, deren schwacher (bzw. schwach - $*$) Limes in M liegt.

Satz 4.16. *Sei $(X, \|\cdot\|_X)$ ein normierter, separabler $\mathbb{K}$-Vektorraum. Dann ist die abgeschlossene Einheitskugel*

$$k_{X^*} := \{f \in X^* \mid \|f\| \leq 1\}$$

schwach - $*$ folgenkompakt.

Beweis. Sei $(x_m)_{m\in\mathbb{N}}$ eine dichte Folge in X und $(f_n)_{n\in\mathbb{N}}$ eine Folge in k_{X^*}, d.h. $\|f_n\| \le 1$. Wegen $|f_n(x_1)| \le \|x_1\|_X$ ist $(f_n(x_1))_{k\in\mathbb{N}}$ beschränkt in $\mathbb{K}$. Somit gibt es ein $\alpha_1 \in \mathbb{K}$ und eine Teilfolge $f_{n_{1,k}}(x_1) \to \alpha_1$ für $k \to \infty$. Ähnlich ist $(f_{n_{1,k}}(x_2))_{k\in\mathbb{N}}$ beschränkt in $\mathbb{K}$ und es lässt sich wiederum ein Häufungspunkt $\alpha_2 \in \mathbb{K}$ auswählen mit $f_{n_{2,k}}(x_j) \to \alpha_j$ für $j = 1, 2$ und $k \to \infty$, wobei $n_{2,k}$ eine Teilfolge von $n_{1,k}$ ist. Iterativ findet man für alle $\ell \in \mathbb{N}$ ein $\alpha_\ell \in \mathbb{K}$ und eine Teilfolge $n_{\ell,k}$, sodass

$$f_{n_{\ell,k}}(x_j) \to \alpha_j \qquad \forall\, j = 1, \dots, \ell \qquad \text{für} \quad k \to \infty.$$

Nun betrachte man die diagonale Folge für $m_\ell := n_{\ell,\ell}$. Dann gilt $f_{m_\ell}(x_j) \to \alpha_j$ für alle $j \in \mathbb{N}$ für $\ell \to \infty$. Für beliebige $x \in X$ erhält man

$$\begin{aligned} |f_{m_\ell}(x) - f_{m_k}(x)| &\le \big(\|f_{m_\ell}\| + \|f_{m_k}\|\big) \cdot \|x - x_j\|_X + |f_{m_\ell}(x_j) - f_{m_k}(x_j)| \\ &\le 2\,\|x - x_j\|_X + |f_{m_\ell}(x_j) - f_{m_k}(x_j)|. \end{aligned}$$

Zu gegebenem $\varepsilon > 0$ liefert die Dichtheit ein $j \in \mathbb{N}$ mit $\|x - x_j\|_X \le \frac{\varepsilon}{4}$ und die Cauchy-Eigenschaft von $(f_{m_\ell}(x_j))_{\ell\in\mathbb{N}}$ impliziert

$$|f_{m_\ell}(x) - f_{m_k}(x)| \le \frac{\varepsilon}{2} + |f_{m_\ell}(x_j) - f_{m_k}(x_j)| \le \varepsilon$$

für $\ell, k \in \mathbb{N}$ groß genug. Deswegen existiert der Grenzwert

$$\lim_{\ell\to\infty} f_{m_\ell}(x) =: f(x) \qquad \forall\, x \in X.$$

Es ist einfach zu sehen, dass f linear ist und wegen

$$|x^*(x)| \le \limsup_{\ell\to\infty} \|f_{m_\ell}\|\,\|x\| \le \|x\|,$$

dass $f \in k_{X^*}$ liegt. Nun folgt aus

$$J_X(x)(f_{m_\ell}) = f_{m_\ell}(x) \to f(x) = J_X(x)(f) \qquad \forall\, x \in X,$$

die Konvergenz $f_{m_\ell} \to f$ bzgl. der $\mathcal{T}_w^*$ - Topologie für $\ell \to \infty$. □

Beispiel 4.17. Sei $X = L^\infty(\Omega)$ für $\Omega = (0, 1) \subset \mathbb{R}$ und für $\varepsilon > 0$ definiere

$$T_\varepsilon f := \frac{1}{\varepsilon} \int_0^\varepsilon f \,\mathrm{d}\lambda \qquad \forall\, f \in L^\infty(\Omega).$$

Dann ist $T_\varepsilon \in (L^\infty(\Omega))^*$ mit $\|T_\varepsilon\| = 1$, jedoch gibt es keine Nullfolge $(\varepsilon_k)_{k\in\mathbb{N}}$, sodass $(T_{\varepsilon_k})_{k\in\mathbb{N}}$ schwach - ∗ konvergent in $(L^\infty(\Omega))^*$ ist.

Im Hausdorff-Raum X^* sind Kompaktheit und Folgenkompaktheit in der Regel verschiedene Konzepte. Um die Kompaktheit der abgeschlossenen Einheitskugel im Dualraum zu zeigen, benötigen wir folgenden Satz, der ohne Beweis gegeben sei.

Satz 4.18 (Tychonoff)**.** *Sei $(X_\lambda)_{\lambda\in\Lambda}$ eine Familie kompakter Räume. Dann ist das Produkt*

$$X := \prod_{\lambda\in\Lambda} X_\lambda$$

kompakt bzgl. der Produkttopologie, die auf X definiert ist.

Bemerkung: Für eine (nicht notwendigerweise abzählbare) Familie topologischer Räume $(X_\lambda, \mathcal{T}_\lambda)$ betrachtet man den Produktraum $X = \prod_{\lambda\in\Lambda} X_\lambda$ und die Projektionen

$$\pi_i : \prod_{\lambda\in\Lambda} X_\lambda \to X_i, \quad (x_\lambda)_{\lambda\in\Lambda} \mapsto x_i.$$

Die kanonische Topologie auf X ist die schwache Topologie, die sogenannte Produkttopologie, d.h. die gröbste Topologie bzgl. welcher alle $\pi_\lambda, \lambda \in \Lambda$, stetig sind. $(X, \mathcal{T}_w)$ ist wiederum eine Hausdorff-Raum, falls alle $X_\lambda, \lambda \in \Lambda$, hausdorff'sch sind.

Satz 4.19 (Banach-Alaoglu)**.** *Sei $(X, \|\cdot\|_X)$ ein normierter $\mathbb{K}$-Vektorraum und X^* sein Dualraum. Dann ist*

$$k_{X^*} = \{f \in X^* \mid \|f\| \le 1\}$$

kompakt bzgl. der $\mathcal{T}_w^*$-Topologie.

Beweis. Für $x \in X$ definiere man den kompakten Raum (mit der durch $\mathbb{K}$ induzierten Topologie)

$$D_x := \{\alpha \in \mathbb{K} \mid |\alpha| \le \|x\|_X\}.$$

Dann ist das Produkt

$$P := \prod_{x\in X} D_x$$

mit der Produkttopologie $\mathcal{T}_P$ versehen ebenfalls nach Satz 4.18 kompakt. Elemente von P sind Familien $(\alpha_x)_{x\in X}$ und können als Abbildungen $\varphi : X \to \mathbb{K}$ mit $\varphi(x) = \alpha_x$ aufgefasst werden. Wegen $|\varphi(x)| \le \|x\|_X$ für jedes $x \in X$ ist

$$k_{X^*} = X^* \cap P.$$

Man zeige nun, dass k_{X^*} eine bzgl. $\mathcal{T}_P$ abgeschlossene Teilmenge von $(P, \mathcal{T}_P)$ ist. Dies impliziert, dass k_{X^*} als abgeschlossene Teilmenge eines kompakten Raumes selbst kompakt ist bzgl. der von $\mathcal{T}_P$ induzierten Teilraumtopologie.
Wir nutzen die Projektionen $p_x : P \to D_x$ mit $p_x(\varphi) = \alpha_x = \varphi(x)$ und definieren für $x, y \in X, \lambda \in \mathbb{K}$ eine Abbildung

$$\sigma_{x,y,\lambda} : P \to \mathbb{K}, \quad \sigma_{x,y,\lambda}(\varphi) := \varphi(x + \lambda\, y) - \varphi(x) - \lambda\, \varphi(y).$$

Wegen $\sigma_{x,y,\lambda} = p_{x+\lambda\, y} - p_x - \lambda\, p_y$ ist diese Abbildung stetig bzgl. der Topologie $\mathcal{T}_P$. Da $\{0\} \subset \mathbb{K}$ abgeschlossen ist, folgt, dass auch $\sigma^{-1}_{x,y,\lambda}(\{0\})$ abgeschlossen in P bzgl. $\mathcal{T}_P$ ist, d.h.

$$k_{X^*} = \bigcap_{\substack{x,y\in X\\ \lambda\in\mathbb{K}}} \sigma^{-1}_{x,y,\lambda}(\{0\})$$

ist abgeschlossen und deswegen auch kompakt bzgl. $\mathcal{T}_P$.
Nun zeige man $\mathcal{T}_w^*|_{k_{X^*}} \subset \mathcal{T}_P|_{k_{X^*}}$, was die Kompaktheit bzgl. $\mathcal{T}_w^*$ zur Folge hat. Eine Umgebungsbasis von $\varphi \in k_{X^*}$ bzgl. $\mathcal{T}_w^*$ besteht aus Mengen der Form

$$V = \{f \in k_{X^*} \mid |(f - \varphi)(x_i)| < \varepsilon \quad \forall\, i \in J\},$$

wobei $\varepsilon > 0, J$ endlich und $x_i \in X$ sind. Wegen $f(x_i) = J_X(x_i)(f)$ ist

$$V = \bigcap_{i\in J} (J_X(x_i))^{-1}\,(B_\varepsilon(\varphi(x_i)))$$

Dann ist aber mit $k_{X^*} \subset P$ auch

$$V \supset \bigcap_{i \in J} p_{x_i}^{-1}\left(B_\varepsilon(\varphi(x_i))\right) =: U_\varphi$$

mit offener Umgebung $U_\varphi \in \mathcal{T}_P$ als Schnitt offener Urbilder der stetigen Projektionen p_{x_i}. Betrachtet man $\Omega \in \mathcal{T}_w^*$ als Vereinigung all dieser Umgebungen U_φ, so zeigt dies $\mathcal{T}_w^*|_{k_{X^*}} \subset \mathcal{T}_P|_{k_{X^*}}$. □

4.3 Reflexive Räume

Im Folgenden betrachten wir spezielle normierte Räume, welche schwache Kompaktheit der abgeschlossenen Einheitskugel in X garantieren werden und damit sehr nützlich sind in vielen Anwendungen.

