

# **Funktionalanalysis I**

**Wintersemester 2001/02**

**Prof. Dr. Michael Růžička**



# Inhaltsverzeichnis

<b>0</b>	<b>Einführung</b>	<b>1</b>
<b>1</b>	<b>Steilkurs: Lebesgue-Maß, Lebesgue-Integral</b>	<b>5</b>
1.1	Das Lebesgue-Maß . . . . .	5
1.2	Messbare Funktionen . . . . .	13
1.3	Das Lebesgue Integral . . . . .	18
1.4	Der Glättungsoperator . . . . .	28
<b>2</b>	<b>Hilberträume</b>	<b>33</b>
2.1	Der Hilbertraum . . . . .	33
2.2	Der Projektionssatz . . . . .	40
2.3	Rieszscher Darstellungssatz . . . . .	42
2.4	Beschränkte lineare Abbildungen . . . . .	45
2.5	Adjungierte Abbildungen . . . . .	52
2.6	Separable Hilberträume und Orthogonalsysteme . . . . .	63
2.7	Schwache Konvergenz und schwache Kompaktheit . . . . .	69
2.8	Kompakte lineare Abbildungen . . . . .	75
2.9	Das Eigenwertproblem . . . . .	86
<b>3</b>	<b>Allgemeine lineare Funktionalanalysis</b>	<b>97</b>
3.1	Grundbegriffe der Funktionalanalysis . . . . .	97
	Metrische Räume . . . . .	97
	Normierte Räume und Banachräume . . . . .	99
3.2	Beschränkte lineare Funktionale und Operatoren . . . . .	114
	Der Dualraum zu $L^p(\Omega)$ . . . . .	115
	Der duale Raum zu $C[a, b]$ . . . . .	116
	Der Dualraum zu $L^\infty(\Omega)$ . . . . .	117
3.3	Das Prinzip der gleichmäßigen Beschränktheit . . . . .	118
3.4	Der Satz von Hahn-Banach . . . . .	125
	Analytische Formulierung des Satzes von Hahn-Banach . . . . .	125
	Geometrische Formulierung des Satzes von Hahn-Banach . . . . .	129
3.5	Hahn-Banach in der konvexen Analysis . . . . .	134
	Konjugiert-konvexe Funktionen . . . . .	134
	Orlicz-Räume . . . . .	137
	Der Kuhn-Tucker-Satz der konvexen Optimierung . . . . .	139
	Alternativsätze für lineare Ungleichungssysteme . . . . .	142

3.6	Adjungierte lineare Operatoren in Banach-Räumen . . . . .	144
3.7	Schwache Topologie und reflexive Räume . . . . .	149
3.8	Konvexität und schwache Topologie . . . . .	155

# Kapitel 0

## Einführung

Die Funktionalanalysis beschäftigt sich mit *unendlich-dimensionalen* Vektorräumen (über  $\mathbb{R}$  oder  $\mathbb{C}$ ), in denen ein *Konvergenzbegriff* (Topologie) gegeben ist, sowie den Abbildungen zwischen ihnen. Besonders angestrebt sind Ergebnisse, die sich auf konkrete Funktionenräume (z.B.  $C[a, b]$ ,  $L^2(\Omega)$  etc.) und konkrete Gleichungen (z.B. Integralgleichungen oder partielle Differentialgleichungen) anwenden lassen.

Die Betrachtungsweise der Funktionalanalysis, Funktionen als Punkte in einem unendlich-dimensionalen Raum aufzufassen und geometrische Überlegungen u.ä. anzuwenden, hat sich als sehr fruchtbar erwiesen und zu einer enormen Bereicherung der Analysis geführt. Die Sätze der Funktionalanalysis sind sehr allgemein und geben ein tieferes Verständnis zahlreicher Zusammenhänge.

Die Funktionalanalysis ist ein Produkt dieses Jahrhunderts (Beginn ca. Anfang dieses Jahrhunderts). Der Name kommt von dem Wort „Funktional“. Ein *lineares Funktional* ist eine lineare Abbildung eines Vektorraumes in seinen Grundkörper. Ist dieser Vektorraum selbst ein Funktionenraum, so sind die Argumente des Funktionals Funktionen. Um nicht von einer Funktion von Funktionen zu sprechen, verwendeten unsere Vorfahren das Wort „Funktional“. (Heute würde man sich daran weniger stören.) Da in den Anwendungen der Funktionalanalysis viel mit solchen Abbildungen - also Funktionen von Funktionen, letztere aufgefaßt als Punkte eines Vektorraumes - gearbeitet wird, erklärt sich auf diese Weise der Name „Funktionalanalysis“.

Ein größeres Teilgebiet der Funktionalanalysis ist die *lineare Funktionalanalysis*. Sie befaßt sich z.B. mit der Lösbarkeit von linearen Gleichungen  $Au = f$  oder Eigenwertproblemen (Spektraltheorie). Ein anderes wichtiges Teilgebiet ist die *nichtlineare Funktionalanalysis*, die sich z.B. mit Fixpunktsätzen, mehrdeutiger Lösbarkeit nichtlinearer Gleichungen  $A(u) = f$  etc. befaßt.

In der Funktionalanalysis, welche mit *unendlich-dimensionalen* Vektorräumen arbeitet, gibt es viele, im Vergleich zur Analysis im  $\mathbb{R}^n$  ungewohnte Effekte. Wir zählen einige von diesen auf:

**a) Surjektive lineare Abbildungen können einen nichttrivialen Kern haben.**

Beispiel: Sei  $V = \{(c_1, c_2, c_3, \dots) \mid c_j \in \mathbb{R}\}$  der unendlich-dimensionale Vektorraum mit der komponentenweisen Addition als additiver Verknüpfung und der komponentenweise Multiplikation mit Skalaren  $\alpha \in \mathbb{R}$  als Skalarmultiplikation. (Man könnte  $V = \mathbb{R}^\infty$  oder besser  $\mathbb{R}^\omega$  schreiben.) Die lineare Abbildung  $A : V \rightarrow V$  sei definiert

durch

$$Ac = (c_2, c_3, c_4, \dots), \quad c = (c_1, c_2, c_3, \dots).$$

Offensichtlich ist  $A(V) = V$ , aber  $A(c_1, 0, 0, \dots) = (0, 0, \dots)$ .

**b) Ist der Kern  $N(A)$  einer linearen Abbildung trivial, so ist  $A$  nicht notwendig surjektiv.**

Beispiel:  $V$  sei wieder wie oben definiert. Es sei  $Ac = (0, c_1, c_2, c_3, \dots)$ . Aus  $Ac = 0$  folgt  $c = 0$ , also  $N(A) = 0$ . Offensichtlich ist aber  $A(V)$  echter Teilraum von  $V$ .

Die in a) und b) beobachteten Effekte können im endlich-dimensionalen Fall nicht eintreten.

Man bemüht sich in der Funktionalanalysis, durch Zusatzbedingungen an  $A$  die im  $\mathbb{R}^n$  bekannten Sätze zu retten.

**c) Lineare Abbildungen - selbst einfachster Art - müssen keinen Eigenwert haben.**

Beispiel:  $C[a, b] =$  Raum der  $\mathbb{R}$ - oder  $\mathbb{C}$ -wertigen stetigen Funktionen auf  $[a, b]$ . Sei  $A$  diejenige lineare Abbildung, welche definiert ist durch

$$(Af)(x) = (\sin x) \cdot f(x), \quad f \in C[a, b], \quad x \in [a, b].$$

$A$  ist also ein sehr simpler Operator, nämlich ein *Multiplikationsoperator* - die Multiplikation mit  $\sin x$ .

$A$  hat *keinen* Eigenwert. Andernfalls existiert  $f \in C[a, b], f \neq 0$  (d. h.  $f$  ist nicht die Null-Funktion) und  $\lambda \in \mathbb{R}$  (oder  $\mathbb{C}$ ) mit  $(\sin x)f(x) = \lambda f(x)$ . Da  $f \neq 0$ , existiert ein  $x_0 \in [a, b]$  mit  $f(x) \neq 0$  für  $x \in U(x_0)$  und es gilt

$$\sin x = \lambda = \text{const}, \quad x \in U(x_0).$$

Dies ist nicht möglich,  $\sin \neq \text{const}$  auf einer offenen Menge.

In der Funktionalanalysis gibt man Klassen von linearen Operatoren (= Abbildungen) an, für die das Eigenwertproblem ähnlich wie bei  $n \times n$ -Matrizen behandelt werden kann. Außerdem werden verschiedene Abschwächungen des Begriffs „Eigenwert“ eingeführt, um eine Analogie zu  $n \times n$ -Matrizen zu bekommen. Zudem sind diese Abschwächungen durch die Quantenphysik motiviert. Hiermit beschäftigt sich die „Spektraltheorie“.

**d) In unendlich-dimensionalen Räumen gibt es häufig mehrere natürliche Konvergenzbegriffe.**

Im  $\mathbb{R}^n$  gibt es außer der euklidischen Topologie natürlich noch beliebig viele andere Topologien - diese sind aber alle äquivalent.

Beispiel:  $V = l^2 = \{ (c_1, c_2, c_3, \dots) \mid \sum_{j=1}^{\infty} |c_j|^2 < \infty \}$ .

Man hat z.B. die Topologie der punktweisen Konvergenz

$$c^j \rightarrow c : \Leftrightarrow c_i^j \rightarrow c_i \quad (j \rightarrow \infty) \quad \text{für alle } i \in \mathbb{N}.$$

Ferner gibt es die starke  $l^2$ -Konvergenz

$$c^j \rightarrow c \Leftrightarrow \|c^j - c\|_{l^2} \rightarrow 0 \quad (j \rightarrow \infty).$$

Hierbei ist  $\|\omega\| = \left(\sum_{k=1}^{\infty} |\omega_k|^2\right)^{1/2}$ .

Schließlich kennt man noch die schwache  $l^2$ -Konvergenz, die wir später kennenlernen. Die drei Konvergenzen sind *nicht* äquivalent.

**e) Lineare Abbildungen sind nicht notwendig stetig.**

Sei  $V$  der Vektorraum der Polynome auf  $[-2, 2]$ , versehen mit der *gleichmäßigen Konvergenz* als Konvergenzbegriff, d.h.  $p^j \rightarrow p \Leftrightarrow \|p^j - p\|_{\infty} \rightarrow 0$  ( $j \rightarrow \infty$ ). Hierbei ist  $\|q\|_{\infty} = \max\{|q(x)| \mid x \in [-2, 2]\}$ .

Wir definieren die lineare Abbildung  $A : V \rightarrow V$ , indem wir sie auf den Basiselementen  $x^n$  definieren:

$$Ax^n = 3^n x^n.$$

Mit  $p_n(x) = \frac{1}{(2,5)^n} x^n$  gilt  $p_n \rightarrow 0$  ( $n \rightarrow \infty$ ), aber

$$\|Ap_n(x)\| = \frac{3^n \cdot 2^n}{(2,5)^n} \rightarrow \infty.$$

Die stetigen linearen Abbildungen und Klassen von unstetigen linearen Abbildungen werden in der Funktionalanalysis besonders studiert.

Schließlich betonen wir noch, daß man auch das Studium konkreter Funktionenräume zur Funktionalanalysis zählen kann, wenn auch die Grenzen fließend sind. Es gibt zahllose Funktionenräume in der Analysis. Der Anfänger kennt vermutlich die Räume  $C[a, b]$ ,  $C^1$ ,  $C^m$ ,  $C^\omega$ ,  $C^\alpha$ ,  $L^2(\Omega)$ ,  $L^p(\Omega)$ . Man benötigt die vielen Funktionenräume u.a., um Sätze der Funktionalanalysis einer konkreten Situation anzupassen.





# Kapitel 1

## Steilkurs: Lebesgue-Maß, Lebesgue-Integral

### 1.1 Das Lebesgue-Maß

Das Maß  $\mu(A)$  von Mengen  $A \subseteq \mathbb{R}^n$  ist eine natürliche Erweiterung der intuitiven Begriffe Länge (1D), Flächeninhalt (2D) und Volumen (3D) von Strecken, Flächen und Körpern auf allgemeinere Mengen. Dabei sollen intuitiv natürliche Eigenschaften, wie

- Positivität des Maßes,
- Übereinstimmung des Maßes mit bekannten Formeln für Würfel und Kugeln,
- Monotonie, d.h.  $A \subseteq B \Rightarrow \mu(A) \leq \mu(B)$ ,
- Additivität, d.h.  $\mu(\bigcup A_i) = \sum \mu(A_i)$ , mit  $A_i$  paarweise disjunkt,

erhalten bleiben. Darüber hinaus soll eine möglichst große Klasse von Teilmengen erfasst werden.

**1.1 Definition.** Für jede beliebige Teilmenge  $A \subseteq \mathbb{R}^n$  definieren wir

$$\mu^*(A) := \inf \left\{ \sum_{k=1}^{\infty} \text{vol}(I_k); A \subseteq \bigcup_{k=1}^{\infty} I_k, I_k \text{ ist ein offenes Intervall} \right\}.$$

Ein **offenes Intervall**  $I \subseteq \mathbb{R}^n$  hat die Form  $I = (a_1, b_1) \times \dots \times (a_n, b_n)$ ,  $a_i < b_i \in \mathbb{R}$ ,  $i = 1, \dots, n$  und das Volumen  $\text{vol}(I) = (b_1 - a_1) \cdot \dots \cdot (b_n - a_n)$ . Die so definierte Funktion  $\mu^* : 2^{\mathbb{R}^n} \rightarrow \mathbb{R} \cup \{\infty\}$  heißt **äußeres Lebesgue Maß**, wobei  $2^{\mathbb{R}^n}$  die Menge aller Teilmengen des  $\mathbb{R}^n$  ist.

Offensichtlich ist  $\mu^*(A) \geq 0$  für alle  $A \subseteq \mathbb{R}^n$  und es gilt:  $A \subseteq B \Rightarrow \mu^*(A) \leq \mu^*(B)$ . Aus der Definition folgt auch, dass  $\mu^*$  **translationsinvariant** ist, d.h. für alle  $\mathbf{x} \in \mathbb{R}^n$  und alle  $A \subseteq \mathbb{R}^n$  gilt:

$$\mu^*(A) = \mu^*(\mathbf{x} + A).$$

Allerdings kann man mit Hilfe des Auswahlaxioms zeigen, dass es disjunkte Mengen  $A, B \subseteq \mathbb{R}^n$  gibt mit (siehe Barner, Flohr, Analysis 2, S. 220 ff)

$$\mu^*(A \cup B) < \mu^*(A) + \mu^*(B).$$

Also müssen wir das Mengensystem, auf dem wir ein Maß erklären können, einschränken.

**1.2 Definition.** Eine Menge  $A \subseteq \mathbb{R}^n$  heißt **Lebesgue messbar**, wenn für alle Teilmengen  $T \subseteq \mathbb{R}^n$  gilt:

$$\mu^*(T) = \mu^*(T \cap A) + \mu^*(T \setminus A).$$

Das System aller Lebesgue - meßbaren Mengen bezeichnen wir mit  $\mathcal{M}$  und definieren das **Lebesgue Maß** von Mengen  $M \in \mathcal{M}$  durch

$$\mu(M) := \mu^*(M).$$

- Die Identität in 1.2 bedeutet, dass die Menge  $A$  messbar ist, wenn sie beliebige Testmengen  $T$  des  $\mathbb{R}^n$  in zwei disjunkte Teilmengen zerlegt, auf denen das äußere Lebesgue Maß  $\mu^*$  additiv ist.

**1.3 Satz.** Für beliebige Teilmengen  $A_j$  des  $\mathbb{R}^n$  gilt:

$$\mu^* \left( \bigcup_{j=1}^{\infty} A_j \right) \leq \sum_{j=1}^{\infty} \mu^*(A_j). \quad (1.4)$$

BEWEIS : Sei  $A = \bigcup_{j=1}^{\infty} A_j$ ,  $\sum_{j=1}^{\infty} \mu^*(A_j) < \infty$  und sei  $\varepsilon > 0$  fest aber beliebig. Nach Definition finden wir offene Intervalle  $I_n^j$  mit

$$A_j \subseteq \bigcup_{n=1}^{\infty} I_n^j, \quad \sum_{n=1}^{\infty} \text{vol}(I_n^j) < \mu^*(A_j) + \frac{\varepsilon}{2^j}.$$

Dann gilt:

$$A \subseteq \bigcup_{j,n=1}^{\infty} I_n^j \quad \text{und} \quad \mu^*(A) \leq \sum_{j=1}^{\infty} \mu^*(A_j) + \varepsilon.$$

■

- Da  $T \subseteq (T \cap A) \cup (T \setminus A)$  gilt, reicht es aufgrund von Satz 1.3 zu überprüfen ob,

$$\mu^*(T) \geq \mu^*(T \cap A) + \mu^*(T \setminus A),$$

um zu zeigen das  $A$  meßbar ist. Man kann sich sogar auf  $T$  mit  $\mu^*(T) < \infty$  beschränken.

- Jede Funktion  $\mu^*$ , die monoton und nichtnegativ ist, sowie (1.4) und  $\mu^*(\emptyset) = 0$  erfüllt, heißt **äußeres Maß**.

- Sei  $X$  eine Menge und  $\mu^*$  ein äußeres Maß auf  $X$ . Für beliebige Teilmengen  $A \subseteq X$  definieren wir die **Restriktion**  $\mu^*|_A$  von  $\mu^*$  durch:

$$\mu^*|_A(B) := \mu^*(A \cap B), \quad \forall B \subseteq X. \quad (1.5)$$

Man beachte, dass für beliebige Teilmengen  $A \subseteq X$  gilt:

$$B \text{ ist bzgl. } \mu^* \text{ messbar} \quad \Rightarrow \quad B \text{ ist bzgl. } \mu^*|_A \text{ messbar.} \quad (1.6)$$

**1.7 Definition.** Ein System  $\mathcal{S}$  von Teilmengen einer Menge  $X$  heißt  **$\sigma$ -Algebra**, falls:

- (i)  $X \in \mathcal{S}$ ,
- (ii)  $A \in \mathcal{S} \Rightarrow X \setminus A \in \mathcal{S}$ ,
- (iii)  $A_j \in \mathcal{S} \Rightarrow \bigcup_{j=1}^{\infty} A_j \in \mathcal{S}$ .

**1.8 Satz.** Das System aller Lebesgue-messbaren Mengen  $\mathcal{M}$  ist eine  $\sigma$ -Algebra.

BEWEIS :

- Offensichtlich ist  $\mathbb{R}^n \in \mathcal{M}$ , da für alle  $T \subseteq \mathbb{R}^n$  gilt:  $T \cap \mathbb{R}^n = T$  und  $T \setminus \mathbb{R}^n = \emptyset$ . Da

$$T \cap (\mathbb{R}^n \setminus A) = T \setminus A, \quad T \setminus (\mathbb{R}^n \setminus A) = T \cap A$$

gilt, sind die Komplemente  $\mathbb{R}^n \setminus A$  von Lebesgue - messbaren Mengen  $A$  wieder Lebesgue - messbar.

- Seien  $A, B \in \mathcal{M}$ . Dann gilt für alle  $T \subseteq \mathbb{R}^n$

$$\begin{aligned} \mu^*(T) &= \mu^*(T \cap A) + \mu^*(T \setminus A), \\ \mu^*(T \cap A) &= \mu^*(T \cap A \cap B) + \mu^*((T \cap A) \setminus B). \end{aligned}$$

Wenn man  $T \setminus (A \cap B)$  als Testmenge wählt und die Messbarkeit von  $A$  benutzt erhält man:

$$\begin{aligned} \mu^*(T \setminus (A \cap B)) &= \mu^*((T \setminus (A \cap B)) \cap A) + \mu^*((T \setminus (A \cap B)) \setminus A) \\ &= \mu^*((T \cap A) \setminus B) + \mu^*(T \setminus A). \end{aligned}$$

Aus diesen drei Identitäten folgt sofort

$$\mu^*(T) = \mu^*(T \cap (A \cap B)) + \mu^*(T \setminus (A \cap B)).$$

Also ist  $\mathcal{M}$  abgeschlossen bzgl. endlichen Durchschnitten (Induktion) und bzgl. Komplementsbildung, also auch bzgl. endlichen Vereinigungen. In der Tat gilt:

$$\bigcup_j A_j = \mathbb{R}^n \setminus (\mathbb{R}^n \setminus \bigcup_j A_j) = \mathbb{R}^n \setminus (\bigcap_j (\mathbb{R}^n \setminus A_j)).$$

• Seien  $A_j \in \mathcal{M}$  paarweise disjunkt. Wählt man  $T = A_1 \cup A_2$  und benutzt die Messbarkeit von  $A_1$  erhält man

$$\mu^*(A_1 \cup A_2) = \mu^*(A_1) + \mu^*(A_2),$$

und über Induktion die analoge Aussage für endliche disjunkte Vereinigungen. Also gilt

$$\sum_{j=1}^{\infty} \mu^*(A_j) = \lim_{k \rightarrow \infty} \sum_{j=1}^k \mu^*(A_j) = \lim_{k \rightarrow \infty} \mu^* \left( \bigcup_{j=1}^k A_j \right) \leq \mu^* \left( \bigcup_{j=1}^{\infty} A_j \right).$$

Die umgekehrte Ungleichung gilt nach Satz 1.3, und somit haben wir bewiesen

$$\mu^* \left( \bigcup_{j=1}^{\infty} A_j \right) = \sum_{j=1}^{\infty} \mu^*(A_j). \quad (1.9)$$

• Sei  $\{A_k\} \subseteq \mathcal{M}$  eine monoton aufsteigende Folge, d.h.  $A_1 \subseteq \dots \subseteq A_k \subseteq A_{k+1} \subseteq \dots$ . Wir haben

$$A_k = A_1 \cup \bigcup_{j=2}^k (A_j \setminus A_{j-1}),$$

wobei  $A_j \setminus A_{j-1}$  paarweise disjunkt sind. Also gilt nach (1.9):

$$\lim_{k \rightarrow \infty} \mu^*(A_k) = \mu^*(A_1) + \sum_{k=2}^{\infty} \mu^*(A_k \setminus A_{k-1}) = \mu^* \left( \bigcup_{k=1}^{\infty} A_k \right). \quad (1.10)$$

• Sei  $\{A_k\} \subseteq \mathcal{M}$  eine monoton fallende Folge mit  $\mu^*(A_1) < \infty$ , d.h.  $A_1 \supseteq \dots \supseteq A_k \supseteq A_{k+1} \supseteq \dots$ . Dann gilt nach (1.10) und der Messbarkeit von  $A_k$

$$\begin{aligned} \mu^*(A_1) - \lim_{k \rightarrow \infty} \mu^*(A_k) &= \lim_{k \rightarrow \infty} \mu^*(A_1 \setminus A_k) \\ &= \mu^* \left( \bigcup_{k=1}^{\infty} (A_1 \setminus A_k) \right) = \mu^* \left( A_1 \setminus \bigcap_{k=1}^{\infty} A_k \right) \\ &\geq \mu^*(A_1) - \mu^* \left( \bigcap_{k=1}^{\infty} A_k \right). \end{aligned}$$

Aufgrund der Monotonie von  $\mu^*$  gilt

$$\mu^* \left( \bigcap_{j=1}^{\infty} A_j \right) \leq \mu^*(A_k)$$

und somit

$$\lim_{k \rightarrow \infty} \mu^*(A_k) = \mu^* \left( \bigcap_{k=1}^{\infty} A_k \right). \quad (1.11)$$

• Seien nun  $A_k \subseteq \mathcal{M}$  beliebig. Wir setzen  $B_j = \bigcup_{k=1}^j A_k$ . Somit ist  $B_j$  eine monoton aufsteigende Folge messbarer Mengen und es gilt  $\bigcup_k A_k = \bigcup_k B_k$ . Für  $T \subseteq \mathbb{R}^n$  beliebig, mit  $\mu^*(T) < \infty$  gilt unter mit Hilfe von (1.6), (1.10), (1.11):

$$\begin{aligned} \mu^* \left( T \cap \bigcup_{k=1}^{\infty} A_k \right) + \mu^* \left( T \setminus \bigcup_{k=1}^{\infty} A_k \right) &= \mu^*|_T \left( \bigcup_{k=1}^{\infty} (B_k) \right) + \mu^*|_T \left( \bigcap_{k=1}^{\infty} (\mathbb{R}^n \setminus B_k) \right) \\ &= \lim_{k \rightarrow \infty} \mu^*|_T (B_k) + \lim_{k \rightarrow \infty} \mu^*|_T (\mathbb{R}^n \setminus B_k) \\ &= \mu(T). \end{aligned}$$

Somit ist  $\bigcup_{k=1}^{\infty} A_k$  messbar und der Satz ist bewiesen. ■

- Aufgrund der Komplementeigenschaft der  $\sigma$ -Algebra  $\mathcal{M}$  sind natürlich auch beliebige Durchschnitte messbarer Mengen messbar.

**1.12 Folgerung.** Seien  $A_k \in \mathcal{M}$ .

(i) Für paarweise disjunkte  $A_k$  gilt:

$$\mu \left( \bigcup_{k=1}^{\infty} A_k \right) = \sum_{k=1}^{\infty} \mu(A_k).$$

(ii) Für  $A_1 \subseteq \dots \subseteq A_k \subseteq A_{k+1} \subseteq \dots$  gilt:

$$\lim_{k \rightarrow \infty} \mu(A_k) = \mu \left( \bigcup_{k=1}^{\infty} A_k \right).$$

(iii) Für  $A_1 \supseteq \dots \supseteq A_k \supseteq A_{k+1} \supseteq \dots$  mit  $\mu(A_1) < \infty$  gilt:

$$\lim_{k \rightarrow \infty} \mu(A_k) = \mu \left( \bigcap_{k=1}^{\infty} A_k \right).$$

BEWEIS : Wurde bereits bewiesen. ■

**1.13 Satz.** Sei  $I \subseteq \mathbb{R}^n$  ein beschränktes Intervall. Dann gilt:

$$\mu^*(I) = \text{vol}(I).$$

BEWEIS : Sei  $Q_j$  eine Überdeckung von  $I$  aus offenen Intervallen. Sei  $J$  ein kompaktes Intervall das in  $I$  enthalten ist. Dann gibt es  $p \in \mathbb{N}$  und  $Q_1, \dots, Q_p$  welche  $J$  überdecken. Wir zerlegen  $J$  in endlich viele nicht überlappende Intervalle  $J_i$  (d.h.

$\text{Int}(J_i) \cap \text{Int}(J_k) = \emptyset$ , für  $i \neq k$  wobei  $\text{Int}(A)$  die Menge der inneren Punkte von  $A$  ist) derart, dass  $J_i$  in einem der  $Q_i$ ,  $i = 1, \dots, p$ , enthalten ist. Wir haben

$$\text{vol}(J) = \sum_i \text{vol}(J_i) \leq \sum_{i=1}^p \text{vol}(Q_i) \leq \sum_{i=1}^{\infty} \text{vol}(Q_i).$$

Da die Differenz  $\text{vol}(I) - \text{vol}(J)$  beliebig klein gemacht werden kann folgt

$$\text{vol}(I) \leq \sum_{i=1}^{\infty} \text{vol}(Q_i),$$

d.h. nach der Definition des äußeren Maßes  $\mu^*$  gilt

$$\text{vol}(I) = \mu^*(I).$$

■

**1.14 Satz.** (a) Jede offene Teilmenge des  $\mathbb{R}^n$  ist messbar.

(b) Falls  $\mu^*(A) = 0$ , dann ist  $A$  messbar.

BEWEIS : Sei  $H = (-\infty, c) \times \mathbb{R}^{n-1}$ , d.h. ein Halbraum und sei  $T \subseteq \mathbb{R}^n$ , mit  $\mu^*(T) < \infty$  beliebig. Für  $\varepsilon > 0$  existieren offene Intervalle  $Q_j$  mit:

$$T \subset \bigcup_{j=1}^{\infty} Q_j, \quad \sum_{j=1}^{\infty} \text{vol}(Q_j) < \mu^*(T) + \varepsilon.$$

Wir nehmen offene Intervalle  $I_j, K_j$  derart

$$\begin{aligned} I_j \cup K_j &\supseteq Q_j, & Q_j \cap H &\subseteq I_j, & Q_j \setminus H &\subseteq K_j, \\ \mu^*(I_j) + \mu^*(K_j) &\leq \mu^*(Q_j) + \varepsilon 2^{-j}. \end{aligned}$$

Nach Folgerung 1.12 gilt:

$$\begin{aligned} \mu^*(T \cap H) + \mu^*(T \setminus H) &\leq \sum_j \text{vol}(I_j) + \sum_j \text{vol}(K_j) \\ &< \mu^*(T) + 2\varepsilon. \end{aligned}$$

Da  $\varepsilon$  beliebig war ist der Halbraum  $H$  messbar. Jede offene Menge kann als Vereinigung von abzählbar vielen Intervallen dargestellt werden. Jedes dieser Intervalle ist der endliche Durchschnitt von Halbräumen. Da  $\mathcal{M}$  eine  $\sigma$ -Algebra ist ist also jede offene Menge messbar. Sei  $A$  derart, dass  $\mu^*(A) = 0$ . Dann gilt für beliebige  $T \subseteq \mathbb{R}^n$

$$\mu^*(T) \leq \mu^*(T \cap A) + \mu^*(T \setminus A) \leq \mu^*(T \setminus A) \leq \mu^*(T),$$

also ist  $A$  messbar.