Definition 4.20. Ein normierter Raum $(X, \|\cdot\|_X)$ heißt **reflexiv**, falls die natürliche Inklusion J_X aus Lemma 4.7 surjektiv ist, d.h. $J_X(X) = X^{**}$.

Bemerkung: X^{**} ist stets ein Banachraum. Ist X reflexiv, so ist daher X ein Banachraum und es gilt $\mathcal{T}_w = \mathcal{T}_w^*$ auf X^*. Jedoch ist J_X im Allgemeinen kein isometrischer Isomorphismus wie nachfolgende Beispiele zeigen.

Satz 4.21. *Sei X ein reflexiver Banachraum und $M \subset X$ ein abgeschlossener linearer Unterraum. Dann ist M reflexiv.*

Beweis. Seien $J_M : M \to M^{**}$ und $J_X : X \to X^{**}$ die natürlichen Inklusionen. Sei $j : M \to X$ die Inklusion von M in X (d.h. $j(x) = x$ für alle $x \in M$), dann ist die duale Abbildung (siehe Definition 3.7) hierzu gegeben durch

$$j^* : X^* \to M^*, \quad j^*(f) = f|_M \qquad \forall\, f \in X^*.$$

und $j^{**} : M^{**} \to X^{**}$ die biduale Abbildung zu j mit $j^{**}(m^{**})(f) = m^{**}(f|_M)$ für alle $m^{**} \in M^{**}, f \in X^*$. Sei $m^{**} \in M^{**}$. Man finde nun ein $m \in M$ mit $J_M(m) = m^{**}$. Da J_X surjektiv ist, findet man zumindest ein $x \in X$ mit $J_X(x) = j^{**}(m^{**})$. Dann ist bereits $x \in M$, denn angenommen, dies wäre nicht der Fall, existiert nach Korollar 3.23 ein $f \in X^*$ mit $f|_M = 0$ und $f(x) = \operatorname{dist}(x, M) > 0$. Damit folgt aber ein Widerspruch

$$0 < f(x) = (J_X(x))\,(f) = (j^{**}(m^{**}))\,(f) = m^{**}(f|_M) = 0.$$

Also ist $x \in M$ und es genügt $J_M(x) = m^{**}$ zu zeigen. Für $m^* \in M^*$ findet man nach Satz 3.22 von Hahn-Banach eine Fortsetzung $f \in X^*$ mit $f|_M = m^*$ und es folgt

$$\begin{aligned}(J_M(x))\,(m^*) &= m^*(x) = f(x) = (J_X(x))\,(f) \\ &= (j^{**}(m^{**}))\,(f) = m^{**}\,(f|_M) \\ &= m^{**}(m^*),\end{aligned}$$

d.h. $J_M(x) = m^{**}$. □

Satz 4.22. *Seien X, Y ein Banachräume. Dann gelten*

(i) Ist $\Phi : X \to Y$ ein Isomorphismus, dann ist X reflexiv genau dann wenn Y reflexiv ist.

(ii) X ist reflexiv $\Leftrightarrow$ X^* ist reflexiv.

Beweis. Wir zeigen hier nur die Richtung X reflexiv $\Rightarrow X^*$ reflexiv. Die übrigen Aussagen folgen aus den Übungen.
Sei also X reflexiv. Man betrachte $J_{X^*}: X^* \to X^{***}$ und zeige, dass J_{X^*} surjektiv ist. Sei $x_0^{***} \in X^{***}$ ein lineares Funktional auf X^{**} und definiere $x_0^* := x_0^{***} \circ J_X \in X^*$. Für ein beliebiges $x^{**} \in X^{**}$ lässt sich aufgrund der Reflexivität ein $x \in X$ finden mit $J_X(x) = x^{**}$. Man erhält dann per Definition von J_{X^*}

$$\begin{aligned}(J_{X^*}(x_0^*))(x^{**}) &= x^{**}(x_0^*) = x^{**}(x_0^{***} \circ J_X)\\ &= (J_X(x))(x_0^{***} \circ J_X) = (x_0^{***} \circ J_X)(x)\\ &= x_0^{***}(x^{**}),\end{aligned}$$

d.h. X^* ist reflexiv. □

Beispiel 4.23.

(a) Endlich dimensionale Banachräume sind nach Linearer Algebra reflexiv: $\dim X = \dim X^* = \dim X^{**}$.

(b) Sei (Ω, Σ, μ) ein Maßraum. Für $1 < p < \infty$ ist $L^p(\Omega, \mu)$ reflexiv.

(c) Sei (Ω, Σ, μ) ein σ-finiter Maßraum. Im Allgemeinen sind $L^1(\Omega, \mu)$ und $L^\infty(\Omega, \mu)$ nicht reflexiv. In der Tat, falls $L^1(\Omega, \mu)$ reflexiv wäre, so wäre $(L^1(\Omega, \mu))^{**}$ separabel. Das impliziert nach Lemma 4.26, dass $(L^1(\Omega, \mu))^*$ auch separabel ist. Jedoch ist $(L^1(\Omega, \mu))^* = L^\infty(\Omega, \mu)$ im Allgemeinen nicht separabel. Ist $L^\infty(\Omega, \mu)$ nicht separabel, so liefert Satz 4.22, dass $(L^1(\Omega, \mu))^* = L^\infty(\Omega, \mu)$ auch nicht reflexiv ist.

Satz 4.24 (Milman-Pettis)**.** *Jeder gleichmäßig konvexe Raum X ist reflexiv.*

Beweis. Ohne Beweis. Siehe etwa Werner: Funktionalanalysis, Satz IV.7.12. □

Satz 4.25. *Sei X ein reflexiver Banachraum und $K \subset X$ eine konvexe, abgeschlossene, beschränkte Teilmenge. Dann ist K kompakt bezüglich der schwachen Topologie $\mathcal{T}_w$.*

Beweis. Da K konvex und abgeschlossen bzgl. der Normtopologie ist, folgt mit Satz 4.11, dass K abgeschlossen ist bzgl. $\mathcal{T}_w$. Weil K beschränkt ist, ist K in einer Normkugel enthalten, d.h. es gibt ein $r > 0$ mit $K \subset B_r(0) = rB_1(0)$ und es genügt die Kompaktheit der abgeschlossenen Einheitskugel

$$k_X := \{x \in X \mid \|x\| \leq 1\}$$

zu zeigen:
Da X reflexiv ist, ist J_X surjektiv, d.h. $J_X(X) = X^{**}$. Da J_X eine Isometrie ist, folgt $J_X(k_X) = k_{X^{**}}$, wobei letztere die abgeschlossene Einheitskugel im Bidualraum bezeichne. Aus Satz 4.19 wissen wir, dass $k_{X^{**}}$ kompakt ist bzgl. der schwach-$*$ Topologie $\mathcal{T}_w^*$ auf X^{**}, welche mit der schwachen Topologie übereinstimmt, denn X^{**} ist ebenfalls reflexiv nach Satz 4.22. Daher genügt es zu zeigen, dass

$$J_X^{-1}: (X^{**}, \mathcal{T}_w) \to (X, \mathcal{T}_w)$$

stetig ist. Nach Lemma 4.2 reicht es für jedes $f \in X^*$ die Stetigkeit der Abbildung

$$L_f = f \circ J_X^{-1} : (X^{**}, \mathcal{T}_w) \to (\mathbb{K}, |\cdot|), \quad x^{**} \mapsto f\left(J_X^{-1}(x^{**})\right)$$

zu prüfen. Per Definition von J_X ist dann $L_f(x^{**}) = x^{**}(f)$ und L_f ist stetig nach Konstruktion der schwachen Topologie auf X^{**}. Damit ist k_X kompakt bzgl. der $\mathcal{T}_w$-Topologie. □

Mit einigem Aufwand lässt sich auch zeigen, dass ein Banachraum reflexiv ist, falls k_X kompakt bzgl. der schwachen Topologie $\mathcal{T}_w$ ist. Wir wollen nun untersuchen, ob k_X auch schwach folgenkompakt ist. Dazu benötigen wir einige Resultate für separable Räume.

Lemma 4.26. *Sei X ein Banachraum und X^* separabel. Dann ist auch X separabel.*

Beweis. Siehe Übungen. □

Korollar 4.27. *Sei X ein Banachraum, dann sind die folgenden Aussagen äquivalent:*

(i) X ist reflexiv und separabel.

(ii) X^* ist reflexiv und separabel.

Beweis. Siehe Übungen. □

Der Begriff der Separabilität ist eng verwandt mit der Metrisierbarkeit schwacher Topologien. Es lässt sich zeigen, dass unter gewissen Voraussetzungen die abgeschlossene Einheitskugel k_X metrisierbar ist und damit Folgenkompaktheit und Kompaktheit äquivalent sind.

Als zentrales Resultat für reflexive Banachräume erhalten wir daher ein Analogon des Satzes von Bolzano-Weierstraß.

Satz 4.28. *Sei X ein reflexiver Banachraum. Dann ist die abgeschlossene Einheitskugel k_X schwach folgenkompakt. Insbesondere besitzt jede beschränkte Folge in X eine schwach konvergente Teilfolge.*

Beweis. Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in k_X. Dann ist

$$Y := \overline{\operatorname{span}\{x_n \mid n \in \mathbb{N}\}}$$

als abgeschlossener Unterraum von X reflexiv und per Definition separabel. Nach Korollar 4.27 ist dann auch Y^* reflexiv und separabel. Dann lässt sich Satz 4.16 auf Y^* und die beschränkte Folge $(J_Y(x_n))_{n\in\mathbb{N}}$ in Y^{**} anwenden. Folglich gibt es ein $y^{**} \in Y^{**}$ und eine Teilfolge, sodass für $k \to \infty$ gilt

$$(J_Y(x_{n_k}))(f) \to y^{**}(f) \qquad \forall\, f \in Y^*.$$

Mit $x := J_Y^{-1} y^{**}$ folgt dann

$$f(x_{n_k}) = (J_Y(x_{n_k}))(f) \to y^{**}(f) = f(x) \qquad \forall\, f \in Y^*.$$

Dann gilt dies auch für alle $f \in X^*$, denn $f|_Y \in Y^*$, und die Teilfolge konvergiert schwach $x_{n_k} \rightharpoonup x$ für $k \to \infty$. □

Die schwache Folgenkompaktheit für abgeschlossene Kugeln ist essentiell für die Lösbarkeit von Minimierungsaufgaben.