■

- Zu jedem System  $\mathcal{T}$  von Teilmengen des  $\mathbb{R}^n$  existiert eine kleinste  $\sigma$ -Algebra  $\mathcal{S}$ , die das gegebene System  $\mathcal{T}$  enthält. (Man nehme den Durchschnitt aller  $\mathcal{T}$  enthaltenden  $\sigma$ -Algebren.) Man sagt, dass  $\mathcal{S}$  von  $\mathcal{T}$  generiert wird. Die von den offenen Mengen generierte  $\sigma$ -Algebra heißt **Borel  $\sigma$ -Algebra**. Satz 1.14 zeigt, dass die  $\sigma$ -Algebra der Lebesgue-messbaren Mengen die Borel  $\sigma$ -Algebra enthält.

**1.15 Satz.** Für  $A \subseteq \mathbb{R}^n$  gilt

$$\mu^*(A) = \inf\{\mu(G); G \text{ offen}, G \supseteq A\}.$$

BEWEIS : Aus der Monotonie von  $\mu^*$  folgt

$$\mu^*(A) \leq \inf\{\mu(G); G \text{ offen}, G \supseteq A\}.$$

Sei  $\mu^*(A) < \infty$ , dann gibt es für alle  $\varepsilon > 0$  offene Intervalle  $Q_j$  mit

$$A \subseteq \bigcup_{j=1}^{\infty} Q_j, \quad \mu\left(\bigcup_{j=1}^{\infty} Q_j\right) \leq \sum_{j=1}^{\infty} \text{vol}(Q_j) \leq \mu^*(A) + \varepsilon.$$

Somit gilt auch die andere Ungleichung, denn

$$\begin{aligned} \inf\{\mu(G), G \text{ offen}\} &\leq \inf\left\{\mu\left(\bigcup_j Q_j\right), Q_j \text{ offene Intervalle}\right\} \\ &< \mu^*(A) + \varepsilon. \end{aligned}$$

und  $\varepsilon > 0$  ist beliebig. ■

Durch den Übergang zum Komplement kann man zeigen (siehe Evans Measure Theory and fine properties of functions S. 8)

**1.16 Satz.** Sei  $A \subseteq \mathbb{R}^n$  messbar. Dann gilt

$$\mu(A) = \sup\{\mu(K); K \text{ kompakt}, K \subseteq A\}.$$

**1.17 Satz.** Sei  $A \subseteq \mathbb{R}^n$ . Es sind äquivalent:

- (i)  $A$  ist messbar,
- (ii) für jedes beschränkte Intervall  $I$  gilt:

$$\mu^*(I) = \mu^*(I \cap A) + \mu^*(I \setminus A),$$

- (iii) für alle  $\varepsilon > 0$  existiert eine offene Menge  $G \supseteq A$  mit  $\mu^*(G \setminus A) < \varepsilon$ ,
- (iv) es gibt eine  $G_\delta$ -Menge  $D \supset A$ , d.h.  $D$  ist der abzählbare Durchschnitt offener Mengen, mit  $\mu^*(D \setminus A) = 0$ ,
- (v) es gibt eine  $F_\delta$ -Menge  $B_i$ , d.h.  $B_i$  ist die abzählbare Vereinigung abgeschlossener Mengen, und eine  $G_\delta$ -Menge  $B_e$  mit  $B_i \subseteq A \subseteq B_e$  und  $\mu^*(B_e \setminus B_i) = 0$ .

BEWEIS : (i)  $\Rightarrow$  (ii) ist klar. Wir nehmen nun an, dass (ii) gilt, wählen ein beliebiges  $\varepsilon > 0$  und bezeichnen  $I_k = (-k, k)^n$ . Aus Satz 1.15 folgt die Existenz offener Mengen  $G_k$  und  $H_k$  mit  $I_k \cap A \subseteq G_k$ ,  $I_k \setminus A \subseteq H_k$

$$\begin{aligned}\mu(G_k) &\leq \mu^*(I_k \cap A) + \varepsilon 2^{-k}, \\ \mu(H_k) &\leq \mu^*(I_k \setminus A) + \varepsilon 2^{-k}.\end{aligned}$$

Wir können annehmen, dass  $G_k$  und  $H_k$  Teilmengen von  $I_k$  sind und also  $G_k \setminus A \subseteq G_k \cap H_k$  gilt. Aus (ii) und der Messbarkeit offener Mengen und  $I_k = H_k \cup (G_k \setminus H_k)$  folgt:

$$\begin{aligned}\mu(I_k) + \mu(G_k \cap H_k) &\leq \mu(H_k) + \mu(G_k \setminus H_k) + \mu(G_k \cap H_k) \\ &= \mu(H_k) + \mu(G_k) \\ &\leq \mu^*(I_k \setminus A) + \mu^*(I_k \cap A) + \varepsilon 2^{-k+1} \\ &= \mu(I_k) + \varepsilon 2^{-k+1}.\end{aligned}$$

Wir setzen  $G \equiv \bigcup_{k=1}^{\infty} G_k$  und da  $\mu(I_k) < \infty$  gilt:

$$\mu^*(G \setminus A) \leq \sum_k \mu(G_k \cap H_k) \leq 2\varepsilon,$$

d.h. (iii) gilt. (iii)  $\Rightarrow$  (iv) ist klar (man setzt  $D = \bigcap_n G_{\frac{1}{n}}$ , mit  $G_{\frac{1}{n}}$  aus (iii)). Um (iv)  $\Rightarrow$  (v) zu zeigen, wendet man (iv) auf  $A$  und  $\mathbb{R}^n \setminus A$  an und erhält Mengen  $B_e \supseteq A$ ,  $D \supseteq \mathbb{R}^n \setminus A$ , mit  $\mu^*(B_e \setminus A) = 0$  und  $\mu^*(D \setminus (\mathbb{R}^n \setminus A)) = 0$ . Wir setzen  $B_i = \mathbb{R}^n \setminus D$  und erhalten

$$0 \leq \mu^*(B_e \setminus B_i) \leq \mu^*((B_e \setminus B_i) \setminus A) + \mu^*((B_e \setminus B_i) \cap A) \leq 0,$$

denn  $(B_e \setminus B_i) \setminus A \subseteq B_e \setminus A$  und  $(B_e \setminus B_i) \cap A = A \setminus B_i = A \setminus (\mathbb{R}^n \setminus D) = A \cap D = D \setminus (\mathbb{R}^n \setminus A)$ . Falls  $A$  (v) erfüllt, haben wir  $A = B_i \cup (A \setminus B_i)$  mit messbaren Mengen  $B_i$  (abgeschlossene Mengen sind messbar und  $\mathcal{M}$  ist abgeschlossen bzgl. abzählbaren Vereinigungen) und  $A \setminus B_i$  ( $A \setminus B_i \subseteq B_e \setminus B_i$  und  $\mu^*(B_e \setminus B_i) = 0$  plus Satz 1.14 b). Also ist  $A$  messbar. ■

Der hier vorgestellte Zugang kann völlig analog auf beliebige Mengen  $X$  verallgemeinert werden. In diesem Falle werden folgende Begriffe und Bezeichnungen benutzt:

- Sei  $\mathcal{S}$  ein System von Teilmengen von  $X$ . Eine nichtnegative Funktion  $\mu : \mathcal{S} \rightarrow [0, \infty]$  heißt Maß, falls
  - (i)  $\mathcal{S}$  eine  $\sigma$ -Algebra ist,
  - (ii)  $\mu(\emptyset) = 0$ ,
  - (iii) für jede Folge paarweiser disjunkter Mengen  $A_j \in \mathcal{S}$  gilt:

$$\mu\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mu(A_j).$$

Das Tripel  $(X, \mathcal{S}, \mu)$  heißt **Maßraum**.



- Aus (iii) folgt, dass  $\mu$  **monoton** ist, d.h.  $A \subseteq B \Rightarrow \mu(A) \leq \mu(B)$ . Die Eigenschaft (iii) heißt  **$\sigma$ -Additivität**.
- Man sagt, dass ein Maß **endlich** ist, falls  $\mu(X) < \infty$ ; im Falle  $\mu(X) = 1$  spricht man von einem **Wahrscheinlichkeitsmaß**. Eine Teilmenge  $A \subseteq X$  heißt  **$\sigma$ -endlich**, falls man Mengen  $B_k$ ;  $\mu(B_k) < \infty$  findet mit  $A = \bigcup_k B_k$ .
- Ein Maß heißt **vollständig**, wenn folgendes gilt:<sup>1</sup>

$$B \in \mathcal{S}, \quad \mu(B) = 0, \quad A \subseteq B \Rightarrow A \in \mathcal{S}.$$

Eine Menge mit Maß Null heißt **Nullmenge**.

- Sei  $\mu$  ein Maß auf  $X$  und  $E \in \mathcal{S}$ . Wir definieren die **Restriktion des Maßes  $\mu$  auf  $E$**  durch  $\mu|_E(A) \equiv \mu(A \cap E)$  für  $A \in \mathcal{S}$  und bezeichnen mit  $\mathcal{S}_E$  die  $\sigma$ -Algebra  $\{M \in \mathcal{S}, M \subseteq E\}$  von Teilmengen von  $E$ . Auf  $\mathcal{S}_E$  definieren wir  $\mu|_E(M) = \mu(M)$  für alle  $M \in \mathcal{S}_E$ . Somit übertragen sich alle Eigenschaften des Maßraumes  $(X, \mathcal{S}, \mu)$  auf  $(E, \mathcal{S}_E, \mu|_E)$ .
- Wir haben also gezeigt, dass das Lebesgue Maß  $\mu$  auf  $\mathcal{M}$  ein vollständiges,  $\sigma$ -endliches Maß ist, wobei alle Borel Mengen Lebesgue - messbar sind. Die Restriktion des Lebesgue Maßes auf beliebige messbare Mengen  $A$  hat völlig analoge Eigenschaften wie das Lebesgue Maß auf  $\mathbb{R}^n$ .

## 1.2 Messbare Funktionen

In Sektion 1.1 haben wir gesehen, dass es gute Gründe gibt das Lebesgue Maß nicht auf allen Teilmengen des  $\mathbb{R}^n$  zu definieren. Analog kann man auch nicht erwarten eine vernünftige Integrationstheorie für beliebige Funktionen zu erhalten. Also muß man sich auf ein angemessenes System von Funktionen beschränken, das natürlich beschränkte glatte Funktionen und Indikatorfunktionen von messbaren Mengen erhalten soll. Für  $A \in \mathcal{M}$  ist die **Indikatorfunktion**  $\chi_A : \mathbb{R}^n \rightarrow \mathbb{R}$  definiert durch

$$\chi_A(x) = \begin{cases} 0 & x \notin A, \\ 1 & x \in A. \end{cases}$$

Im folgenden sei  $(X, \mathcal{S}, \mu)$  ein vollständiger Massraum. Man stelle sich das Lebesgue Maß auf  $\mathbb{R}^n$  oder messbaren Teilmengen des  $\mathbb{R}^n$  vor.

**2.1 Definition.** Sei  $D \in \mathcal{S}$ . Eine Funktion  $f : D \rightarrow \overline{\mathbb{R}} \equiv \mathbb{R} \cup \{\pm\infty\}$  heißt  **$\mu$ -messbar** auf  $D$ , falls die Menge

$$\{x \in D; \quad f(x) > \alpha\} =: \{f > \alpha\}$$

$\mu$ -messbar ist für alle  $\alpha \in \mathbb{R}$ .

<sup>1</sup>Sei  $X$  eine Menge und  $\mathcal{S}$  eine  $\sigma$ -Algebra, die nicht alle Teilmengen von  $X$  enthält. Das Maß  $\mu$  definiert durch  $\mu(A) = 0$  für alle  $A \in \mathcal{S}$  ist nicht vollständig.

- Da  $\sigma$ -Algebren abgeschlossen bzgl. des Komplements sind, ist  $f$  messbar genau dann, wenn  $\{f \leq \alpha\} \in \mathcal{S}$  für alle  $\alpha \in \mathbb{R}$  ist. Da

$$\{f \geq \alpha\} = \bigcap_{n=1}^{\infty} \left\{f > \alpha - \frac{1}{n}\right\}$$

gilt, kann man in der obigen Definition die Bedingung  $\{f > \alpha\} \in \mathcal{S}$  durch  $\{f \geq \alpha\} \in \mathcal{S}$  ersetzen.

**2.2 Satz.** *Seien  $f, g$  auf  $D$  definierte messbare Funktionen. Dann gilt:*

- (i)  $\{f < g\} := \{x \in D; f(x) < g(x)\} \in \mathcal{S}$ ,
- (ii)  $f^{-1}(\infty), f^{-1}(-\infty) \in \mathcal{S}$ .

BEWEIS : Da

$$\begin{aligned} \{f < g\} &= \bigcup_{q \in \mathbb{Q}} (\{f < q\} \cap \{g > q\}), \\ f^{-1}(\infty) &= \bigcap_{n \in \mathbb{N}} \{f > n\}, \end{aligned}$$

gilt, folgen die Behauptungen sofort. ■

**2.3 Satz (Eigenschaften messbarer Funktionen).** *Seien  $f, g : X \rightarrow \mathbb{R}$ ,  $f_n : X \rightarrow \mathbb{R}$  messbare Funktionen,  $\lambda \in \mathbb{R}$  und  $\varphi$  eine stetige Funktion definiert auf einer offenen Menge  $G \subseteq \mathbb{R}$ . Dann sind auch die folgenden Funktionen messbar*

- (i)  $\lambda f, f + g, \max(f, g), \min(f, g), f^+, f^-, |f|, fg, f/g$  falls  $g \neq 0$ ,
- (ii)  $\sup_n f_n, \inf_n f_n, \limsup_n f_n, \liminf_n f_n, \lim f_n$  falls der Grenzwert existiert,
- (iii)  $\varphi \circ f$ , falls  $f(X) \subseteq G$ .

BEWEIS : (iii) Da  $\{\varphi \circ f < \alpha\} = \{x; f(x) \in \varphi^{-1}(-\infty, \alpha)\}$  und  $\varphi^{-1}(-\infty, \alpha)$  eine offene Menge ist, die also als abzählbare Vereinigung von offenen Intervallen dargestellt werden kann, welche wiederum Durchschnitte von Halbgeraden der Form  $(-\infty, a), (b, \infty)$  sind, ist die Behauptung klar.