Satz 4.29. *Sei $(X, \|\cdot\|)$ ein reflexiver Banachraum und sei $A \subset X$ konvex, abgeschlossen und nicht–leer. Dann gibt es zu jedem $x \in X$ ein $y \in A$ mit*

$$\|x - y\| = \inf_{a \in A} \|x - a\|.$$

Beweis. Sei $x \in X$ und $(z_n)_{n\in\mathbb{N}}$ eine Minimalfolge in A mit $\|x - z_n\| \to \inf_{a\in A} \|x - a\| =: d$. Wegen $\|z_n\| \leq \|z_n - x\| + \|x\|$ ist $(z_n)_{n\in\mathbb{N}}$ beschränkt und besitzt daher nach Satz 4.28 eine schwach konvergente Teilfolge $(z_{n_k})_{k\in\mathbb{N}}$ mit $z_{n_k} \rightharpoonup y$. Da A abgeschlossen und konvex ist, ist A auch schwach abgeschlossen, also $y \in A$. Schließlich ist $x - z_{n_k} \rightharpoonup x - y$, d.h. nach Lemma 4.8 $\|x - y\| \leq \liminf_{k\to\infty} \|x - z_{n_k}\|$. □

Im Allgemeinen erhält man keine eindeutige Projektion auf konvexe, abgeschlossene Teilmengen wie man etwa für $\mathbb{R}^2$ mit der Maximumsnorm sieht. Jedoch gilt Eindeutigkeit im Fall der L^p-Räume wie Satz 3.14 beweist.

5 Spektraltheorie

Die Spektraltheorie verallgemeinert die Eigenwerttheorie der Linearen Algebra für Operatoren auf endlich dimensionalen Räumen, d.h. für Matrizen, auf unendlich dimensionale Banachräume. Im letzteren Fall muss jedoch eine injektive Abbildung im Allgemeinen nicht surjektiv sein und wir benötigen eine allgemeineres Konzept, welches mehr als nur Eigenwerte charakterisiert.

Definition 5.1. Sei X ein Banachraum über $\mathbb{K}$ und $T \in \mathscr{L}(X)$. Man definiere die **Resolventenmenge**

$$\varrho(T) := \{\lambda \in \mathbb{K} \mid \ker(\lambda I - T) = \{0\} \text{ und } \operatorname{ran}(\lambda I - T) = X\},$$

d.h. $\varrho(T)$ ist die Menge aller $\lambda \in \mathbb{K}$, sodass $\lambda I - T$ invertierbar (und mit Satz 3.38 $(\lambda I - T)^{-1} \in \mathscr{L}(X)$) ist. Damit ist die **Resolvente** von T gegeben durch

$$R_T : \varrho(T) \to \mathscr{L}(X), \quad \lambda \mapsto (\lambda I - T)^{-1}$$

wohldefiniert. Das **Spektrum** von T ist definiert als

$$\sigma(T) := \mathbb{K} \setminus \varrho(T).$$

Genauer zerlegt man $\sigma(T)$ in das

- **Punktspektrum**

$$\sigma_p(T) := \{\lambda \in \mathbb{K} \mid \ker(\lambda I - T) \neq \{0\}\},$$

- **kontinuierliche Spektrum**

$$\sigma_c(T) := \{\lambda \in \mathbb{K} \mid \ker(\lambda I - T) = \{0\} \text{ und } \operatorname{ran}(\lambda I - T) \neq X, \overline{\operatorname{ran}(\lambda I - T)} = X\},$$

- **Residualspektrum**

$$\sigma_r(T) = \{\lambda \in \mathbb{K} \mid \ker(\lambda I - T) = \{0\} \text{ und } \overline{\operatorname{ran}(\lambda I - T)} \neq X\}.$$

Abkürzend schreiben wir $\lambda I - T =: \lambda - T$.

Bemerkung: Es ist offensichtlich $\sigma(T) = \sigma_p(T) \cup \sigma_c(T) \cup \sigma_r(T)$.

(i) In den Übungen wurde gezeigt, dass $\varrho(T)$ offen und $\sigma(T)$ kompakt in $\mathbb{K}$ sind und für den **Spektralradius** $r(T) := \limsup_{m\to\infty} \|T^m\|^{1/m}$ von T gilt $|\lambda| \leq r(T)$ für alle $\lambda \in \sigma(T)$.

(ii) $\lambda \in \sigma_p(T)$ wird **Eigenwert** genannt, denn äquivalent hierzu gibt es ein $x \in X \setminus \{0\}$ mit $Tx = \lambda x$. Dabei heißt ein solches x **Eigenvektor** von T zum Eigenwert λ. $E_\lambda := \ker(\lambda - T)$ heißt **Eigenraum** von T zum Eigenwert λ.

Wie schon in Linearer Algebra bemerkt, kann im Fall $\mathbb{K} = \mathbb{R}$ das Spektrum $\sigma(T)$ durchaus leer sein.

Satz 5.2. *Sei X ein Banachraum über $\mathbb{K}$ und $T \in \mathscr{L}(X)$. Dann ist die Resolvente $R_T : \varrho(T) \to \mathscr{L}(X)$ ist eine analytische Abbildung mit*

$$\|R_T(\lambda)\| \geq \frac{1}{\operatorname{dist}(\lambda, \sigma(T))}.$$

Im Fall $\mathbb{K} = \mathbb{C}$ ist für $X \neq \{0\}$ stets $\sigma(T) \neq \emptyset$ und es gilt

$$\max_{\lambda \in \sigma(T)} |\lambda| = r(T) \leq \|T\|.$$

Beweis. In den Übungen wurde die strikte Positivität der Resolvente sowie deren Analytizität gezeigt, d.h. dass es für jedes $\lambda_0 \in \varrho(T)$ ein $r_0 > 0$ und eine Potenzreihe gibt mit

$$R_T(\lambda) = \sum_{n=0}^{\infty} A_n\, (\lambda - \lambda_0)^n,$$

wobei $A_n \in \mathscr{L}(X)$ für $n \in \mathbb{N}$ ist, und die Reihe konvergiert absolut für $\lambda \in B_{r_0}(\lambda_0)$. Sei also $\mathbb{K} = \mathbb{C}$, dann genügt es aufgrund der Schranke von $\sigma(T)$ die Abschätzung

$$s := \sup_{\lambda \in \sigma(T)} |\lambda| \geq r(T)$$

zu zeigen. Für $\mu \in \mathbb{C}$ mit $|\mu| > s$ zeigen wir $|\mu| \geq r(T)$, was $s = r(T)$ impliziert. Für $|\lambda| > s$ ist aufgrund der Analytizität für beliebiges $\ell \in \mathscr{L}(X)^*$ auch $f_\ell(\lambda) := \ell(R_T(\lambda))$ analytisch in λ. Für $|\lambda| > r(T) \geq 0$ folgt nach Lemma 3.8 die Darstellung

$$R_T(\lambda) = \sum_{n=0}^{\infty} T^n \lambda^{-(n+1)}$$

und damit insbesondere für jedes $\ell \in \mathscr{L}(X)^*$

$$f_\ell(\lambda) = \sum_{n=0}^{\infty} \ell(T^n) \lambda^{-(n+1)}.$$

Nach einem Satz der Funktionentheorie konvergiert die Reihe im größten offenen Kreisring, in welchem f_ℓ analytisch ist. Das heißt sie konvergiert insbesondere für μ und $(\ell(T^n)\mu^{-(n+1)})_{n\in\mathbb{N}}$ ist notwendig eine Nullfolge in $\mathbb{C}$. Da dies für alle $\ell \in \mathscr{L}(X)^*$ gilt, ist $(T^n \mu^{-(n+1)})_{n\in\mathbb{N}}$ schwach konvergent in $\mathscr{L}(X)$ und folglich beschränkt nach Lemma 4.8. Damit folgt für ein $C > 0$

$$\|T^n\|^{1/n} \leq C^{1/n} |\mu|^{(n+1)/n} \to |\mu| \qquad \text{für} \quad n \to \infty,$$

d.h. $|\mu| \geq r(T)$. $\square$

Beispiel 5.3.

(a) Für $\dim(X) < \infty$ ist $\sigma(T) = \sigma_p(T)$, d.h. das Spektrum besteht nur aus Eigenwerten.

(b) Betrachte den Rechtsshift auf ℓ^2, d.h.

$$R : \ell^2 \to \ell^2, \quad (x_1, x_2, x_3, \ldots) \mapsto (0, x_1, x_2, \ldots)$$

sowie den Linksshift

$$L : \ell^2 \to \ell^2, \quad (x_1, x_2, x_3, \dots) \mapsto (x_2, x_3, \dots).$$

Dann gilt $\|R\| = 1 = \|L\|$ und somit $\sigma(R), \sigma(L) \subset \overline{B_1(0)}$. Man macht sich leicht klar, dass $\sigma_p(R) = \emptyset$ gilt und $\sigma_p(L) = B_1(0)$ aufgrund der Konvergenz der geometrischen Reihe. Da das Spektrum abgeschlossen ist, folgt $\sigma(L) = \overline{B_1(0)}$.

(c) Für $X = C([0,1])$ betrachte den Operator $T \in \mathscr{L}(X)$ mit

$$(Tx)(t) := \int_0^t x(s)\,\mathrm{d}s \qquad \forall\, x \in X, t \in [0,1].$$

Nach dem Hauptsatz der Differential- und Integralrechnung ist T injektiv und T hat kein dichtes Bild, da stets $(Tx)(0) = 0$ gilt. Daher ist $0 \in \sigma_r(T)$. Genauer ist $\sigma(T) = \{0\} = \sigma_r(T)$, denn für $\lambda \neq 0$ gibt es genau eine Lösung x der Gleichung $\lambda x - Tx = y$ für $y \in X$. Dies folgt aus einer Anwendung der Variation der Konstanten auf $z := Tx$ mit $z(0) = 0$:

$$z'(t) + \frac{1}{\lambda} z(t) = \frac{1}{\lambda} y(t), \qquad z(0) = 0.$$

Die eindeutige Lösung des Anfangswertproblems ist dann gegeben durch

$$z(t) = \frac{1}{\lambda} \int_0^t \mathrm{e}^{(t-s)/\lambda} y(s)\,\mathrm{d}s$$

und $x(t) = z'(t)$ liefert die Surjektivität. Da für $y = 0$ auch $z = 0$ sowie $x = 0$ folgt, ist $\lambda - T$ bijektiv und $\lambda \in \rho(T)$.
Betrachtet man T auf $Y := \{x \in X \mid x(0) = 0\}$, so ist

$$\operatorname{ran}(T) = \{x \in C^1([0,1]) \mid x(0) = 0 = x'(0)\},$$

aber $\overline{\operatorname{ran}(T)} = Y$ nach dem Satz von Stone-Weierstraß, d.h. $\sigma(T) = \{0\} = \sigma_c(T)$.

Betrachtet man den adjungierten Operator T^* gemäß Definition 3.7, so erhält man neben Korollar 3.24 folgende Eigenschaften

Lemma 5.4. *Seien X, Y, Z Banachräume und für $T \in \mathscr{L}(X,Y)$ die adjungierte Abbildung*

$$T^* : Y^* \to X^* \qquad \text{mit} \qquad (T^* y^*)(x) := y^*(Tx) \qquad \forall\, x \in X, y^* \in Y^*,$$

dann gelten für $S, T \in \mathscr{L}(X,Y)$

(i) $(\alpha S + T)^* = \alpha S^* + T^*$ für alle $\alpha \in \mathbb{K}$.

(ii) $(UT)^* = T^* U^*$ für $U \in \mathscr{L}(Y,Z)$.

(iii) $T^{**} J_X = J_Y T$ für die kanonischen Einbettungen J_X, J_Y in die Bidualräume.

(iv) $T^{-1} \in \mathscr{L}(Y,X)$ existiert genau dann wenn $(T^*)^{-1} \in \mathscr{L}(X^*, Y^*)$ existiert und in diesem Fall gilt $(T^{-1})^* = (T^*)^{-1}$.

Beweis. Siehe Übungen für (i)-(iii). Ist in (iv) T invertierbar, so folgt nach (ii)

$$I = I^* = (T^{-1}T)^* = T^*(T^{-1})^* \qquad \text{sowie} \qquad I = (T^{-1})^* T^*,$$

d.h. T^* ist invertierbar und nach Satz 3.38 ist $(T^{-1})^* = (T^*)^{-1} \in \mathscr{L}(X^*, Y^*)$.
Sei nun T^* invertierbar, dann ist nach dem eben Gezeigten T^{**} invertierbar und bildet nach Satz 3.37 abgeschlossene Mengen auf abgeschlossene Mengen ab. Da J_X Isometrie ist, folgt mit (iii), dass auch

$$\operatorname{ran}(J_Y T) = \operatorname{ran}(T^{**} J_X) = T^{**}(\operatorname{ran}(J_X))$$

abgeschlossen in Y^{**}, denn $J_X(X) \subset X^{**}$ ist (folgen)abgeschlossen. Da J_Y injektiv und stetig ist, folgt schließlich

$$\operatorname{ran}(T) = J_Y^{-1}(\operatorname{ran}(J_Y T))$$

ist abgeschlossen in Y. Da T^* injektiv ist, gilt

$$\{0\} = \ker(T^*) = \{f \in Y^* \mid f(y) = 0 \quad \forall\, y \in \operatorname{ran}(T)\}.$$

Nach Korollar 3.23 folgt dann $\operatorname{ran}(T) = \overline{\operatorname{ran}(T)} = Y$ und T ist surjektiv. Die Injektivität von T^{**}, J_X sowie J_Y impliziert die von T, denn nach (iii) ist

$$\{0\} = \ker(T^{**} J_X) = \ker(J_Y T) = \ker(T).$$

Nach dem Satz von der inversen Abbildung ist dann wiederum T^{-1} stetig. $\square$

Die letzte Eigenschaft impliziert daher sofort $\sigma(T) = \sigma(T^*)$. Neben der Adjungierten in einem Banachraum, gibt es noch den Spezialfall eines Hilbertraums $(H, (\cdot,\cdot))$, wobei wir nach dem Riesz'schen Darstellungssatz 3.26 H mit H^* identifizieren können vermöge des antilinearen isometrischen Isomorphismus

$$\phi : H \to H^*, \quad (\phi(u))\,(v) := (u, v) \qquad \forall\, u, v \in H.$$

Definition 5.5. Sei $(H, (\cdot,\cdot))$ ein Hilbertraum über $\mathbb{K}$ und $T \in \mathscr{L}(H)$, dann definiert

$$T^\dagger := \phi^{-1} T^* \phi \in \mathscr{L}(H)$$

die **Hilbertraum-Adjungierte**, welche durch die Bedingung

$$(T^\dagger x, y) = (x, Ty) \qquad \forall\, x, y \in H$$

in H charakterisiert ist. $T \in \mathscr{L}(H)$ heißt **selbstadjungiert**, falls $T^\dagger = T$ gilt bzw.

$$(Tx, y) = (x, Ty) \qquad \forall\, x, y \in H.$$

$T \in \mathscr{L}(H)$ heißt **normal**, falls $T^\dagger T = T T^\dagger$ gilt.

Satz 5.6. *Sei X ein Banachraum, H ein Hilbertraum und $T \in \mathscr{L}(X)$ bzw. $T \in \mathscr{L}(H)$. Dann ist $\sigma(T) = \sigma(T^*)$ und $\sigma(T^\dagger) = \{\overline{\lambda} \mid \lambda \in \sigma(T)\}$. Außerdem gelten die Aussagen von Lemma 5.4 (ii) - (iv), wobei anstatt (i)*

$$(\alpha S + T)^\dagger = \overline{\alpha} S^\dagger + T^\dagger \qquad \forall\, \alpha \in \mathbb{K}$$

gilt sowie $T^{\dagger\dagger} = T$ in (iii). Außerdem gilt $\sigma_r(T) = \sigma_p(T^*)$ sowie $\sigma_r(T) = \{\overline{\lambda} \mid \lambda \in \sigma_p(T^\dagger)\}$.

Beweis. Siehe Übungen. Analog zum Beweis von Lemma 5.4 (iv) beachte für H die Identität
$$((\lambda - T)^{-1})^\dagger = ((\lambda - T)^\dagger)^{-1} = (\overline{\lambda} - T^\dagger)^{-1} \qquad \forall\, \lambda \in \varrho(T).$$
Sei nun $\lambda \in \sigma_r(T)$, so ist $\operatorname{ran}(\lambda - \mathrm{T})$ ein echter Unterraum von X und nach Korollar 3.23 gibt es ein $f \in X^* \backslash \{0\}$ mit $f|_{\operatorname{ran}(\lambda - T)} = 0$. Per Definition von $(\lambda - T)^*$ folgt $f \in \ker(\lambda - T)^*$. Für $T^\dagger$ betrachte man $\phi^{-1} f \in H \setminus \{0\}$ mit $(\phi^{-1} f, y) = 0$ für alle $y \in \operatorname{ran}(\lambda - T)$ und die Aussage folgt aus dem ersten Teil dieses Satzes.
Ist umgekehrt $f \in \ker(\lambda - T^*) \setminus \{0\}$, so ist $f(\operatorname{ran}(\lambda - T)) = 0$. Für $U := \operatorname{ran}(\lambda - T)$ ist folglich, $f|_U = 0 \in U^*$. Dann ist U nicht dicht in X, denn angenommen, es gelte $\overline{U} = X$, so gibt es nach Satz 3.22 eine stetige Fortsetzung von $f|_U \in U^*$ auf ganz X, wobei diese nach Lemma 1.28 eindeutig ist, d.h. $f = 0$ im Widerspruch zur Annahme. Somit ist $\overline{\operatorname{ran}(\lambda - T)} \neq X$ und $\lambda \in \sigma_r(T)$. Der Rest folgt analog. □

Beispiel 5.7. Vergleiche Beispiel 5.3 und betrachte den Rechtsshift auf ℓ^2, d.h.
$$R : \ell^2 \to \ell^2, \quad (x_1, x_2, x_3, \dots) \mapsto (0, x_1, x_2, \dots).$$
Dann ist der entsprechende Linksshift der adjungierte Operator $L = R^\dagger$,
$$L : \ell^2 \to \ell^2, \quad (x_1, x_2, x_3, \dots) \mapsto (x_2, x_3, \dots),$$
und es folgt $\sigma(R) = \sigma(L) = \overline{B_1(0)}$. Wäre nun $\lambda \in \sigma_r(L) \cap \partial B_1(0)$, so folgte mit Satz 5.6, dass $\overline{\lambda} \in \sigma_p(R) = \emptyset$ wäre, im Widerspruch. Daher ist $\sigma_r(L) = \emptyset$ und $\sigma_c(L) = \partial B_1(0)$. Wegen $L^\dagger = R$ folgt ähnlich $\sigma_r(R) = B_1(0)$ sowie $\sigma_c(R) = \partial B_1(0)$.

5.1 Kompakte Operatoren

Im Folgenden betrachten wir einen Unterraum der stetigen linearen Operatoren, für welchen eine relativ einfache Charakterisierung des Spektrums möglich ist.

Definition 5.8. Seien X, Y Banachräume und $B_X := \{x \in X \mid \|x\| < 1\}$ die offene Einheitskugel in X. Ein Operator $T \in \mathscr{L}(X, Y)$ heißt **kompakt**, falls $\overline{T(B_X)}$ kompakt in Y ist. Man bezeichnet dann $\mathscr{K}(X) := \mathscr{K}(X, X)$ bzw. allgemein
$$\mathscr{K}(X, Y) := \{T \in \mathscr{L}(X, Y) \mid T \text{ ist kompakt}\}.$$

Bemerkung: Da Y vollständig ist, lässt sich nach Satz 1.25 und den Übungen die Kompaktheit von T äquivalent definieren durch eine der folgenden Aussagen:

(i) $M \subset X$ beschränkt $\Rightarrow T(M)$ ist präkompakt in Y.

(ii) Für jede beschränkte Folge $(x_n)_{n \in \mathbb{N}}$ in X besitzt $(Tx_n)_{n \in \mathbb{N}}$ in Y eine konvergente Teilfolge.