(i) Sei  $\alpha \in \mathbb{R}$ . Die Messbarkeit von  $\lambda f$  ist klar. Es gilt  $\{f + g > \alpha\} = \{f > \alpha - g\}$ . Aber  $\alpha - g$  ist messbar und Satz 2.2 liefert die Behauptung. Die Funktionen  $|f|, f^2$  und  $\frac{1}{g}$  sind nach (iii) messbar. Wir haben

$$\begin{aligned} \max(f, g) &= \frac{1}{2}(f + g + |f - g|), & \min(f, g) &= \frac{1}{2}(f + g - |f - g|), \\ f^+ &= \max(f, 0), & f^- &= \max(-f, 0), \\ fg &= \frac{1}{2}((f + g)^2 - f^2 - g^2). \end{aligned}$$

(ii) Wir haben

$$\begin{aligned}\{\sup_n f_n > \alpha\} &= \bigcup_n \{f_n > \alpha\}, \\ \{\inf_n f_n > \alpha\} &= \bigcap_n \{f_n > \alpha\},\end{aligned}$$

welches die Messbarkeit von  $\sup_n f_n$  und  $\inf_n f_n$  liefert. Die Formeln

$$\begin{aligned}\liminf_{n \rightarrow \infty} f_n &= \sup_{m \geq 1} \inf_{n \geq m} f_n, \\ \limsup_{n \rightarrow \infty} f_n &= \inf_{m \geq 1} \sup_{n \geq m} f_n\end{aligned}$$

beenden den Beweis. ■

- Der **Träger** einer messbaren Funktion  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  ist die Menge

$$\text{supp}(f) \equiv \overline{\{x \in \mathbb{R}^n; f(x) \neq 0\}}.$$

- Sei  $\Omega \subseteq \mathbb{R}^n$  eine offene Menge. Wir bezeichnen mit  $C(\Omega)$ ,  $C^k(\Omega)$  bzw.  $C^\infty(\Omega)$  die Menge der stetigen,  $k$ -mal stetig differenzierbaren bzw. unendlich oft differenzierbaren Funktionen. Mit  $C_0(\Omega)$ ,  $C_0^k(\Omega)$  bzw.  $C_0^\infty(\Omega)$  bezeichnen wir die Teilmengen der obigen Räume bestehend aus Funktionen mit kompakten Trägern, d.h.  $\text{supp}(f) \subseteq \subseteq \Omega$ . Hierbei bedeutet  $A \subseteq \subseteq \Omega$ , dass  $\overline{A} \subseteq \Omega$  und  $\overline{A}$  eine kompakte Menge ist.

**2.4 Satz (Lusin).** Sei  $A \subseteq \mathbb{R}^n$  eine offene Menge mit  $\mu(A) < \infty$  und  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  messbar. Für alle  $\varepsilon > 0$  existiert eine kompakte Menge  $K = K(\varepsilon) \subseteq A$  so, dass

- (i)  $\mu(A \setminus K) < \varepsilon$ ;
- (ii)  $f|_K$  ist stetig.

BEWEIS : Für alle  $i \in \mathbb{N}$  seien  $\{B_{ij}\}_{j=1}^\infty \subseteq \mathbb{R}$  paarweise disjunkte Borelmengen mit  $\bigcup_{j=1}^\infty B_{ij} = \mathbb{R}$  so, dass für den Durchmesser gilt:  $\text{diam}(B_{ij}) < \frac{1}{i}$ . (Man nehme Intervalle der Form  $(\frac{j}{2^i}, \frac{(j+1)}{2^i}]$ .) Dann sind die Mengen  $A_{ij} \equiv A \cap f^{-1}(B_{ij})$  messbar und es gilt  $A = \bigcup_{j=1}^\infty A_{ij}$ . Satz 1.16 liefert die Existenz kompakter Mengen  $K_{ij} \subseteq A_{ij}$  mit  $\mu(A_{ij} \setminus K_{ij}) < \frac{\varepsilon}{2^{i+j}}$ . Dann haben wir

$$\begin{aligned}\mu\left(A \setminus \bigcup_{j=1}^\infty K_{ij}\right) &= \mu\left(\bigcup_{j=1}^\infty A_{ij} \setminus \bigcup_{j=1}^\infty K_{ij}\right) \\ &\leq \mu\left(\bigcup_{j=1}^\infty (A_{ij} \setminus K_{ij})\right) < \frac{\varepsilon}{2^i}.\end{aligned}$$

Da  $\lim_{N \rightarrow \infty} \mu \left( A \setminus \bigcup_{j=1}^N K_{ij} \right) = \mu \left( A \setminus \bigcup_{j=1}^{\infty} K_{ij} \right)$  gilt, existiert eine Zahl  $N(i)$  mit

$$\mu \left( A \setminus \bigcup_{j=1}^{N(i)} K_{ij} \right) < \frac{\varepsilon}{2^i}.$$

Die Menge  $D_i \equiv \bigcup_{j=1}^{N(i)} K_{ij}$  ist kompakt. Für alle  $i, j$  wählen wir ein  $b_{ij} \in B_{ij}$  und definieren  $g_i : D_i \rightarrow \mathbb{R}$  durch  $g_i(x) = b_{ij}$  für  $x \in K_{ij}$  ( $j \leq N(i)$ ). Da  $K_{i1}, \dots, K_{iN(i)}$  kompakte disjunkte Mengen sind haben sie einen positiven Abstand voneinander. Also ist  $g_i$  stetig. Wir haben auch

$$|f(x) - g_i(x)| \leq \frac{1}{i} \quad \forall x \in D_i. \quad (*)$$

Wir setzen  $K \equiv \bigcap_{i=1}^{\infty} D_i$ . Diese Menge ist kompakt und es gilt

$$\mu(A \setminus K) \leq \sum_{i=1}^{\infty} \mu(A \setminus D_i) < \varepsilon.$$

Aus (\*) und der Definition von  $K$  folgt, dass  $g_i$  gleichmäßig gegen  $f$  auf  $K$  konvergiert. Also ist  $f|_K$  stetig. ■

Es gilt folgende, stärkere Version vom Satz von Lusin.

**2.5 Folgerung.** *Unter der Voraussetzung von Satz 2.4 gibt es ein  $g \in C_0(A)$  mit*

$$\begin{aligned} \sup_{x \in \mathbb{R}^n} |g(x)| &\leq \sup_{x \in \mathbb{R}^n} |f(x)|, \\ \mu(\{x \in \mathbb{R}^n; f(x) \neq g(x)\}) &< \varepsilon. \end{aligned}$$

BEWEIS : Evans S. 16. ■

**2.6 Definition.** *Eine **Treppenfunktion**  $f$  ist eine endliche Linearkombination von Indikatorfunktionen messbarer Mengen, d.h. es gibt  $\alpha_i \in \mathbb{R}$  und  $A_i \in \mathcal{S}$  mit*

$$f = \sum_{i=1}^m \alpha_i \chi_{A_i}.$$

• *Treppenfunktionen nehmen also nur endlich viele Werte an.*

**2.7 Satz.** *Sei  $f \geq 0$  eine messbare Funktion. Dann existiert eine monoton steigende Folge  $\{f_n\}$  nichtnegativer Treppenfunktionen mit  $f_n \nearrow f$ .*

BEWEIS : Für  $m \in \mathbb{N}$  und  $k = 1, 2, \dots, m2^m$  setzen wir

$$F_{m,k} = \left\{ x \in X; \frac{k-1}{2^m} \leq f(x) < \frac{k}{2^m} \right\}$$

und definieren

$$f_m(x) = \begin{cases} \frac{k-1}{2^m}, & \text{falls } x \in F_{m,k}, \\ m, & \text{falls } x \in X \setminus \bigcup_k F_{m,k}. \end{cases}$$

Offensichtlich sind  $f_m$  Treppenfunktionen für die gilt:  $f_m(x) \nearrow f(x)$  für alle  $x \in X$ . ■  
Da im Allgemeinen Nullmengen für die Integration keine Rolle spielen ist folgender Begriff von fundamentaler Bedeutung.

**2.8 Definition.** Wir sagen eine Funktion  $h$  ist **fast überall** definiert, falls der Definitionsbereich  $D \in \mathcal{S}$  die Bedingung  $\mu(X \setminus D) = 0$  erfüllt. Seien  $f$  und  $g$  Funktionen, die fast überall auf  $X$  definiert sind. Wir sagen  $f(x) \leq g(x)$  **fast überall**, falls es eine Menge  $N \in \mathcal{S}$  gibt mit  $\mu(N) = 0$  und  $f(x) \leq g(x)$  für alle  $x \in X \setminus N$  gilt.

- Analog wird der Begriff „fast überall“ oder „fast alle“ in anderen Zusammenhängen verstanden, z.B. **fast überall Konvergenz**.
- Die Relation „ $f = g$  fast überall“ ist offensichtlich eine Äquivalenzrelation auf der Menge der messbaren Funktionen. Man kann also eine gegebene Funktion  $f$  auf einer beliebigen Nullmenge beliebig umdefinieren und bleibt in der gleichen Äquivalenzklasse.

In den weiteren Kapiteln werden wir oft nicht zwischen der Funktion  $f$  und ihrer Äquivalenzklasse  $[f]$  unterscheiden. (Etwas ungenau aber sehr praktisch.)

**2.9 Satz (Egoroff).** Sei  $A \subseteq \mathbb{R}^n$  messbar mit  $\mu(A) < \infty$  und seien  $f_n : A \rightarrow \mathbb{R}$  messbare Funktionen mit  $f_n \rightarrow g$  fast überall. Dann existiert für alle  $\varepsilon > 0$  eine messbare Menge  $B \subseteq A$  mit

$$(i) \quad \mu(A \setminus B) < \varepsilon,$$

$$(ii) \quad f_k \rightrightarrows g \text{ gleichmäßig auf } B.$$

BEWEIS : Sei  $C_{i,j} \equiv \bigcup_{k=j}^{\infty} \{x \in A \mid |f_k(x) - g(x)| > 2^{-i}\}$ ,  $i, j = 1, 2, \dots$ . Dann ist  $C_{i,j+1} \subseteq C_{i,j}$  für alle  $i, j \in \mathbb{N}$ . Da  $\mu(A) < \infty$  haben wir mit Folgerung 1.12 (iii)

$$\lim_{j \rightarrow \infty} \mu(C_{i,j}) = \mu\left(\bigcap_{j=1}^{\infty} C_{i,j}\right) = 0, \quad i = 1, 2, \dots$$

Also gibt es ein  $N(i)$  mit  $\mu(C_{i,N(i)}) < \varepsilon 2^{-i}$ . Wir setzen  $B = A \setminus \bigcup_{i=1}^{\infty} C_{i,N(i)}$  und erhalten

$$\mu(A \setminus B) \leq \sum_{i=1}^{\infty} \mu(C_{i,N(i)}) < \varepsilon.$$

Für alle  $i$ , alle  $x \in B$  und alle  $n > N(i)$  gilt

$$|f_n(x) - g(x)| \leq 2^{-i},$$

d.h.  $f_n \rightrightarrows g$  auf  $B$ . ■

### 1.3 Das Lebesgue Integral

Das Riemann Integral der Indikatorfunktion eines beliebigen Intervalls ist gleich seiner Länge. Demzufolge ist es natürlich für das Lebesgue Integral

$$\int \chi_A d\mu = \mu(A), \quad A \in \mathcal{S},$$

zu fordern. Darüber hinaus sollten Additivität und Monotonie des Integrals gelten und das System der integrierbaren Funktionen so groß wie möglich sein. Im folgenden sei  $(X, \mathcal{S}, \mu)$  ein vollständiger Maßraum. Im Falle des Lebesgue Maßes  $\mu$  im  $\mathbb{R}^n$  werden wir die traditionelle Integralnotation, d.h.

$$\int_E f dx := \int_E f d\mu$$

benutzen. Treppenfunktionen haben eine Darstellung

$$f = \sum_{j=1}^n \beta_j \chi_{B_j}$$

mit  $\beta_j \in \mathbb{R}$ ,  $B_j \in \mathcal{S}$ , wobei  $B_j$  paarweise disjunkt sind. Allerdings ist diese Darstellung nicht eindeutig.

**3.1 Lemma.** *Seien  $A_1, \dots, A_m \in \mathcal{S}$  und  $B_1, \dots, B_n \in \mathcal{S}$  paarweise disjunkte Mengen und seien  $\alpha_i, \beta_j$  nichtnegative reelle Zahlen. Dann gilt:*

$$\sum_{i=1}^m \alpha_i \chi_{A_i} \leq \sum_{j=1}^n \beta_j \chi_{B_j} \quad \Rightarrow \quad \sum_{i=1}^m \alpha_i \mu(A_i) \leq \sum_{j=1}^n \beta_j \mu(B_j).$$

BEWEIS : Wir setzen  $\alpha_0 = \beta_0 = 0$ ,  $A_0 = X \setminus \bigcup_{i=1}^m A_i$ ,  $B_0 = X \setminus \bigcup_{j=1}^n B_j$ . Für  $i \in \{0, \dots, m\}, j \in \{0, \dots, n\}$  gilt entweder  $A_i \cap B_j = \emptyset$  oder  $\alpha_i \leq \beta_j$  und somit<sup>2</sup>:

$$\begin{aligned} \sum_{i=0}^m \alpha_i \mu(A_i) &= \sum_{i=0}^m \sum_{j=0}^n \alpha_i \mu(A_i \cap B_j) \leq \sum_{i=0}^m \sum_{j=0}^n \beta_j \mu(A_i \cap B_j) \\ &\leq \sum_{j=0}^n \beta_j \mu(B_j). \end{aligned}$$

■

**3.2 Definition.** *Sei  $D \in \mathcal{S}$  und  $s$  eine nichtnegative Treppenfunktion mit einer Darstellung  $s = \sum_{j=1}^n \beta_j \chi_{B_j}$ , wobei  $B_j$  paarweise disjunkt sind und  $\beta_j \geq 0$ . Wir setzen*

$$\int_D s d\mu := \sum_{j=1}^n \beta_j \mu(D \cap B_j).$$

<sup>2</sup>Hier und im Folgenden benutzen wir die Konvention  $0 \cdot \infty = \infty \cdot 0 = 0$

• Lemma 3.1 zeigt, dass die Definition des Integrals einer Treppenfunktion unabhängig von ihrer Darstellung ist.

**3.3 Definition.** Sei  $f \geq 0$  eine messbare Funktion und  $D \in \mathcal{S}$ . Wir definieren das **Lebesgue Integral**

$$\int_D f d\mu := \sup \left\{ \int_D s d\mu; 0 \leq s \leq f \text{ auf } D, s \text{ Treppenfunktion} \right\}.$$

• Nach Satz 2.7 existiert für jede nichtnegative messbare Funktion  $f$  eine Folge von Treppenfunktionen  $s_n \geq 0$ ,  $s_n \nearrow f$ . Dies zusammen mit Satz 3.6 ist die Idee hinter unserer Definition des Lebesgue Integrals.

• Lemma 3.1 zeigt wiederum, dass die Definitionen 3.2 und 3.3 im Falle einer Treppenfunktion  $f$  übereinstimmen.

**3.4 Definition.** Sei  $f$  eine auf  $D \in \mathcal{S}$  definierte messbare Funktion. Wir setzen

$$\int_D f d\mu := \int_D f^+ d\mu - \int_D f^- d\mu,$$

falls wenigstens eins der Integrale auf der rechten Seite endlich ist.