Satz 5.9. *Seien X, Y, Z Banachräume über $\mathbb{K}$. Dann gelten*

(i) $\mathscr{K}(X, Y)$ ist ein abgeschlossener Unterraum von $\mathscr{L}(X, Y)$.

(ii) Seien $T \in \mathscr{K}(X, Y)$, $S \in \mathscr{L}(Y, Z)$ und $R \in \mathscr{L}(Z, X)$. Dann ist $ST \in \mathscr{K}(X, Z)$ und $TR \in \mathscr{K}(Z, Y)$.

(iii) $T \in \mathscr{K}(X,Y)$ *ist* ***vollstetig***, *d.h.*

$$x_n \rightharpoonup x \quad in \quad X \qquad \Rightarrow \qquad Tx_n \to Tx \quad in \quad Y \qquad für \quad n \to \infty.$$

Ist X ein reflexiver Banachraum, so ist auch jeder vollstetige Operator kompakt.

Beweis. (i) Seien $T, S \in \mathscr{K}(X,Y)$, $\lambda \in \mathbb{K}$ und $(x_n)_{n\in\mathbb{N}}$ eine beschränkte Folge in X. Dann ist $(Tx_{n_j})_{j\in\mathbb{N}}$ konvergent für eine Teilfolge $(x_{n_j})_{j\in\mathbb{N}}$. Wählt man aus der weiterhin beschränkten Folge $(x_{n_j})_{j\in\mathbb{N}}$ eine gleich bezeichnete Teilfolge aus, so konvergiert auch $(Sx_{n_j})_{j\in\mathbb{N}}$. Dann ist die Linearkombination $((T+\lambda S)(x_{n_j}))_{j\in\mathbb{N}}$ ebenso konvergent, also ist $\mathscr{K}(X,Y)$ ein linearer Unterraum.
Sei nun $(T_n)_{n\in\mathbb{N}}$ eine Folge in $\mathscr{K}(X,Y)$ mit $T_n \to T$ in $\mathscr{L}(X,Y)$. Sei $\varepsilon > 0$ fest und $n \in \mathbb{N}$ groß genug, sodass $\|T_n - T\| < \varepsilon$. Da T_n kompakt ist, gibt es $x_1, \dots, x_m \in B_X$, sodass

$$T_n(B_X) \subset \bigcup_{k=1}^{m} B_\varepsilon(T_n(x_k)),$$

d.h. für jedes $x \in B_X$ gibt es ein $j \in \{1, \dots, m\}$ mit $\|T_n x - T_n x_j\| < \varepsilon$. Für $T \in \mathscr{K}(X,Y)$ zeigen wir, dass $T(B_X)$ präkompakt ist und durch $3\,\varepsilon$-Kugeln um $Tx_1, \dots, Tx_m$ eine Überdeckung von $T(B_X)$ gegeben ist. In der Tat, für $x \in B_X$ wähle man ein $j \in \{1, \dots, m\}$ wie oben und erhält

$$\begin{aligned}
\|Tx - Tx_j\| &\le \|Tx - T_n x\| + \|T_n x - T_n x_j\| + \|T_n x_j - Tx_j\| \\
&\le 2\,\|T_n - T\| + \varepsilon \\
&< 3\,\varepsilon.
\end{aligned}$$

(ii) Sei $(x_n)_{n\in\mathbb{N}}$ eine beschränkte Folge in X. Dann existiert eine Teilfolge $(x_{n_j})_{j\in\mathbb{N}}$, sodass $(Tx_{n_j})_{j\in\mathbb{N}}$ in Y konvergiert. Dann ist $(STx_{n_j})_{j\in\mathbb{N}}$ für jedes stetige $S \in \mathscr{L}(Y,Z)$ konvergent in Z. Analog zeigt man $TR \in \mathscr{K}(Z,Y)$.

(iii) Es gelte $x_n \rightharpoonup x$. Es folgt zunächst $y_n := Tx_n \rightharpoonup Tx =: y$, denn für jedes $L \in Y^*$ ist $LT \in X^*$. Konvergiert nun $(y_n)_{n\in\mathbb{N}}$ nicht gegen y, so existieren ein $\varepsilon > 0$ und eine Teilfolge $(y_{n_k})_{k\in\mathbb{N}}$ mit $\|y_{n_k} - y\|_Y \ge \varepsilon$ für alle $k \in \mathbb{N}$. Aus der schwachen Konvergenz folgt nach Lemma 4.8 $\|x_n\|_X \le c$, d.h. $(y_{n_k})_{k\in\mathbb{N}}$ besitzt aufgrund der Kompaktheit von T eine konvergente Teilfolge. Wegen $y_n \rightharpoonup y$ muss der Grenzwert dieser Teilfolge ebenso y sein, ein Widerspruch.
Der zweite Teil der Aussage folgt aus Satz 4.28, da eine beschränkte Folge in X eine schwach konvergente Teilfolge besitzt.

□

Beispiel 5.10.

(a) Man sagt, $T \in \mathscr{L}(X,Y)$ hat einen **endlichen Rang**, falls $\dim T(X) < \infty$. Dann ist $T \in \mathscr{K}(X,Y)$, denn falls $(x_n)_{n\in\mathbb{N}}$ eine beschränkte Folge in X ist, so ist $(Tx_n)_{n\in\mathbb{N}}$ beschränkt in $\operatorname{ran}(T)$ und Satz 1.37 liefert die Aussage. Beispielsweise ist jedes Funktional $f \in X^*$ kompakt, d.h. $f \in \mathscr{K}(X, \mathbb{K})$.

(b) Sei $I = [0,1]$ und $K : I \times I \to \mathbb{K}$ stetig. Für $f \in C(I)$ setze man

$$(Tf)(x) := \int_0^x K(x,y)\, f(y)\; \mathrm{d}y \qquad \forall\, x \in I$$

(K heißt **Integralkern** von T). Dann ist nach dem Satz von Arzelà-Ascoli aus den Übungen $T \in \mathscr{K}(C(I))$.

(c) Für den Links- bzw. Rechtsshift aus Beispiel 5.3 gilt $LR = I$ und nach Satz 5.9 (ii) sind weder L noch R kompakt, denn I ist genau dann kompakt, wenn X endlich dimensional ist (siehe Satz 1.37).

Für den adjungierten Operator erhält man

Satz 5.11. *Seien X, Y Banachräume und $T \in \mathscr{L}(X,Y)$ bzw. H ein Hilbertraum und $T \in \mathcal{L}(H)$. Dann gelten*

$$T \in \mathscr{K}(X,Y) \qquad \Leftrightarrow \qquad T^* \in \mathscr{K}(Y^*, X^*)$$

beziehungsweise

$$T \in \mathscr{K}(H) \qquad \Leftrightarrow \qquad T^\dagger \in \mathscr{K}(H).$$

Beweis. Sei $T \in \mathscr{L}(X,Y)$ kompakt, d.h. $K := \overline{T(B_X)} \subset Y$ ist kompakt. Für $f \in Y^*$ folgt aufgrund der Stetigkeit

$$\|T^* f\| = \sup_{x \in B_X} |(T^* f)(x)| = \sup_{x \in B_X} |f(Tx)| = \sup_{y \in K} |f(y)|$$

und definiert man $Rf := f|_K$ als Einschränkung von f auf K, so ist $\|T^* f\| = \|Rf\|_{C(K)}$. Außerdem ist $Rf \in C_{\mathbb{K}}(K)$ und

$$\sup_{y \in K} |f(y)| \leq \|f\| \sup_{y \in K} \|y\|_Y \leq \|f\| \|T\|.$$

Damit ist die Abbildung $R : Y^* \to C_{\mathbb{K}}(K)$ linear und beschränkt mit $\|R\| \leq \|T\|$. Die Abbildung $T^* f \mapsto Rf$ definiert also einen isometrischen Isomorphismus zwischen $T^*(B_{Y^*})$ und $R(B_{Y^*}) \subset C(K)$. Um zu zeigen, dass $T^*(B_{Y^*})$ präkompakt ist, genügt es also zu zeigen, dass $R(B_{Y^*})$ präkompakt in $C(K)$ ist. Dazu werden wir den Satz von Arzelà-Ascoli für kompakte Räume benutzen, was analog zu den Übungen bewiesen werden kann. Die Beschränktheit von $R(B_{Y^*})$ folgt aus $\|R\| \leq \|T\|$. Die gleichgradige Stetigkeit folgt für $y_1, y_2 \in K$ aus

$$|Rf(y_1) - Rf(y_2)| = |f(y_1 - y_2)| \leq \|f\| \, \|y_1 - y_2\| \leq \|y_1 - y_2\|,$$

wobei die rechte Seite nicht von $f \in B_{Y^*}$ abhängt.
Sei nun $T^* \in \mathscr{K}(Y^*, X^*)$. Aus dem bereits Gezeigten folgt $T^{**} \in \mathscr{K}(X^{**}, Y^{**})$, d.h. $T^{**}(B_{X^{**}})$ ist präkompakt in Y^{**}. Durch die kanonische Einbettung $J : X \to X^{**}$ lässt sich B_X als eine Teilmenge von $B_{X^{**}}$ auffassen. Nach Lemma 5.4 (iii) ist

$$J_Y(T(B_X)) = T^{**}(J_X(B_X)) \subset T^{**}(B_{X^{**}})$$

als Teilmenge präkompakt. Da J_Y isometrisch ist, folgt die Präkompaktheit von $T(B_X)$. Somit ist T kompakt.
Per Definition von $T^\dagger$ und Satz 5.9 (ii) folgt die zweite Äquivalenz. □

Im Folgenden betrachten wir zu $T \in \mathscr{L}(X), \lambda \in \mathbb{K}$ und $y \in X$ die Lösbarkeit der Gleichung

$$\lambda x - Tx = y,$$

d.h. wir suchen alle Lösungen $x \in X$ hierzu. Dieses Problem ist eng verknüpft mit dem **Eigenwertproblem** für $y = 0$ wie wir sehen werden. Ist $\lambda \in \varrho(T)$, so existiert genau eine

Lösung $x \in X$ dieser Gleichung. Ist jedoch $\lambda \in \sigma_p(T)$, so ist eine Lösung $x \in X$ (existiert z.B. für $y \in \operatorname{ran}(\lambda - T)$) nicht eindeutig bestimmt, denn jedes Element der Form $x + \overline{x}$ mit $\overline{x} \in \ker(\lambda - T)$ ist auch eine Lösung. Die folgende Klasse von Operatoren besitzt dabei nur eine überschaubare Anzahl von Freiheitsgraden für x sowie y.

Definition 5.12. Seien X, Y Banachräume. $T \in \mathscr{L}(X,Y)$ heißt **Fredholm-Operator**, falls gilt:

(i) $\dim(\ker(T)) < \infty$,

(ii) $\operatorname{ran}(T)$ ist abgeschlossen,

(iii) $\operatorname{codim}(\operatorname{ran}(T)) < \infty$.

Der **Index** von T ist definiert durch

$$\operatorname{ind}(T) := \dim(\ker(T)) - \operatorname{codim}(\operatorname{ran}(T)).$$

Bemerkung: Die Kodimension $\operatorname{codim}(Z)$ eines abgeschlossenes Unterraums $Z \subset X$ ist die Dimension des Quotientenraums $X/_\sim$ (wobei $x \sim y \Leftrightarrow x - y \in Z$). Hat Z die Kodimension $m < \infty$, so gibt es linear unabhängige $x_1, \dots, x_m \in X$ mit

$$X = Z \oplus \operatorname{span}\{x_1, \dots, x_m\}.$$

Satz 5.13. *Sei X ein Banachraum und $T \in \mathscr{K}(X)$. Dann ist $S = I - T$ ein Fredholm-Operator mit Index* 0.

Beweis. Der Beweis erfolgt in mehreren Schritten und wurde in der Vorlesung ausgelassen. Insbesondere folgt die bemerkenswerte Eigenschaft wie im endlich dimensionalen Fall:

$$\ker(S) = \{0\} \Leftrightarrow \operatorname{ran}(S) = X.$$

(i) Sei $Sx = 0$. Dann ist $Tx = x$, also $B_1(0) \cap \ker(S) \subset T(B_1(0))$. Da T kompakt ist, ist die offene Einheitskugel in $\ker(S)$ präkompakt. Das ist aber nur dann möglich, falls $\dim(\ker(S)) < \infty$.

(ii) Sei $x \in \overline{\operatorname{ran}(S)}$ und entsprechend $Sx_n \to x$ für $n \to \infty$. Man kann annehmen, dass

$$\|x_n\| \leq 2\, d_n \qquad \text{mit} \quad d_n := \operatorname{dist}(x_n, \ker(S)).$$

(Falls dies nicht der Fall ist, so findet man ein $a_n \in \ker(S)$ mit $\|x_n - a_n\| \leq 2\, d_n$ und man ersetze x_n durch $\tilde{x}_n := x_n - a_n$ mit $S\tilde{x}_n = Sx_n$.)
Wir zeigen, dass $(d_n)_{n\in\mathbb{N}}$ beschränkt ist. Dazu nehme an, dass es eine Teilfolge $d_{n_j} \to \infty$ für $j \to \infty$ gibt, und setze

$$y_j := \frac{x_{n_j}}{d_{n_j}}.$$

Dann ist $(y_j)_{j\in\mathbb{N}}$ eine beschränkte Folge in X und aufgrund der Kompaktheit gibt es eine Teilfolge $(Ty_{j_i})_{i\in\mathbb{N}}$ und ein $y \in X$ mit $Ty_{j_i} \to y$ für $i \to \infty$. Dies impliziert

$$y_{j_i} = Sy_{j_i} + Ty_{j_i} \to y \qquad \text{für} \quad i \to \infty.$$

Da $(Sy_j)_{j\in\mathbb{N}}$ eine Nullfolge in Y ist und S stetig ist, muss auch $Sy = 0$ gelten. Damit liegt $y \in \ker(S)$ und somit

$$\|y_j - y\| \geq \operatorname{dist}(y_j, \ker(S)) = \operatorname{dist}\left(\frac{x_{n_j}}{d_{n_j}}, \ker(S)\right) = 1,$$

was einen Widerspruch zu $y_{j_i} \to y$ für $i \to \infty$ darstellt.
Mit $(d_n)_{n\in\mathbb{N}}$ ist daher auch $(x_n)_{n\in\mathbb{N}}$ beschränkt. Die Kompaktheit von T liefert wieder eine Teilfolge und $z \in X$, sodass $Tx_{n_j} \to z$ für $j \to \infty$. Daraus folgt

$$x = \lim_{j\to\infty} Sx_{n_j} = \lim_{j\to\infty} S\left(Sx_{n_j} + Tx_{n_j}\right) = S(x+z),$$

also $x \in \operatorname{ran}(S)$.

(iii) Wir zeigen $\operatorname{ind}(S) = 0$ in 3 Schritten:

- $\ker S = \{0\}$ impliziert $\operatorname{ran}(S) = X$, denn gibt es ein $x \in X \setminus \operatorname{ran}(S)$, so ist $S^n x \in \operatorname{ran}(S^n) \setminus \operatorname{ran}(S^{n+1})$. In der Tat, falls $S^n x = S^{n+1} y$ für ein $y \in X$, so gilt $S^n(x - Sy) = 0$ und die Injektivität liefert $x = Sy$, ein Widerspruch. Außerdem ist $\operatorname{ran}(S^{n+1})$ nach (ii) abgeschlossen, denn

  $$S^{n+1} = (I-T)^{n+1} = I + \sum_{k=1}^{n+1} \binom{n+1}{k} (-T)^k =: I - K$$

  wobei K kompakt ist nach Satz 5.9. Deswegen ist $\operatorname{dist}(S^n x, \operatorname{ran}(S^{n+1})) > 0$ und es gibt ein $a_{n+1} \in \operatorname{ran}(S^{n+1})$ mit

  $$0 < \|S^n x - a_{n+1}\| \leq 2 \operatorname{dist}(S^n x, \operatorname{ran}(S^{n+1})).$$

  Setze

  $$x_n := \frac{S^n x - a_{n+1}}{\|S^n x - a_{n+1}\|} \in \operatorname{ran}(S^n).$$

  Es gilt nun $\operatorname{dist}(x_n, \operatorname{ran}(S^{n+1})) \geq \frac{1}{2}$, da für $y \in \operatorname{ran}(S^{n+1})$ gilt:

  $$\|x_n - y\| = \frac{\|S^n x - (a_{n+1} + \|S^n x - a_{n+1}\| y)\|}{\|S^n x - a_{n+1}\|} \geq \frac{\operatorname{dist}(x_n, \operatorname{ran}(S^{n+1}))}{\|S^n x - a_{n+1}\|} \geq \frac{1}{2}.$$

  Per Definition von $S = I - T$ ist dann für $m > n$

  $$\|Tx_n - Tx_m\| = \|x_n - (Sx_n + x_m - Sx_m)\| \geq \operatorname{dist}(x_n, \operatorname{ran}(S^{n+1})) \geq \frac{1}{2}.$$

  So besitzt $(Tx_n)_{n\in\mathbb{N}}$ keine konvergente Teilfolge, im Widerspruch zur Kompaktheit von T, denn $\|x_n\| \leq 1$ ist beschränkt.

- $\operatorname{codim}(\operatorname{ran}(S)) \leq \dim(\ker(S))$, denn:
  Nach (i) gibt es eine Basis $\{x_1, \ldots, x_n\}$ von $\ker(S)$. Wäre die Behauptung falsch, so gäbe es linear unabhängige $y_1, \ldots, y_n \in X$, sodass

  $$\operatorname{span}\{y_1, \ldots, y_n\} \oplus \operatorname{ran}(S)$$

ein echter Teilraum von X ist. Die endlich dimensionalen und damit abgeschlossenen Unterräume span $\{x_j \mid j \neq i\}$ liefern nach Korollar 3.23 $f_1, \dots, f_n \in X^*$ mit $f_i(x_l) = \delta_{il}$ für $i, l \in \{1, \dots, n\}$. Man definiere

$$\tilde{T}x := Tx + \sum_{k=1}^{n} f_k(x) y_k \qquad \forall\, x \in X.$$

Dann ist $\tilde{T}$ als Summe kompakter Operatoren wiederum kompakt. Definiere $\tilde{S} = I - \tilde{T}$, so ist $\ker(\tilde{S}) = \{0\}$, denn:

$$\tilde{S}x = 0 \qquad \Leftrightarrow \qquad Sx - \sum_{k=1}^{n} f_k(x) y_k = 0$$

impliziert wegen span $\{y_1, \dots, y_n\} \cap \operatorname{ran}(S) = \{0\}$ (y_i linear unabhängig)

$$Sx = 0 \qquad \text{und} \qquad f_k(x) = 0 \quad \forall\, k = 1, \dots, n.$$

Wegen $x \in \ker(S)$ folgt $x = \sum_{k=1}^{n} \alpha_k x_k$ und somit $0 = f_k(x) = \alpha_k$ für alle $k = 1, \dots, n$, d.h. $x = 0$.
Das bereits Gezeigte impliziert nun $\operatorname{ran}(\tilde{S}) = X$ und die Konstruktion von $\tilde{S}$ zeigt $X = \operatorname{ran}(\tilde{S}) \subset \operatorname{span}\{y_1, \dots, y_n\} \oplus \operatorname{ran}(S)$, ein Widerspruch zur Annahme.

- $\dim(\ker(S)) \leq \operatorname{codim}(\operatorname{ran}(S))$, denn:
  Sei $S^* = I - T^* \in \mathscr{L}(X^*)$ die zu S adjungierte Abbildung, welche nach Satz 5.11 wieder kompakt ist. Somit impliziert der vorherige Beweis

$$\operatorname{codim}(\operatorname{ran}(S^*)) \leq \dim(\ker(S^*)).$$

Dann genügt es die beiden Ungleichungen zu zeigen:

$$\dim(\ker(S)) \leq \dim(\ker(S^{**})) \tag{5.1}$$
$$\dim(\ker(S^*)) = \operatorname{codim}(\operatorname{ran}(S)) \tag{5.2}$$

In der Tat würde dann $\dim(\ker(S^{**})) \leq \operatorname{codim}(\operatorname{ran}(S^*))$ folgen und die Ungleichungen liefern die Behauptung und sogar Gleichheit in (5.1).
Um (5.1) zu zeigen, verwende man Lemma 5.4 (iii) mit $S^{**}J_X = J_X S$. Da J_X injektiv ist, erhält man $\ker(S) = \ker(J_X S) = \ker(S^{**}J_X)$ und

$$\dim(\ker(S)) = \dim(\ker(S^{**}J_X)) \leq \dim(\ker(S^{**}))$$

Um (5.2) zu beweisen, bemerke man wie im Beweis von Lemma 5.4

$$\ker(S^*) = \{f \in X^* \mid f(y) = 0 \quad \forall\, y \in \operatorname{ran}(S)\}$$

und wir zeigen $\dim(\ker(S^*)) = \operatorname{codim}(\operatorname{ran}(S))$. Da $n := \operatorname{codim}(\operatorname{ran}(S)) < \infty$ seien $x_1, \dots, x_n \in X$ linear unabhängig mit

$$X = \operatorname{ran}(S) \oplus \operatorname{span}\{x_1, \dots, x_n\}.$$

Zum abgeschlossenen Unterraum $X_i := \operatorname{ran}(S) \oplus \operatorname{span}\{x_j \mid j \neq i\}$ wählen wir nach Korollar 3.23 $f_i \in X^*$ mit $f_i(x_l) = \delta_{il}$ für $i, l \in \{1, \dots, n\}$ und $f_i|_{\operatorname{ran}(S)} = 0$.