• Sei  $f$  eine auf  $X$  definierte Funktion und sei  $M \in \mathcal{S}$ . Dann gilt offensichtlich

$$\begin{aligned} \int_M f d\mu &= \int_X f \chi_M d\mu, \\ \int_M f d\mu &= \int_X f d\nu, \end{aligned}$$

wobei  $\nu = \mu|_M$  die Restriktion von  $\mu$  auf die  $\sigma$ -Algebra  $\mathcal{S}_M = \{A \in \mathcal{S}, A \subseteq M\}$  ist. Also ist es keine Einschränkung der Allgemeinheit wenn wir im weiteren nur Integrale über ganz  $X$  betrachten.

• Falls  $f$  nur fast überall auf  $X$  definiert ist, d.h. es existiert ein  $D \in \mathcal{S}$ , so dass  $f$  auf  $D$  definiert ist und  $\mu(X \setminus D) = 0$ , dann setzen wir

$$\int_X f d\mu := \int_D f d\mu,$$

falls das Integral auf der rechten Seite definiert ist.

**3.5 Definition.** Die Menge aller messbaren fast überall auf  $X$  definierten Funktionen  $f$ , deren Integral definiert ist wird mit  $\mathcal{L}^*(\mu)$  oder  $\mathcal{L}^*$  bezeichnet. Weiter bezeichnen wir

$$\mathcal{L}^1 = \mathcal{L}^1(\mu) := \left\{ f \in \mathcal{L}^*(\mu); \int_X f d\mu \in \mathbb{R} \right\},$$

und sagen, dass Elemente  $f \in \mathcal{L}^1(\mu)$  **integrierbare Funktionen** sind.

Das Herzstück der Lebesgueschen Integraltheorie ist der folgende Spezialfall des Satzes über monotone Konvergenz.

**3.6 Satz.** *Seien  $f_n \geq 0$  messbare Funktionen auf  $X$  und gelte  $f_n \nearrow f$ , dann gilt:*

$$\lim_{n \rightarrow \infty} \int_X f_n d\mu = \int_X f d\mu.$$

BEWEIS :  $f$  ist als Grenzwert meßbarer Funktionen wiederum meßbar und nicht negativ. Also sind alle Integrale definiert. Aus  $f_n \leq f_{n+1}$  und der Definition 3.3 folgt, dass  $\{\int_X f_n d\mu\}$  eine nichtfallende Folge reeller Zahlen ist, deren Grenzwert wir mit  $\alpha$  bezeichnen, d.h.  $\alpha = \lim \int_X f_n d\mu$ . Da  $f_n \leq f$  gilt haben wir also  $\alpha \leq \int_X f d\mu$ . Falls  $\alpha = \infty$  ist die Behauptung des Satzes klar. Sei nun  $\alpha < \infty$  und sei  $s$  eine Treppenfunktion mit  $0 \leq s \leq f$ . Wir werden in mehreren Schritten zeigen, dass  $\int_X s d\mu \leq \alpha$  ist, was sofort die Behauptung liefert.

- (a) Sei  $\tau \in (0, 1)$  und sei  $E_n := \{x \in X; f_n(x) \geq \tau s(x)\}$ . Dann ist  $E_n \in \mathcal{S}$ ,  $E_n \subseteq E_{n+1}$  und  $\bigcup_{n=1}^{\infty} E_n = X$ . In der Tat, falls  $f(x) = 0$  dann ist  $x \in E_1$ ; falls  $f(x) > 0$ , dann ist  $\tau s(x) < \lim_{n \rightarrow \infty} f_n(x)$ . Also gilt nach Folgerung 1.12 (ii) für alle  $A \in \mathcal{S}$

$$\lim_{n \rightarrow \infty} \mu(E_n \cap A) = \mu(A).$$

- (b) Sei  $s = \sum_{j=1}^k \beta_j \chi_{B_j}$ , wobei  $B_j$  paarweise disjunkt sind. Dann haben wir

$$\begin{aligned} \int_X f_n d\mu &\geq \int_{E_n} f_n d\mu \geq \int_{E_n} \tau s d\mu = \tau \int_{E_n} \sum_{j=1}^k \beta_j \chi_{B_j} d\mu \\ &= \tau \sum_{j=1}^k \beta_j \mu(B_j \cap E_n). \end{aligned}$$

- (c) Wenn wir in dieser Ungleichung zur Grenze  $n \rightarrow \infty$  übergehen erhalten wir unter Benutzung von (a)

$$\alpha \geq \tau \sum_{j=1}^k \beta_j \mu(B_j) = \tau \int_X s d\mu.$$

Da  $\tau \in (0, 1)$  beliebig war erhalten wir  $\alpha \geq \int_X s d\mu$ . ■

**3.7 Lemma.** *Sei  $g \in \mathcal{L}^*$  und  $f$  eine messbare Funktion mit  $f = g$  fast überall. Dann gilt:*

$$f \in \mathcal{L}^*(\mu) \text{ und } \int_X f d\mu = \int_X g d\mu.$$



BEWEIS : Sei  $D \in \mathcal{S}$  mit  $f = g$  auf  $D$ ,  $\mu(X \setminus D) = 0$ . Dann gilt:

$$\int_X f d\mu = \int_D f d\mu = \int_D g d\mu$$

ist definiert. ■

Dies ist die genaue Formulierung der Aussage: „Mengen mit Maß Null spielen für das Integral keine Rolle.“

**3.8 Lemma.** *Seien  $f_1, f_2$  nichtnegative messbare Funktionen. Dann gilt:*

$$\int_X (f_1 + f_2) d\mu = \int_X f_1 d\mu + \int_X f_2 d\mu.$$

BEWEIS :

(a) Seien  $f_1, f_2$  Treppenfunktionen. Dann gibt es paarweise disjunkte Mengen  $A_0, \dots, A_m, B_0, \dots, B_n$  und nichtnegative reelle Zahlen  $\alpha_i, \beta_j$  mit  $\bigcup_{i=0}^m A_i =$

$\bigcup_{j=0}^n B_j = X$  und  $f_1 = \sum_{i=0}^m \alpha_i \chi_{A_i}, f_2 = \sum_{j=0}^n \beta_j \chi_{B_j}$ . Es gilt:

$$\begin{aligned} \int_X (f_1 + f_2) d\mu &= \sum_{i=0}^m \sum_{j=0}^n (\alpha_i + \beta_j) \mu(A_i \cap B_j) \\ &= \sum_{i=0}^m \alpha_i \mu(A_i) + \sum_{j=0}^n \beta_j \mu(B_j) \\ &= \int_X f_1 d\mu + \int_X f_2 d\mu. \end{aligned}$$

(b) Im allgemeinen Falle finden wir Folgen von nichtnegativen Treppenfunktionen  $\{s_n^1\}, \{s_n^2\}$  mit  $s_n^1 \nearrow f_1, s_n^2 \nearrow f_2$  und benutzen Teil (a) sowie Satz 3.6. ■

**3.9 Satz.** *Es gelten:*

(a) Falls  $f \in \mathcal{L}^1$ , dann ist  $f$  fast überall endlich.

(b) Für  $f, g \in \mathcal{L}^1$  und  $\alpha, \beta \in \mathbb{R}$  gilt:

$$\int_X (\alpha f + \beta g) d\mu = \alpha \int_X f d\mu + \beta \int_X g d\mu.$$

(c) Aus  $f \in \mathcal{L}^1$  folgt  $|f| \in \mathcal{L}^1$  und es gilt die Abschätzung:

$$\left| \int_X f d\mu \right| \leq \int_X |f| d\mu.$$

(d) Aus  $f, g \in \mathcal{L}^1$  folgt  $\max(f, g), \min(f, g) \in \mathcal{L}^1$ .

(e) Sei  $f$  messbar,  $g \in \mathcal{L}^1$  und gelte  $|f| \leq g$ , dann gilt auch  $f \in \mathcal{L}^1$ .

BEWEIS :

- (a) Sei  $f \in \mathcal{L}^1$  und  $A := \{x \in X; f(x) = \infty\}$ . Dann ist  $A$  messbar und es gilt für alle  $n \in \mathbb{N}$  :  $0 \leq n\chi_A \leq f^+$ . Also haben wir  $0 \leq \int_X n\chi_A d\mu \leq \int_X f^+ d\mu < \infty$  und somit

$$\mu(A) \leq \frac{1}{n} \int_X f^+ d\mu, \quad \forall n \in \mathbb{N},$$

was  $\mu(A) = 0$  liefert. Analog geht für die Menge  $B := \{x \in X; f(x) = -\infty\}$  vor.

- (b) Aus der Definition des Integrals folgt sofort  $\int_X \alpha f d\mu = \alpha \int_X f d\mu$  für alle  $f \in \mathcal{L}^1$ ,  $\alpha \in \mathbb{R}$ . Für  $f, g \in \mathcal{L}^1$  schreiben wir  $f = f^+ - f^-$ ,  $g = g^+ - g^-$  und  $h = f + g$  (aufgrund von (a) ist  $h$  fast überall definiert). Es gilt:

$$h = h^+ - h^- = f^+ - f^- + g^+ - g^-$$

und Lemma 3.8 liefert, da alle Funktionen messbar sind,

$$\int_X h^+ d\mu + \int_X f^- d\mu + \int_X g^- d\mu = \int_X f^+ d\mu + \int_X g^+ d\mu + \int_X h^- d\mu.$$

Die Behauptung folgt nun, wenn wir beachten, das  $\int_X h^+ d\mu$  und  $\int_X h^- d\mu$  endlich sind. Dies ist aber der Fall, da

$$0 \leq h^+ = (f + g)^+ \leq f^+ + g^+$$

und die rechte Seite eine integrierbare Funktion ist.

- (c) Für  $f \in \mathcal{L}^1$  gilt:  $|f| = f^+ + f^- \in \mathcal{L}^1$  nach (b). Also

$$\begin{aligned} \left| \int_X f d\mu \right| &= \left| \int_X f^+ d\mu - \int_X f^- d\mu \right| \\ &\leq \left| \int_X f^+ d\mu \right| + \left| \int_X f^- d\mu \right| \\ &= \int_X f^+ d\mu + \int_X f^- d\mu = \int_X |f| d\mu. \end{aligned}$$

- (d) Die Behauptung folgt sofort aus (b) und (c), wenn man beachtet, dass

$$\max(f, g) = \frac{1}{2}(f + g + |f - g|).$$

Analog für  $\min(f, g)$ .

(e) Für eine messbare Funktion  $f$  ist auch  $f^+$  messbar und es gilt:

$$0 \leq \int_X f^+ d\mu \leq \int_X g d\mu < \infty,$$

da  $0 \leq f^+ \leq |f| \leq g$ . Also ist  $f^+ \in \mathcal{L}^1$  und analog zeigt man  $f^- \in \mathcal{L}^1$ . Da  $f = f^+ - f^-$  gilt, folgt die Behauptung. ■

Sätze über das Vertauschen von Integral und Grenzwert haben eine zentrale Bedeutung in der Integrationstheorie und ihren Anwendungen.

**3.10 Satz (Levi, monotone Konvergenz).** Sei  $\{f_n\}$  eine Folge messbarer Funktionen mit  $f_n \nearrow f$  fast überall und sei  $\int_X f_1 d\mu > -\infty$ . Dann gilt:

$$\lim_{n \rightarrow \infty} \int_X f_n d\mu = \int_X f d\mu.$$

BEWEIS : Der Satz wurde schon im Spezialfall  $f_n \geq 0$  und  $f_n \nearrow f$  überall bewiesen (Satz 3.6). Der allgemeine Fall kann darauf zurückgeführt werden. Wir ändern  $f_n, f$  auf einer Menge vom Maß Null derart, dass  $f_n \nearrow f$  überall. Aus  $f_n \geq f_1$  und  $f_1 \in \mathcal{L}^*$  folgt  $f_n \in \mathcal{L}^*$  für alle  $n \in \mathbb{N}$ . Falls für  $n \geq n_0$   $\int_X f_n d\mu = \infty$  ist die Behauptung klar. Sei also  $f_n \in \mathcal{L}^1$  für  $n \in \mathbb{N}$ . Wir setzen  $g_n \equiv f_n + f_1^- \in \mathcal{L}^1$ , da  $f_1^- \in \mathcal{L}^1$  nach Voraussetzung. Die Folge  $g_n$  besteht aus nichtnegativen integrierbaren Funktionen mit  $g_n \nearrow f + f_1^-$ . Satz 3.6 liefert also

$$\int_X g_n d\mu \longrightarrow \int_X (f + f_1^-) d\mu.$$

Da  $f^- \leq f_1^- \in \mathcal{L}^1$  und  $f + f_1^- = f^+ + (f_1^- - f^-)$  liefert Lemma 3.8

$$\begin{aligned} \int_X (f - f_1^-) d\mu &= \int_X f^+ d\mu + \int_X (f_1^- - f^-) d\mu \\ &= \int_X f^+ d\mu - \int_X f^- d\mu + \int_X f_1^- d\mu \\ &= \int_X f d\mu + \int_X f_1^- d\mu. \end{aligned}$$

Dies zusammen mit

$$\int_X f_n d\mu = \int_X g_n d\mu - \int_X f_1^- d\mu \rightarrow \int_X (f + f_1^-) d\mu - \int_X f_1^- d\mu$$

liefert die Behauptung. ■

**3.11 Satz.** Für nichtnegative messbare Funktionen  $f_n$  gilt

$$\int_X \sum_{n=1}^{\infty} f_n d\mu = \sum_{n=1}^{\infty} \int_X f_n d\mu.$$

BEWEIS : Folgt sofort aus Satz 3.6 und Lemma 3.8. ■

**3.12 Satz (dominierte Konvergenz).** Sei  $\{f_n\}$  eine Folge messbarer Funktionen mit  $f_n \rightarrow f$  fast überall. Wenn es eine Funktion  $h \in \mathcal{L}^1$  gibt mit  $|f_n| \leq h$  fast überall für alle  $n \in \mathbb{N}$ , dann gilt  $f \in \mathcal{L}^1$  und

$$\lim_{n \rightarrow \infty} \int_X f_n d\mu = \int_X f d\mu.$$

BEWEIS : Der Beweis kann auf Satz 3.10 zurückgeführt werden, wenn man beachtet, dass  $f = \limsup f_n = \liminf f_n$ . Wir setzen  $s_n = \sup\{f_n, f_{n+1}, \dots\}$ ,  $t_n = \inf\{f_n, f_{n+1}, \dots\}$ . Dann gilt  $-h \leq t_n \leq s_n \leq h$ ,  $s_n \searrow f, t_n \nearrow f$  fast überall und

$$-\infty < \int_X -h d\mu \leq \int_X t_n d\mu \leq \int_X s_n d\mu \leq \int_X h d\mu < \infty.$$

Satz 3.10 liefert also

$$\int_X f d\mu = \lim_{n \rightarrow \infty} \int_X t_n d\mu = \lim_{n \rightarrow \infty} \int_X s_n d\mu,$$

was zusammen mit

$$\int_X t_n d\mu \leq \int_X f_n d\mu \leq \int_X s_n d\mu$$

die Behauptung gibt. ■

**3.13 Folgerung.** Sei  $\{h_n\}$  eine Folge messbarer Funktionen und  $g \in \mathcal{L}^1$ . Falls die Reihe  $\sum_{j=1}^{\infty} h_j$  fast überall konvergiert und falls für alle  $n \in \mathbb{N}$  gilt

$$\left| \sum_{j=1}^n h_j \right| \leq g \quad \text{fast überall,}$$

dann gilt auch

$$\int_X \sum_{j=1}^{\infty} h_j d\mu = \sum_{j=1}^{\infty} \int_X h_j d\mu.$$

BEWEIS : Folgt sofort aus 3.12. ■

**3.14 Lemma (Fatou).** Sei  $\{f_n\}$  eine Folge messbarer Funktionen und  $g \in \mathcal{L}^1$ . Aus  $f_n \geq g$  fast überall für alle  $n \in \mathbb{N}$  folgt

$$\int_X \liminf_{n \rightarrow \infty} f_n d\mu \leq \liminf_{n \rightarrow \infty} \int_X f_n d\mu.$$

BEWEIS : Wir setzen  $g_k := \inf\{f_n; k \leq n\}$ . Dann haben wir  $g_k \nearrow \liminf_{n \rightarrow \infty} f_n$  und wir können Satz 3.10 anwenden. ■

**3.15 Satz.** Sei  $f \geq 0$  eine messbare Funktion. Aus  $\int_X f d\mu = 0$  folgt  $f = 0$  fast überall.

BEWEIS : Wir setzen  $A_n \equiv \{x \in X; f(x) \geq \frac{1}{n}\}$ . Aus  $\bigcup_{n=1}^{\infty} A_n = \{x \in X, f(x) > 0\}$  und  $0 \leq \chi_{A_n} \leq nf$  folgt sofort  $\mu(A_n) = 0$  und somit die Behauptung. ■

**3.16 Folgerung.** Sei  $f \in \mathcal{L}^1$  und  $\int_E f d\mu = 0$  für alle messbaren Mengen  $E$ . Dann folgt  $f = 0$  fast überall.