Dann sind $f_1, \dots, f_n \in \ker(S^*)$ linear unabhängig, d.h. $\dim(\ker(S^*)) \geq n$. Ist umgekehrt $f \in \ker(S^*)$ ein Funktional, so folgt mit

$$x = Sx_0 + \sum_{k=1}^{n} \alpha_k x_k \in X \qquad \text{für} \quad x_0 \in X, \alpha_k \in \mathbb{K}$$

auch $\dim(\ker(S^*)) \leq n$, denn f ist Linearkombination der $f_1, \dots, f_n$:

$$f(x) = \sum_{k=1}^{n} \alpha_k f(x_k) = \sum_{j,k=1}^{n} \alpha_k f(x_j) f_j(x_k) = \left(\sum_{j=1}^{n} f(x_j) f_j \right)(x).$$

□

Satz 5.14 (Fredholm'sche Alternative)**.** *Sei X ein Banachraum, $T \in \mathscr{K}(X)$ kompakt und $S = I - T$. Dann gilt genau eine der folgenden Aussagen:*

(i) Die Gleichung $Sx = y$ besitzt für jedes $y \in X$ genau eine Lösung und $S^{-1} \in \mathscr{L}(X)$.

(ii) Die Gleichung $Sx = 0$ hat eine nichttriviale Lösung und $\operatorname{ran}(S) \neq X$.

Beweis. Man unterscheidet zwei Fälle

(i) $\ker(S) = \{0\}$. Dann ist $\operatorname{ran}(S) = X$ und die Gleichung $Sx = y$ besitzt eine eindeutige Lösung für jedes $y \in X$, denn $S : X \to X$ ist bijektiv. Aus dem Satz vom inversen Operator folgt $S^{-1} \in \mathscr{L}(X)$.

(ii) $\ker(S) \neq \{0\}$. Dann hat $Sx = 0$ eine nichttriviale Lösung $x \neq 0$. Dann ist

$$\operatorname{codim}(\operatorname{ran}(S)) \geq 1$$

und $\operatorname{ran}(S) \neq X$.

□

5.2 Spektralsätze für kompakte Operatoren

Die Hauptsätze für kompakte Operatoren spiegeln vor allem für normale Operatoren die Hauptachsentransformation bzw. Diagonalisierung für Matrizen in endlich dimensionalen Räumen wider und lassen eine sehr präzise Beschreibung des Spektrums zu. Wie man sich leicht mit Satz 1.37 klar macht, spielt $\lambda = 0$ eine besondere Rolle im Spektrum, denn ist X unendlich dimensional, so folgt $0 \in \sigma(T)$. Im Übrigen erhält man

Satz 5.15 (Spektralsatz für kompakte Operatoren, Riesz-Schander)**.** *Sei X ein Banachraum und $T \in \mathscr{K}(X)$. Dann besteht $\sigma(T) \setminus \{0\}$ aus höchstens abzählbar vielen Eigenwerten und 0 ist einzig möglicher Häufungspunkt von $\sigma(T)$, falls $\sigma(T) \setminus \{0\}$ unendlich ist. Außerdem gilt für $\lambda \in \sigma(T) \setminus \{0\}$*

(i) $\dim(\ker(\lambda - T)) < \infty$ *und*

$$1 \leq n_\lambda := \max\left\{ n \in \mathbb{N} \mid \ker(\lambda - T)^{n-1} \neq \ker(\lambda - T)^n \right\} < \infty.$$

n_λ *heißt die* ***Ordnung*** *von λ und* $\dim(\ker(\lambda - T))$ *die* ***Vielfachheit*** *von λ.*

(ii) Es ist
$$X = \operatorname{ran}(\lambda - T)^{n_\lambda} \oplus \ker(\lambda - T)^{n_\lambda}.$$
Beide Unterräume sind abgeschlossen und invariant bzgl. T und es gilt
$$\dim(\ker(\lambda - T)^{n_\lambda}) < \infty.$$

(iii) Es ist
$$\sigma(T|_{\operatorname{ran}(\lambda - T)^{n_\lambda}}) = \sigma(T) \setminus \{\lambda\}.$$

(iv) Sei P_λ die Projektion auf $\ker(\lambda - T)^{n_\lambda}$ gemäß (ii), dann gilt
$$P_\lambda P_\mu = \delta_{\lambda,\mu} P_\lambda \qquad \forall\, \lambda, \mu \in \sigma(T) \setminus \{0\}.$$

Beweis. Für $\lambda \notin \sigma_p(T), \lambda \neq 0$ ist $\ker(I - \lambda^{-1}T) = \{0\}$ und nach Satz 5.13 mit kompaktem $\lambda^{-1}T$ auch $\operatorname{ran}(I - \lambda^{-1}T) = X$ und $\lambda \in \varrho(T)$. Dies beweist $\sigma(T) \setminus \{0\} \subset \sigma_p(T)$.
Ist $\sigma(T) \setminus \{0\}$ unendlich, wählen wir paarweise verschiedene $\lambda_n \in \sigma(T) \setminus \{0\}$ für $n \in \mathbb{N}$ mit jeweiligen Eigenvektoren $e_n \neq 0$ zu λ_n. Dann sind die Eigenvektoren linear unabhängig wie man analog zur Linearen Algebra zeigt. Daher ist
$$X_n := \{e_1, \ldots, e_n\} \qquad \text{mit} \quad X_n \subsetneq X_{n+1}, \quad \dim X_n = n.$$
Nach Lemma 1.36 lässt sich jeweils ein $x_n \in X_n$ wählen mit
$$\|x_n\| = 1 \qquad \text{und} \qquad \operatorname{dist}(x_n, X_{n-1}) \geq \frac{1}{2}.$$
Dabei gibt es ein $\alpha_n \in \mathbb{K}, \tilde{x}_n \in X_{n-1}$ mit $x_n = \alpha_n e_n + \tilde{x}_n$ und es folgt
$$Tx_n - \lambda_n x_n = T\tilde{x}_n - \lambda_n \tilde{x}_n \in X_{n-1}.$$
Für $m < n$ erhält man mit $y_m := x_m/\lambda_m \in X_{n-1}$
$$\|T(y_n) - T(y_m)\| = \left\| x_n + \frac{1}{\lambda_n}(Tx_n - \lambda_n x_n) - Ty_m \right\| \geq \frac{1}{2}.$$
Folglich besitzt $(Ty_n)_{n\in\mathbb{N}}$ keinen Häufungspunkt und aufgrund der Kompaktheit von T kann $(y_n)_{n\in\mathbb{N}}$ keine beschränkte Teilfolge enthalten. Deshalb divergiert $\|y_n\| \to \infty$ bzw.
$$|\lambda_n| = \frac{1}{\|y_n\|} \to 0 \qquad \text{für} \quad n \to \infty.$$
Damit ist 0 einziger Häufungspunkt von $\sigma(T) \setminus \{0\}$ und $\sigma(T) \setminus B_{1/n}(0)$ ist endlich für jedes $n \in \mathbb{N}$, d.h. $\sigma(T) \setminus \{0\}$ abzählbar.
Die Aussagen (ii)-(iv) sind Folgerungen aus Satz 5.13 und können z.B. in H.W. Alt: Funktionalanalysis, Satz 9.9 nachgelesen werden. Es bleibt (i) zu zeigen:
Für $\lambda \in \sigma(T) \setminus \{0\}$ setzen wir $A := \lambda - T$, so ist
$$\{0\} \subset \ker(A) \subset \ldots \subset \ker(A^n) \qquad \forall\, n \in \mathbb{N}.$$
Wir zeigen, dass die Nullräume stationär werden und nehmen daher an
$$\ker(A^{n-1}) \subsetneq \ker(A^n) \qquad \forall\, n \in \mathbb{N}.$$

Analog zum Beweis der vorherigen Aussage wähle man nach Lemma 1.36

$$x_n \in \ker(A^n), \qquad \|x_n\| = 1 \qquad \text{und} \qquad \operatorname{dist}(x_n, \ker(A^{n-1})) \geq \frac{1}{2}.$$

und die Kompaktheit von T liefert einen Widerspruch, denn für $m < n$ ist

$$\|Tx_n - Tx_m\| = |\lambda|\|x_n - \lambda^{-1}(Ax_n + \lambda x_m - Ax_m)\| \geq |\lambda| \operatorname{dist}(x_n, \ker(A^{n-1})) \geq \frac{|\lambda|}{2} > 0.$$

Damit sind zwei aufeinanderfolgende Nullräume $\ker(A^{n-1}) = \ker(A^n)$ für ein $n \in \mathbb{N}$ identisch und man macht sich leicht klar, dass auch alle nachfolgenden identisch sind. Folglich ist die Ordnung n_λ endlich und damit auch die Vielfachheit von λ aufgrund der aufsteigenden Kette. Wegen $\ker(A) \neq \{0\}$ folgt schließlich $n_\lambda \geq 1$. $\square$

Das Spektrum eines kompakten Operators besteht also bis auf möglicherweise 0 aus einer endlichen Menge oder einer Nullfolge. Falls jedoch $\sigma(T) \setminus \{0\} = \emptyset$ liefert dieser Spektralsatz leider keine Aussage über das Spektrum, wie auch das Beispiel 5.3 (c) zeigt. Satz 5.15 (ii) besagt außerdem, dass sich $\lambda - T$ für $\lambda \neq 0$ in zwei einfachere Bestandteile zerlegen lässt, denn zum einen ist

$$(\lambda - T)|_{\operatorname{ran}(\lambda-T)^{n_\lambda}} : \operatorname{ran}(\lambda - T)^{n_\lambda} \to \operatorname{ran}(\lambda - T)^{n_\lambda}$$

ein Isomorphismus aufgrund der direkten Summe und zum anderen gibt es den endlich dimensionalen Anteil (der durch eine Matrix beschrieben werden kann) $(\lambda - T)|_{\ker(\lambda-T)^{n_\lambda}}$.