BEWEIS : Wir setzen  $E \equiv \{f \geq 0\}$ . Dann haben wir  $0 = \int_E f d\mu = \int_E f^+ d\mu$ . Aus Satz 3.15 folgt  $f^+ = 0$  fast überall. Analog für  $f^-$ . ■

**3.17 Folgerung.** Seien  $f, g \in \mathcal{L}^1$ . Aus

$$\int_E f d\mu \leq \int_E g d\mu$$

für alle messbaren Mengen  $E$  folgt  $f \leq g$  fast überall.

BEWEIS : Wir setzen  $h = (f - g)^+$ . Dann gilt  $h \geq 0$  und

$$0 \leq \int_E h d\mu = \int_{E \cap \{h > 0\}} (f - g) d\mu \leq 0.$$

Satz 3.15 liefert  $(f - g)^+ = 0$  fast überall, d.h.  $f - g \leq 0$ . ■

**3.18 Definition.** Sei  $1 \leq p < \infty$ . Wir bezeichnen mit  $\mathcal{L}^p = \mathcal{L}^p(X, \mathcal{S}, \mu)$  die Menge aller messbaren Funktionen  $f$  mit

$$\int_X |f|^p d\mu < \infty.$$

Die Größe

$$\|f\|_p \equiv \left( \int_X |f|^p d\mu \right)^{\frac{1}{p}} \quad (3.19)$$

heißt  $L^p$ -Norm der Funktion  $f \in \mathcal{L}^p$ . Mit  $\mathcal{L}^\infty = \mathcal{L}^\infty(X, \mathcal{S}, \mu)$  bezeichnen wir die Menge aller messbaren Funktionen  $f$  für die eine Konstante  $K \in \mathbb{R}$  existiert mit

$$|f| \leq K \quad \text{fast überall.}$$

Die Größe

$$\|f\|_\infty \equiv \operatorname{ess\,sup}_{x \in X} |f(x)| \equiv \inf_{\substack{M \in \mathcal{S}, \\ \mu(M) = 0}} \sup_{x \in X \setminus M} |f(x)| \quad (3.20)$$

heißt  $L^\infty$ -Norm der Funktion  $f \in \mathcal{L}^\infty$ .

Im folgenden werden wir zeigen, dass die so definierten Größen wirklich Normen sind.

**3.21 Lemma (Young–Ungleichung).** Sei  $p, q \in (1, \infty)$  mit  $q^{-1} + p^{-1} = 1$ . Für nichtnegative Zahlen  $a, b$  gilt

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

BEWEIS : Sei  $ab > 0$ . Die Konkavität der Funktionen  $\ln$  liefert

$$\ln \left( \frac{a^p}{p} + \frac{b^q}{q} \right) \geq \frac{1}{p} \ln a^p + \frac{1}{q} \ln b^q = \ln a + \ln b = \ln(ab).$$

Da  $\ln$  monoton wachsend ist folgt die Behauptung. ■

**3.22 Lemma (Hölder–Ungleichung).** Sei  $f \in \mathcal{L}^p$  und  $g \in \mathcal{L}^q$  mit  $p, q \in [1, \infty]$ ,  $q^{-1} + p^{-1} = 1$ .<sup>3</sup> Dann ist  $fg \in \mathcal{L}^1$  und es gilt:

$$\left| \int_X fg \, d\mu \right| \leq \|f\|_p \|g\|_q.$$

BEWEIS : Der Fall  $p = 1$  ist klar. Sei also  $p \in (1, \infty)$ . Sei  $s = \|f\|_p, t = \|g\|_q$ . Wir können annehmen, dass  $st > 0$ . Aus Lemma 3.21 mit  $a = |f(x)|/s, b = |g(x)|/t$  folgt für alle  $x \in X$

$$\frac{f(x)g(x)}{st} \leq \frac{|f(x)||g(x)|}{st} \leq \frac{|f(x)|^p}{ps^p} + \frac{|g(x)|^q}{qt^q}.$$

Also gilt:

$$\left| \frac{1}{st} \int_X fg \, d\mu \right| \leq \frac{1}{ps^p} \int_X |f|^p \, d\mu + \frac{1}{qt^q} \int_X |g|^q \, d\mu = \frac{1}{p} + \frac{1}{q} = 1,$$

was die Behauptung ist. ■

**3.23 Lemma (Minkowski–Ungleichung).** Sei  $p \in [1, \infty]$  und  $f, g \in \mathcal{L}^p$ . Dann ist  $f + g \in \mathcal{L}^p$  und es gilt:

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

<sup>3</sup>Wir benutzen die Konvention, dass  $p = 1, q = \infty$  in der Identität  $q^{-1} + p^{-1} = 1$  enthalten ist.

BEWEIS : Der Fall  $p = 1$  folgt aus der Dreiecksungleichung in  $\mathbb{R}$ . Für  $p = \infty$  sei  $|f| \leq M$  fast überall und  $|g| \leq N$  fast überall. Dann ist  $|f + g| \leq M + N$  fast überall und die Behauptung folgt. Sei also  $p \in (1, \infty)$ . Aus Lemma 3.22 folgt mit  $q = \frac{p}{p-1}$

$$\begin{aligned} \int_X |f| |f + g|^{p-1} d\mu &\leq \left( \int_X |f|^p d\mu \right)^{\frac{1}{p}} \left( \int_X |f + g|^{(p-1)q} d\mu \right)^{\frac{1}{q}} \\ &= \|f\|_p \|f + g\|_p^{p-1} \end{aligned}$$

und analog

$$\int_X |g| |f + g|^{p-1} d\mu \leq \|g\|_p \|f + g\|_p^{p-1}.$$

Somit erhalten wir

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

welches die Behauptung ist. ■

**3.24 Die  $L^p$ -Räume:** Lemma 3.23 zeigt, dass die Funktion  $f \rightarrow \|f\|_p$  definiert auf  $\mathcal{L}^p$ ,  $p \in [1, \infty]$  alle Eigenschaften einer Norm hat (Nichtnegativität, positive Homogenität, Dreiecksungleichung). Allerdings gibt es nichttriviale Funktionen  $g \neq 0$  mit  $\|g\|_p = 0$ , nämlich alle Funktionen aus

$$N = \{g \in \mathcal{L}^p; g = 0 \text{ fast überall}\}.$$

Um dieses Problem zu lösen geht man zum Restklassenraum über, d.h. man betrachtet

$$[f] \equiv \{g \in \mathcal{L}^p; g = f \text{ fast überall}\}$$

und definiert

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

Auf  $L^p$  definiert man die Operationen

$$[f] + [g] \equiv [f + g], \quad \alpha[f] \equiv [\alpha f]$$

und die Norm

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

Diese Operationen hängen offensichtlich nicht vom Repräsentanten ab. Im weiteren werden wir nicht zwischen  $[f]$  und  $f$  unterscheiden.

Wir sagen, dass eine Folge  $\{f_n\} \in L^p$  in  $L^p$  gegen  $f \in L^p$  **konvergiert**<sup>4</sup> genau dann, wenn

$$\|f_n - f\|_p \rightarrow 0 \quad (n \rightarrow \infty).$$

Man schreibt  $f_n \rightarrow f$  in  $L^p$  und spricht von der durch die Norm definierten Konvergenz oder von der **starken Konvergenz** in  $L^p$ .

<sup>4</sup>Man sagt auch bezüglich  $\|\cdot\|_p$  konvergiert.

## 1.4 Der Glättungsoperator

Der Satz über dominierte Konvergenz hat einfache aber wichtige Konsequenzen für Parameterintegrale.

**4.1 Satz.** Sei  $(X, \mathcal{S}, \mu)$  ein Maßraum,  $P$  ein metrischer Raum und  $U$  eine offene Umgebung eines Punktes  $a \in P$ . Sei  $F : U \times X \rightarrow \mathbb{R}$  eine Funktion mit den Eigenschaften:

- i) es existiert eine Nullmenge  $N \subseteq X$ , so dass für alle Punkte  $x \in X \setminus N$  die Funktion  $F(., x)$  stetig in  $a$  ist;
- ii) für alle  $t \in U$  ist die Funktion  $F(t, .)$  messbar;
- iii) es existiert eine Funktion  $g \in L^1(X)$  so, dass für alle  $t \in U$

$$|F(t, .)| \leq g$$

fast überall gilt.

Dann ist für alle  $t \in U$  die Funktion  $F(t, .)$  integrierbar, d.h.  $F(t, .) \in L^1(X)$ , und die Funktion

$$f : t \mapsto \int_X F(t, .) d\mu$$

ist stetig in  $a$ .

BEWEIS : Aus den Voraussetzungen des Satzes folgt sofort, dass  $F(t, .) \in L^1(X)$  nach Satz 3.9 (e). Die Stetigkeit der Funktion  $f$  im Punkt  $a$  ist äquivalent zu

$$\lim_{j \rightarrow \infty} \int_X F(t_j, .) d\mu = \int_X F(a, .) d\mu$$

für alle Folgen  $t_j \rightarrow a$ ,  $t_j \in U$ . Dies folgt aber sofort aus Satz 3.12. ■

**4.2 Satz.** Sei  $(X, \mathcal{S}, \mu)$  ein Maßraum,  $N \subseteq X$  eine Nullmenge und  $I \subseteq \mathbb{R}$  ein offenes Intervall. Sei  $F : I \times X \rightarrow \mathbb{R}$  eine Funktion mit den Eigenschaften:

- i) für alle  $x \in X \setminus N$  ist  $F(., x)$  differenzierbar in  $I$ ;
- ii) für alle  $t \in I$  ist  $F(t, .)$  messbar;
- iii) es existiert eine Funktion  $g \in L^1(X)$  mit

$$\left| \frac{d}{dt} F(t, x) \right| \leq g(x) \quad \text{für alle } x \in X \setminus N, \text{ für alle } t \in I;$$

- iv) es gibt ein  $t_0 \in I$  so, dass  $F(t_0, .) \in L^1(X)$ .



Dann ist  $F(t, \cdot)$  integrierbar für alle  $t \in I$  und die Funktion

$$f : t \mapsto \int_X F(t, \cdot) d\mu$$

ist differenzierbar auf  $I$  mit der Ableitung

$$f'(t) = \frac{d}{dt} \int_X F(t, \cdot) d\mu = \int_X \frac{d}{dt} F(t, \cdot) d\mu.$$

BEWEIS : Seien  $a, b \in I$ ,  $b \neq a$ ,  $x \in X \setminus N$ . Aus dem Mittelwertsatz folgt die Existenz von  $\xi$  zwischen  $a, b$  mit

$$\left| \frac{F(b, x) - F(a, x)}{b - a} \right| = \left| \frac{d}{dt} F(\xi, x) \right| \leq g(x).$$

Für  $a = t_0$  erhalten wir also aus Satz 3.9 (e), dass die Funktion

$$x \mapsto \frac{F(b, x) - F(a, x)}{b - a}$$

integrierbar ist und somit auch  $F(b, \cdot)$ . Für  $a \in I$  folgt aus dem Satz über dominierte Konvergenz

$$\lim_{j \rightarrow \infty} \int_X \frac{F(t_j, \cdot) - F(a, \cdot)}{t_j - a} d\mu = \int_X \frac{d}{dt} F(a, \cdot) d\mu$$

für beliebige Folgen  $t_j \rightarrow a$ ,  $t_j \in I \setminus \{a\}$ . Dies liefert sofort

$$\frac{d}{dt} \int_X F(a, \cdot) d\mu = \lim_{t \rightarrow a} \int_X \frac{F(t, \cdot) - F(a, \cdot)}{t - a} d\mu = \int_X \frac{d}{dt} F(a, \cdot) d\mu$$

■

**4.3 Definition.** Sei  $\Omega \subseteq \mathbb{R}^n$  eine offene Menge. Sei  $J$  eine nichtnegative Funktion aus  $C_0^\infty(\Omega)$  mit den Eigenschaften

i)  $J(x) = 0$ , falls  $\|x\| \geq 1$ ;

ii)  $\int_{\mathbb{R}^n} J(x) dx = 1$ .

Für  $\varepsilon > 0$  definieren wir  $J_\varepsilon(x) \equiv \varepsilon^{-n} J(x/\varepsilon)$ , welche die Eigenschaften

i)  $J_\varepsilon(x) = 0$ , falls  $\|x\| \geq \varepsilon$ ;

ii)  $\int_{\mathbb{R}^n} J_\varepsilon(x) dx = 1$

hat.  $J_\varepsilon$  heißt **Glättungsoperator** und die Konvolution

$$(J_\varepsilon * f)(x) \equiv \int_{\mathbb{R}^n} J_\varepsilon(x-y)f(y) dy,$$

welche für Funktionen  $f$  definiert ist, für die die rechte Seite Sinn macht, heißt **Regularisierung von  $f$** .

- Ein typisches Beispiel für  $J$  ist

$$J(x) = \begin{cases} k \exp\left(\frac{-1}{1-\|x\|^2}\right), & \text{falls } \|x\| \leq 1, \\ 0, & \text{falls } \|x\| \geq 1, \end{cases}$$

wobei  $k$  eine Normierungskonstante ist.

Bevor wir die Eigenschaften des Glättungsoperators untersuchen, zeigen wir, dass man  $L^p$ -Funktionen beliebig gut durch stetige Funktionen annähern kann.

**4.4 Satz.** Sei  $\Omega \subseteq \mathbb{R}^n$  eine offene Menge und  $1 \leq p < \infty$ . Die Menge der stetigen Funktionen mit kompakten Trägern  $C_0(\Omega)$  ist dicht in  $L^p(\Omega)$ , d.h. für alle  $f \in L^p(\Omega)$  und alle  $\varepsilon > 0$  existiert ein  $\varphi \in C_0(\Omega)$  mit  $\|f - \varphi\|_p \leq \varepsilon$ .

BEWEIS : Aufgrund der Zerlegung  $f = f^+ - f^-$  reicht es, den Satz für nichtnegative Funktionen, zu zeigen. Sei  $0 \leq f \in L^p(\Omega)$  und  $\varepsilon > 0$ . Satz 2.7 liefert die Existenz einer Folge nichtnegativer Treppenfunktionen  $s_n$ , mit  $s_n \nearrow f$  punktweise. Da  $0 \leq (s_n)^p \leq (f)^p$  gilt, erhalten wir  $s_n \in L^p(\Omega)$ . Aus  $(f(x) - s_n(x))^p \leq f(x)^p$  und dem Satz über dominierte Konvergenz erhalten wir  $s_n \rightarrow f$  in  $L^p(\Omega)$ . Sei  $s_{n_0}$  derart gewählt, dass

$$\|f - s_{n_0}\|_p \leq \frac{\varepsilon}{2}.$$

Da  $s_{n_0}$  eine Treppenfunktion ist und  $p < \infty$ , muss der Träger von  $s_{n_0}$  endliches Maß haben. Wir können annehmen, dass  $s_{n_0}(x) = 0$  für  $x \in \mathbb{R}^n \setminus \Omega$ . Folgerung 2.5 liefert also eine Funktion  $\varphi \in C_0(\Omega)$  mit

$$\begin{aligned} |\varphi(x)| &\leq \|s_{n_0}\|_\infty \quad \forall x \in \Omega \\ \mu(\{x \in \Omega; s_{n_0}(x) \neq \varphi(x)\}) &< \left(\frac{\varepsilon}{4\|s_{n_0}\|_\infty}\right)^p. \end{aligned}$$

Die Hölder-Ungleichung liefert

$$\begin{aligned} \|s_{n_0} - \varphi\|_p &= \left(\int_{\Omega \cap \{s_{n_0} \neq \varphi\}} |s_{n_0} - \varphi|^p \cdot 1 d\mu\right)^{\frac{1}{p}} \\ &\leq \|s_{n_0} - \varphi\|_\infty \mu(\{x \in \Omega; s_{n_0}(x) \neq \varphi(x)\})^{\frac{1}{p}} \\ &\leq 2 \|s_{n_0}\|_\infty \frac{\varepsilon}{4\|s_{n_0}\|_\infty} = \frac{\varepsilon}{2}. \end{aligned}$$

Und somit haben wir  $\|f - \varphi\|_p \leq \varepsilon$ . ■

**4.5 Satz.** Sei  $\Omega$  eine offene Menge des  $\mathbb{R}^n$ ,  $f \in \mathcal{L}^*(\mathbb{R}^n)$  sei außerhalb  $\Omega$  identisch Null und  $\varepsilon > 0$ .

(i) Für  $f \in L^1(\Omega)$  gilt  $J_\varepsilon * f \in C^\infty(\mathbb{R}^n)$ .

(ii) Falls <sup>5</sup>  $\text{supp}(f) \subseteq \subseteq \Omega$ , dann ist  $J_\varepsilon * f \in C_0^\infty(\Omega)$  für  $\varepsilon < \text{dist}(\text{supp}(f), \partial\Omega)$ .

(iii) Für  $f \in L^p(\Omega)$ ,  $1 \leq p < \infty$ , gilt:  $J_\varepsilon * f \in L^p(\Omega)$  und

$$\|J_\varepsilon * f\|_p \leq \|f\|_p, \\ \lim_{\varepsilon \rightarrow 0^+} \|J_\varepsilon * f - f\|_p = 0.$$

(iv) Für  $f \in C(\Omega)$  und  $K \subseteq \subseteq \Omega$  gilt:

$$\lim_{\varepsilon \rightarrow 0^+} J_\varepsilon * f(x) = f(x) \quad \text{gleichmäßig auf } K.$$

BEWEIS : Da  $J_\varepsilon(x-y)$  unendlich oft differenzierbar ist, liefert die wiederholte Anwendung von Satz 4.2 die Identität

$$D^\alpha(J_\varepsilon * f)(x) = \int_{\mathbb{R}^n} D_x^\alpha J_\varepsilon(x-y) f(y) dy, \quad (4.6)$$

wobei  $\alpha = (\alpha_1, \dots, \alpha_n)$  ein Multiindex ist und  $D^\alpha \equiv \partial_{x_1}^{\alpha_1}, \dots, \partial_{x_n}^{\alpha_n}$  gilt. Aus (4.6) folgt sofort die Behauptung (i). Da  $J_\varepsilon(x-y)$  für  $\|x-y\| \geq \varepsilon$  verschwindet erhält man auch (ii). Sei  $f \in L^p(\Omega)$  und  $1 < p < \infty$ . Wir setzen  $p' = p/(p-1)$ , d.h.  $p^{-1} + (p')^{-1}$ , und erhalten aus der Hölder-Ungleichung

$$|J_\varepsilon * f(x)| \leq \left( \int_{\mathbb{R}^n} J_\varepsilon(x-y) dy \right)^{\frac{1}{p'}} \left( \int_{\mathbb{R}^n} J_\varepsilon(x-y) |f(y)|^p dy \right)^{\frac{1}{p}} \\ = \left( \int_{\mathbb{R}^n} J_\varepsilon(x-y) |f(y)|^p dy \right)^{\frac{1}{p}}.$$

Diese Ungleichung zusammen mit dem Satz von Fubini über das Vertauschen der Integrationsreihenfolge (siehe Evans S.22) liefert

$$\int_{\Omega} |J_\varepsilon * f(x)|^p dx = \int_{\Omega} \left| \int_{\mathbb{R}^n} J_\varepsilon(x-y) f(y) dy \right|^p dx \\ \leq \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} J_\varepsilon(x-y) |f(y)|^p dy dx \quad (4.7) \\ = \int_{\mathbb{R}^n} |f(y)|^p \int_{\mathbb{R}^n} J_\varepsilon(x-y) dx dy = \|f\|_p^p.$$

<sup>5</sup>Man schreibt  $K \subseteq \subseteq \Omega$  falls  $\overline{K} \subseteq \Omega$  und  $\overline{K}$  kompakt ist.

Sei  $\eta > 0$ . Nach Satz 4.4 existiert  $\varphi \in C_0(\Omega)$  so, dass

$$\|f - \varphi\|_p < \frac{\eta}{3}.$$

Also impliziert (4.7) angewandt auf  $\varphi - f$

$$\|J_\varepsilon * \varphi - J_\varepsilon * f\|_p < \frac{\eta}{3}.$$

Weiter gilt:

$$\begin{aligned} |J_\varepsilon * \varphi(x) - \varphi(x)| &= \left| \int_{\mathbb{R}^n} J_\varepsilon(x-y)(\varphi(y) - \varphi(x)) dy \right| \\ &\leq \sup_{\|x-y\| < \varepsilon} |\varphi(y) - \varphi(x)|. \end{aligned} \quad (4.8)$$

Da  $\varphi$  gleichmäßig stetig auf  $\Omega$  ist, konvergiert die rechte Seite in (4.8) gegen Null für  $\varepsilon \rightarrow 0^+$ . Da der Träger von  $\varphi$  kompakt ist, können wir  $\varepsilon$  derart wählen, dass

$$\|J_\varepsilon * \varphi - \varphi\|_p < \frac{\eta}{3}.$$

Insgesamt haben wir also  $\|f - J_\varepsilon * f\|_p < \eta$ . Somit ist (iii) für  $1 < p < \infty$  bewiesen. Im Falle  $p = 1$  folgt (4.7) direkt aus der Definition von  $J_\varepsilon * f$  ohne die Hölder-Ungleichung. Danach verläuft der Beweis analog. Die Behauptung (iv) folgt sofort, wenn man in (4.8)  $\varphi$  durch  $f$  ersetzt. ■

**4.9 Satz.** Sei  $\Omega \subseteq \mathbb{R}^n$  eine offene Menge. Dann ist  $C_0^\infty(\Omega)$  dicht in  $L^p(\Omega)$  für  $1 \leq p < \infty$ .

BEWEIS : Dies folgt sofort aus Satz 4.4 und Satz 4.5 (ii), (iv). ■

# Kapitel 2

## Hilberträume

### 2.1 Der Hilbertraum

**1.1 Definition.** Ein **Prä-Hilbertraum**  $H$  („inner product space“) ist ein Vektorraum über  $\mathbb{R}$  bzw.  $\mathbb{C}$ , in dem ein **Skalarprodukt**  $(\cdot, \cdot)$  gegeben ist, d.h. eine Abbildung von  $H \times H$  nach  $\mathbb{R}$  bzw.  $\mathbb{C}$  mit den Eigenschaften

(i) *Bilinearität bzw. Sesquilinearität*

$$\begin{aligned}(\alpha u + \beta v, w) &= \alpha(u, w) + \beta(v, w), \\(u, \alpha z + \beta w) &= \bar{\alpha}(u, z) + \bar{\beta}(u, w),^1\end{aligned}$$

für alle  $\alpha, \beta \in \mathbb{R}$  bzw.  $\mathbb{C}$  und alle  $u, v, w, z \in H$ .

(ii)  $(\cdot, \cdot)$  ist *symmetrisch bzw. hermitesch*, d.h.

$$(u, v) = \overline{(v, u)} \quad \text{für alle } u, v \in H.$$

(iii)  $(\cdot, \cdot)$  ist *positiv definit*, d.h.

$$(u, u) \geq 0 \quad \text{für alle } u \in H,$$

und  $(u, u) = 0$  genau für  $u = 0$ .

• Wir benutzen die Bezeichnung  $\|u\| \equiv \sqrt{(u, u)}$  und werden zeigen, dass  $\|\cdot\|$  alle Eigenschaften einer Norm hat.

Wir notieren einige Standardfolgerungen aus den Eigenschaften des Skalarproduktes. Zunächst sei darauf hingewiesen, dass im komplexen Prä-Hilbertraum die Identität

$$\|a + b\|^2 = \|a\|^2 + (a, b) + (b, a) + \|b\|^2 = \|a\|^2 + 2 \operatorname{Re}(a, b) + \|b\|^2$$

( $\operatorname{Re} =$  Realteil) gilt, da  $(b, a) = \overline{(a, b)}$ .

**1.2 Lemma (Cauchy-Schwarzsche Ungleichung).** Sei  $H$  ein Prä-Hilbertraum. Dann gilt für alle  $u, v \in H$

$$|(u, v)| \leq \|u\| \|v\|.$$

---

<sup>1</sup> $\bar{\alpha}$  ist die konjugiert komplexe Zahl zu  $\alpha$ .

BEWEIS (für den reellen Fall): Es gilt

$$0 \leq \left( \alpha u - \frac{1}{\alpha} v, \alpha u - \frac{1}{\alpha} v \right) = |\alpha|^2 \|u\|^2 + \frac{1}{|\alpha|^2} \|v\|^2 - 2 \left( \alpha u, \frac{1}{\alpha} v \right),$$

und somit

$$2(u, v) \leq |\alpha|^2 \|u\|^2 + \frac{1}{|\alpha|^2} \|v\|^2. \quad (1.3)$$

Mit  $\alpha = (\|v\|/\|u\|)^{\frac{1}{2}}$ ,  $u \neq 0$ , folgt

$$(u, v) \leq \|u\| \|v\|.$$

Übergang von  $u$  zu  $-u$  ergibt die behauptete Ungleichung. ■

BEWEIS (für den reellen und komplexen Fall): Es gilt mit  $\alpha \in \mathbb{R} \setminus \{0\}$ ,  $\rho \in \mathbb{C}$

$$0 \leq \left( \alpha u - \frac{\rho}{\alpha} v, \alpha u - \frac{\rho}{\alpha} v \right) = |\alpha|^2 \|u\|^2 - 2 \operatorname{Re}(u, \rho v) + \frac{|\rho|^2}{|\alpha|^2} \|v\|^2$$

Man setze  $\alpha = (\|v\|/\|u\|)^{\frac{1}{2}}$ . Dann folgt

$$\operatorname{Re}(u, \rho v) \leq \|u\| \|v\| \quad \text{und} \quad \operatorname{Re}(\bar{\rho}(u, v)) \leq \|u\| \|v\|$$

Setze  $\rho = \frac{(u, v)}{|(u, v)|}$ , falls  $(u, v) \neq 0$ . Dann folgt

$$|(u, v)| \leq \|u\| \|v\|.$$

■

**1.4 Lemma.** Sei  $H$  ein Prä-Hilbert Raum mit Skalarprodukt  $(\cdot, \cdot)$ .  $\|\cdot\|$  besitzt folgende Eigenschaften:

- (i)  $\|\alpha u\| = |\alpha| \|u\|$ , (Homogenität),
- (ii)  $\|u + v\| \leq \|u\| + \|v\|$ , (Dreiecksungleichung).

BEWEIS :

$$(i) \quad \|\alpha u\|^2 = (\alpha u, \alpha u) = \alpha \bar{\alpha} (u, u) = |\alpha|^2 \|u\|^2$$

$$\begin{aligned} (ii) \quad \|u + v\|^2 &= \|u\|^2 + 2 \operatorname{Re}(u, v) + \|v\|^2 \\ &\leq \|u\|^2 + 2 |(u, v)| + \|v\|^2 \\ &\leq \|u\|^2 + 2 \|u\| \|v\| + \|v\|^2 = (\|u\| + \|v\|)^2. \end{aligned}$$

■

Lemma 1.4 und Definition 1.1 rechtfertigen folgende Definition.

**1.5 Definition.**  $\|u\| := \sqrt{(u, u)}$  ist die durch das Skalarprodukt in  $H$  induzierte Norm.

**1.6 Definition.** Eine Folge  $(u_j)_{j \in \mathbb{N}}$ ,  $u_j \in H$  **konvergiert bezüglich**  $\|\cdot\|$  (man sagt auch „in  $H$ “) gegen  $u \in H$  genau dann, wenn

$$\|u_j - u\| \rightarrow 0 \quad (j \rightarrow \infty).$$

Man schreibt  $u_j \rightarrow u$  in  $H$  und spricht von der durch die Norm definierten Konvergenz oder von der **starken Konvergenz** in  $H$ .