Im Fall eines kompakten normalen Operators auf einem Hilbertraum lässt sich der Operator besonders einfach darstellen mithilfe der Eigenwerte. Zunächst ergibt eine genauere Betrachtung von Satz 5.2 für das Spektrum

Lemma 5.16. *Sei $H \neq \{0\}$ ein Hilbertraum und $T \in \mathscr{L}(H)$. Dann gelten*

(i) Ist $\mathbb{K} = \mathbb{C}$ und T normal, so gibt es ein $\lambda \in \sigma(T)$ mit $|\lambda| = \|T\|$.

(ii) Ist $\mathbb{K} = \mathbb{R}$ und $T \in \mathscr{K}(H)$ selbstadjungiert, so gibt es ein $\lambda \in \sigma_p(T)$ mit $|\lambda| = \|T\|$.

Beweis. (i) Nach Satz 5.2 genügt es die Ungleichung $r(T) \geq \|T\|$ zu zeigen. Dies folgt jedoch aus den Übungen.

(ii) O.B.d.A. sei $\|T\| > 0$, sonst ist die Aussage trivial. Zunächst ist für selbstadjungiertes $T \in \mathscr{L}(H)$ stets $\|T\| = \sup_{x \in k_H} |(Tx, x)|$ wie die Übungen zeigen. Damit gibt es eine Maximalfolge $(x_n)_{n\in\mathbb{N}}$ in k_H mit $|(Tx_n, x_n)| \to \|T\|$ für $n \to \infty$. Die Kompaktheit von T impliziert, dass es eine konvergente Bildfolge der beschränkten Folge $(x_n)_{n\in\mathbb{N}}$ gibt mit $Tx_{n_j} \to y$ für $j \to \infty$. Für eine (gleich bezeichnete) Teilfolge gibt es außerdem ein $\alpha \in \mathbb{R}$ mit $|\alpha| = \|T\|$ und

$$\alpha = \lim_{j\to\infty} (Tx_{n_j}, x_{n_j}).$$

Nun ist wegen $\|x_{n_j}\| \leq 1$ für $j \to \infty$

$$\begin{aligned}
\|Tx_{n_j} - \alpha x_{n_j}\|^2 &= \|Tx_{n_j}\|^2 - 2\alpha(Tx_{n_j}, x_{n_j}) + \alpha^2\|x_{n_j}\|^2 \\
&\leq \|T\|^2 - 2\alpha(Tx_{n_j}, x_{n_j}) + \alpha^2 \\
&\to \|T\|^2 - \alpha^2 = 0.
\end{aligned}$$

Folglich ist $y = \lim_{j\to\infty} \alpha x_{n_j}$ mit $Ty = \alpha y$. Wegen $|\alpha| = \|T\|$ genügt es $y \neq 0$ zu zeigen, sodass $\alpha \in \sigma_p(T)$ gilt. Wäre $y = 0$, so wäre $(Tx_{n_j})_{j\in\mathbb{N}}$ eine Nullfolge und demnach $\alpha = 0$ sowie $\|T\| = 0$ im Widerspruch zur Annahme.

□

Satz 5.17 (Spektralsatz für kompakte normale Operatoren). *Sei $(H, (\cdot,\cdot))$ ein nicht trivialer $\mathbb{K}$-Hilbertraum und $T \in \mathscr{K}(H)$ normal (falls $\mathbb{K} = \mathbb{C}$) bzw. selbstadjungiert (falls $\mathbb{K} = \mathbb{R}$). Dann gelten*

(i) Es gibt ein (evtl. endliches) Orthonormalsystem $(e_k)_{k\in\mathcal{N}}$ in H und eine (evtl. abbrechende) Folge $(\lambda_k)_{k\in\mathcal{N}}$ in $\mathbb{K}\setminus\{0\}$ mit

$$Te_k = \lambda_k e_k \qquad \forall\, k \in \mathcal{N}$$

für $\mathcal{N} \subset \mathbb{N}$, wobei für $|\mathcal{N}| = \infty$ gilt $\lim_{k\to\infty} \lambda_k = 0$. Außerdem ist

$$\sigma(T) \setminus \{0\} = \{\lambda_k \mid k \in \mathcal{N}\},$$

wobei die Eigenwerte λ_k nicht notwendigerweise verschieden sein müssen. Zudem folgt $\|T\| = \max_{k\in\mathcal{N}} |\lambda_k|$.

(ii) $H = \ker(T) \oplus \overline{\operatorname{span}\{e_k \mid k \in \mathcal{N}\}}$, wobei das Orthonormalsystem $(e_k)_{k\in\mathcal{N}}$ auch orthogonal zu $\ker(T)$ ist.

(iii) Für alle $x \in H$ gilt

$$Tx = \sum_{k\in\mathcal{N}} \lambda_k (e_k, x)\, e_k.$$

Beweis. (i) Nach Satz 5.15 ist für ein $\mathcal{M} \subset \mathbb{N}$

$$\sigma(T) \setminus \{0\} = \{\mu_k \mid \mu_k \text{ ist Eigenwert}, k \in \mathcal{M}\},$$

für paarweise verschiedene μ_k, wobei $\lim_{k\to\infty} \mu_k = 0$, falls $\mathcal{M}$ unendlich ist. Des weiteren ist $E_k := \ker(\mu_k - T)$ endlich dimensional für $k \in \mathcal{M}$. Setzt man $\mu_0 = 0$, so ist $E_0 = \ker(T)$ nicht notwendigerweise endlich dimensional, jedoch gilt nach den Übungen für die normalen Operator $\mu_k - T$

$$E_k = \ker(\mu_k - T) = \ker(\overline{\mu_k} - T^\dagger) \qquad \forall\, k \in \mathcal{M} \cup \{0\}.$$

Damit sind die Eigenräume paarweise senkrecht zueinander, d.h. $E_k \perp E_\ell$ für alle $k, \ell \in \mathcal{M} \cup \{0\}$ mit $k \neq \ell$, denn für $x_k \in E_k, x_\ell \in E_\ell$ ist

$$\mu_k(x_\ell, x_k) = (x_\ell, Tx_k) = (T^\dagger x_\ell, x_k) = (\overline{\mu_\ell} x_\ell, x_k) = \mu_\ell(x_\ell, x_k).$$

Für $k \neq \ell$ ist daher $(x_\ell, x_k) = 0$ bzw. $E_k \perp E_\ell$. Wählt man jeweils entsprechend der Vielfachheit d_k von μ_k eine Folge $(\lambda_k)_{k\in\mathcal{N}}$ durch wiederholte Nennung der μ_k,

$$(\lambda_k)_{k\in\mathcal{N}} = (\mu_1, \ldots, \mu_1, \mu_2, \ldots, \mu_2, \mu_3, \ldots, \mu_3, \ldots),$$

und eine Orthonormalbasis $\{e_1^k, \ldots, e_{d_k}^k\}$ zum Eigenraum E_k, so erhält man ein Orthonormalsystem $(e_k)_{k\in\mathcal{N}}$ durch Aufzählung

$$(e_k)_{k\in\mathcal{N}} = (e_1^1, \ldots, e_{d_1}^1, e_1^2, \ldots, e_{d_2}^2, \ldots).$$

Nach Lemma 5.16 gilt wegen $\|T\| \geq \max_{k\in\mathcal{N}} |\mu_k| = \max_{k\in\mathcal{N}} |\lambda_k|$ auch Gleichheit.

(ii) Die Orthogonalität der Eigenräume zeigt, dass die direkte Summe

$$H_1 := \ker(T) \oplus \overline{\operatorname{span}\{e_k \mid k \in \mathcal{N}\}},$$

der abgeschlossenen Räume wieder ein abgeschlossener Unterraum von H ist. Wir zeigen nun $H_1 = H$ und betrachten hierfür $H_2 := H_1^\perp$. Es ist H_2 invariant unter T, denn $T(H_2) \subset H_2$ gemäß

$$((x + e_k), Ty) = (T^\dagger(x + e_k), y) = (T^\dagger x, y) + \lambda_k(e_k, y) = 0,$$

für $y \in H_1^\perp$, $(x + e_k) \in H_1$ und $\ker(T) = \ker(T^\dagger)$ aufgrund der Normalität von T. Damit lässt sich $S := T|_{H_2}$ als kompakter Operator $S \in \mathscr{K}(H_2)$ auffassen. Wäre $S \neq 0$, so liefert Lemma 5.16 sowie der Spektralsatz 5.15 für kompakte Operatoren ein Eigenwert $\lambda \neq 0$ sowie ein $x \in H_2 \setminus \{0\}$ mit $Sx = \lambda x$. Dann wäre aber auch $\lambda \in \sigma(T) \setminus \{0\}$ und per Konstruktion $x \in \operatorname{span}\{e_k \mid k \in \mathcal{N}\} \subset H_2^\perp$. Folglich ist $0 \neq x \in H_2 \cap H_2^\perp = \{0\}$, ein Widerspruch. Also ist $S = 0$ und $H_2 \subset \ker(T)$. Per Konstruktion ist $\ker(T) \subset H_1 = H_1^{\perp\perp} = H_2^\perp$, d.h. $H_2 \subset H_2^\perp$ und schließlich $H_2 = 0$, denn $H_2 \cap H_2^\perp = \{0\}$. Dies zeigt die Darstellung $H = H_1$.

(iii) Nach Satz 1.53 lässt sich dann jedes $x \in H$ darstellen als

$$x = y + \sum_{k \in \mathcal{N}} (e_k, x) e_k$$

für ein $y \in \ker(T)$. Die Stetigkeit von T liefert dann

$$Tx = Ty + \sum_{k \in \mathcal{N}} (e_k, x) T e_k = \sum_{k \in \mathcal{N}} \lambda_k (e_k, x) e_k.$$

□