**1.7 Lemma.** Das Skalarprodukt ist bezüglich der Normkonvergenz stetig.

BEWEIS : Seien  $u_j \rightarrow u$ ,  $w_j \rightarrow w$  in  $H$ . Somit ist  $\|u_j\| \leq K$  für eine geeignete Konstante  $K$ . Also gilt:

$$\begin{aligned} |(u_j, w_j) - (u, w)| &\leq |(u_j, w_j) - (u_j, w) + (u_j, w) - (u, w)| \\ &\leq \|u_j\| \|w_j - w\| + \|w\| \|u_j - u\| \\ &\leq K \|w_j - w\| + \|w\| \|u_j - u\|. \end{aligned}$$

Daher gilt  $(u_j, w_j) \rightarrow (u, w)$ , welches die Behauptung ist. ■

**1.8 Definition.** Eine Folge  $(u_j)_{j \in \mathbb{N}}$ ,  $u_j \in H$  heisst **Cauchyfolge** (bezüglich  $\|\cdot\|$ ) genau dann, wenn  $\forall \varepsilon > 0 \exists l_0 \in \mathbb{N}$  mit:  $\forall j, k \geq l_0$  gilt

$$\|u_j - u_k\| < \varepsilon.$$

Man schreibt auch etwas unpräzise  $\|u_j - u_k\| \rightarrow 0$  für  $j, k \rightarrow \infty$ .

**1.9 Definition.** Ein **Hilbertraum** ist ein Prä-Hilbertraum, in dem jede Cauchyfolge einen Limes besitzt. Man sagt auch: „Der Hilbertraum ist vollständig.“

**Beispiele:**

- a)  $\mathbb{R}^n$  oder  $\mathbb{C}^n$  mit dem Euklidischen Skalarprodukt  $(\cdot, \cdot)$  sind Hilberträume.
- b)  $\mathbb{R}^n$  oder  $\mathbb{C}^n$  mit dem Skalarprodukt

$$(u, v)_A = (u, A\bar{v})$$

mit einer hermiteschen, positiv definiten Matrix  $A$ .

- c)  $C[a, b]$  mit dem Skalarprodukt

$$(f, g) = \int_a^b f\bar{g} dx$$

ist ein Prä-Hilbertraum, aber kein Hilbertraum wie das folgende Beispiel zeigt.

Sei  $[a, b] = [0, 1]$  und  $f_n(x) = 0$ ,  $x \in [\frac{1}{n}, 1]$ ,  $f_n(x) = 1 - nx$ ,  $x \in [0, \frac{1}{n}]$ . Dann gilt:  $f_n \rightarrow f$ , wobei  $f(x) = 0$   $x \in (0, 1]$ ,  $f(0) = 1$ . Aus dem Satz über dominierte Konvergenz folgt  $f_n \rightarrow f$  in  $L^2(0, 1)$ . Also ist  $f_n$  eine Cauchyfolge deren Limes keine stetige Funktion ist.

d)  $l^2$  ist ein Hilbertraum. Hierbei ist

$$l^2 = \{ (c_1, c_2, \dots) \mid c_j \in \mathbb{R} \text{ oder } \mathbb{C}, \sum_{j=1}^{\infty} |c_j|^2 < \infty \}$$

$$(b, c)_{l^2} = (b, c) = \sum_{j=1}^{\infty} b_j \bar{c}_j$$

e)  $L^2(\Omega)$ ,  $\Omega$  offene Teilmenge des  $\mathbb{R}^n$ , ist ein Hilbertraum. Das Skalarprodukt ist erklärt durch

$$(u, v)_{L^2} = (u, v) = \int_{\Omega} u \bar{v} \, dx.$$

Die Vollständigkeit wurde auf dem Übungsblatt 3 Aufgabe 2 bewiesen.

f) Der Sobolevraum  $H^{1,2}(\Omega)$ ,  $\Omega$  offene Teilmenge des  $\mathbb{R}^n$  (siehe nachfolgende Definitionen).

Wir erläutern im Folgenden *drei* Definitionen des Sobolevraumes  $H^{1,2}(\Omega)$ . Im Folgenden sei  $\Omega \subseteq \mathbb{R}^n$  eine offene Menge und alle Funktionen seien reellwertig.

**1.10 Definition.** Sei  $\tilde{H}^{1,2}$  die Menge aller Funktionen aus  $C^1(\Omega)$  mit  $\int_{\Omega} (|u|^2 + |\nabla u|^2) \, dx < \infty$ . Bezüglich des (sogenannten  $H^{1,2}$ -Skalarproduktes)

$$(u, v)_1 = \int_{\Omega} u v \, dx + \int_{\Omega} \nabla u \cdot \nabla v \, dx = \int_{\Omega} u v + \int_{\Omega} \sum_{i=1}^n D_i u D_i v \, dx$$

ist  $\tilde{H}^{1,2}$  ein Prä-Hilbertraum.

**1.11 Definition (Vervollständigung).**  $H^{1,2}(\Omega)$  ist der Restklassenraum der Cauchyfolgen  $(u_j)_{j \in \mathbb{N}} \in \tilde{H}^{1,2}$  nach den Nullfolgen (alles bezüglich der  $H^{1,2}$ -Norm).

- Man kann sich  $H^{1,2}$  als die Menge aller Funktionen  $u$  vorstellen, für die eine Cauchyfolge  $(u_j)$  in der  $H^{1,2}$ -Norm existiert, so dass  $u_j \rightarrow u$  in  $L^2(\Omega)$ . Die Grösse  $L^2\text{-}\lim_{j \rightarrow \infty} \nabla u_j$  kann man sich als verallgemeinerte Ableitung von  $u$  vorstellen und schreibt  $\nabla u = \lim_{j \rightarrow \infty} \nabla u_j$ . Durch die Restklassenbildung sieht man Limites von Cauchyfolgen  $(u_j), (v_j)$  als gleich an, wenn  $u_j - v_j \rightarrow 0$  in  $\tilde{H}^{1,2}$ . Diese Art der Definition des Sobolevraumes  $H^{1,2}$  nennt man die „Definition durch Vervollständigung“.

Für die zweite Definition von  $H^{1,2}(\Omega)$  (für den Anfänger am empfehlenswertesten) benötigen wir zunächst den Begriff der verallgemeinerten Ableitung.

**1.12 Definition.** Sei  $u \in L^2(\Omega)$ . Die Funktion  $u_i \in L^2(\Omega)$  heisst (erste) **verallgemeinerte Ableitung von  $u$  bezüglich der  $i$ -ten Variablen**, wenn

$$(u, D_i \varphi)_{L^2} = -(u_i, \varphi)_{L^2} \quad \text{für alle } \varphi \in C_0^\infty(\Omega). \quad (1.13)$$



- Die Funktion  $u_i$  aus der Definition wird meist mit  $D_i u$  bezeichnet.
- Hierbei ist  $C_0^\infty(\Omega)$  der sogenannte **Raum der Testfunktionen**.
- Die Definition 1.12 ist motiviert durch die partielle Integration, falls  $u \in C^1(\Omega)$ , denn dann gilt

$$(u, D_i \varphi)_{L^2} = -(D_i u, \varphi)_{L^2}.$$

**1.14 Lemma (Eindeutigkeit der verallgemeinerten Ableitung).** *Es seien  $u_i, \tilde{u}_i \in L^2(\Omega)$  verallgemeinerte Ableitungen von  $u \in L^2(\Omega)$ . Dann gilt  $u_i = \tilde{u}_i$  fast überall.*

BEWEIS : Aus Definition 1.12 folgt, dass für alle  $\varphi \in C_0^\infty(\Omega)$

$$\int_{\Omega} (u_i - \tilde{u}_i) \varphi \, dx = 0. \quad (1.15)$$

Da nach Satz 4.9  $C_0^\infty(\Omega)$  dicht in  $L^2(\Omega)$  ist, existieren  $\varphi_j \in C_0^\infty(\Omega)$ , mit  $|\varphi_j| \leq 2$ , die fast überall gegen  $H(u_i - \tilde{u}_i)$  konvergieren. Hierbei ist die Heavyside Funktion  $H$  definiert durch

$$H(x) = \begin{cases} 1, & \text{falls } s > 0, \\ 0, & \text{falls } s = 0, \\ -1, & \text{falls } s < 0. \end{cases}$$

Einsetzen von  $\varphi_j$  in (1.15) und Grenzübergang  $j \rightarrow \infty$ , der aufgrund von Satz 1.3.12 möglich ist, liefern

$$\int_{\Omega} |u_i - \tilde{u}_i| \, dx = 0$$

Nun liefert Satz 1.3.15 die Behauptung. ■

**1.16 Definition (mit verallgemeinerten Ableitungen).**

$$H^{1,2}(\Omega) = \{u \in L^2(\Omega) \mid u \text{ besitzt erste verallgemeinerte Ableitungen in } L^2(\Omega)\}.$$

Das Skalarprodukt wird dann mit Hilfe der verallgemeinerten Ableitungen definiert durch

$$(u, v)_1 = (u, v)_{L^2} + (\nabla u, \nabla v)_{L^2}.$$

**1.17 Satz.**  $H^{1,2}(\Omega)$  ist vollständig.

BEWEIS : Sei  $(u_j)$  Cauchy-Folge in  $H^{1,2}(\Omega)$ . Dann ist  $(u_j)$  auch eine Cauchy-Folge in  $L^2(\Omega)$  mit Limes  $u \in L^2(\Omega)$ . Also gilt nach Lemma 1.7:  $(D_i \varphi, u_j) \xrightarrow{j} (D_i \varphi, u)$  für alle  $\varphi \in C_0^\infty(\Omega)$ . Wir zeigen, dass  $u$  verallgemeinerte Ableitungen in  $L^2(\Omega)$  hat. Da  $(D_i u_j)$  ebenfalls eine Cauchy-Folge in  $L^2(\Omega)$  ist, gilt  $D_i u_j \xrightarrow{j} \tilde{u}_i \in L^2(\Omega)$  in  $L^2(\Omega)$  und  $-(D_i \varphi, u_j) = (\varphi, D_i u_j) \xrightarrow{j} (\varphi, \tilde{u}_i)$  für alle  $\varphi \in C_0^\infty(\Omega)$ . Daher gilt  $\tilde{u}_i = D_i u$ . ■

**1.18 Definition (Tonnelli).** Es sei  $M$  die Menge aller Funktionen  $f \in L^2(\Omega)$  mit der Eigenschaft, dass die durch

$$\varphi_i(\xi) = f(x_1, x_2, \dots, x_{i-1}, \xi, x_{i+1}, \dots, x_n)$$

definierten Restriktionen  $\varphi_i$  bis auf eine Ausnahmemenge von Punkten  $(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$  mit dem  $(n-1)$ -dimensionalen Mass Null **absolut stetig**<sup>1</sup> ist. Die fast überall definierte Ableitung  $\frac{d}{d\xi}\varphi_i$  bezeichnen wir mit  $D_i f$ .  $D_i f$  ist eine messbare Funktion. Gilt überdies  $D_i f \in L^2(\Omega)$ ,  $i = 1, \dots, n$ , so schreiben wir  $f \in H^{1,2}(\Omega)$ . Das Skalarprodukt ist erwartungsgemäß definiert durch

$$(u, v)_{H^{1,2}} = (u, v)_{L^2} + (\nabla u, \nabla v)_{L^2}.$$

- Damit die Definitheitsvoraussetzung gilt, muss die übliche Klassenbildung durchgeführt werden, d.h. die Elemente von  $H^{1,2}(\Omega)$  sind Klassen von Funktionen, die sich auf einer Menge vom Mass Null unterscheiden.
- Ein **Gebiet**  $\Omega$ , d.h. eine offene, zusammenhängende Menge, deren Rand lokal durch Lipschitz-stetige Funktionen beschrieben werden kann, heisst **Gebiet mit Lipschitz-stetigem Rand** und man schreibt  $\partial\Omega \in C^{0,1}$ , wobei  $C^{0,1}$  die Menge der Lipschitz-stetigen Funktionen bezeichnet. Man kann zeigen, dass die drei Definitionen für  $H^{1,2}(\Omega)$  im Falle von Gebieten mit  $\partial\Omega \in C^{0,1}$  übereinstimmen (siehe Nečas, Les méthodes directes en théorie des équations elliptiques, S. 63 ff, Alt, S. 55 ff)

In der Theorie der partiellen Differentialgleichungen ist es von grosser Bedeutung, zum Ausdruck zu bringen, dass eine  $H^{1,2}$ -Funktion am Rand des Gebietes in einem verallgemeinerten Sinne verschwindet. Dies leistet der Raum

$$H_0^{1,2}(\Omega) = \overline{C_0^\infty(\Omega)}^{H^{1,2}} = \{ u \in H^{1,2} \mid \exists u_j \in C_0^\infty \text{ mit } u_j \rightarrow u \text{ in } H^{1,2} \}.$$

**1.19 Lemma.** Sei  $\Omega \subset \mathbb{R}^n$  offen und beschränkt,  $0 \in \Omega$ , und  $s < \frac{n}{2} - 1$ . Dann ist

$$|x|^{-s} \in H^{1,2}.$$

Für  $s \geq \frac{n}{2} - 1$  ist  $|x|^{-s} \notin H^{1,2}$ .

BEWEIS : Der Beweis des Lemmas ist vielleicht am einfachsten, wenn man die Definition der Sobolevräume nach Tonnelli durchführt. Die Funktion  $|x|^{-s}$  ist dann fast überall klassisch differenzierbar mit der Ableitung

$$\nabla |x|^{-s} = -s x |x|^{-s-2}$$

<sup>1</sup>Eine Funktion  $f : (a, b) \rightarrow \mathbb{R}$  heisst **absolut stetig**, falls für alle  $\varepsilon > 0$  ein  $\delta > 0$  existiert so, dass

$$\sum_{i=1}^N (\beta_i - \alpha_i) < \delta \quad \Rightarrow \quad \sum_{i=1}^N |f(\beta_i) - f(\alpha_i)| < \varepsilon,$$

für alle paarweise disjunkten Intervalle  $(\alpha_i, \beta_i) \subseteq (a, b)$ .

und mit Hilfe von Polarkoordinaten erhält man  $x_i |x|^{-s-2} \sim |x|^{-s-1} \in L^2(\Omega)$  genau dann, wenn  $s \leq \frac{n}{2} - 1$ . In der Tat gilt

$$\begin{aligned} \int_{B_r(0)} |x|^{-2(s+1)} dx &= \int_0^R \int_{S_r(0)} r^{-2(s+1)} dH_{n-1} dr \\ &= c_n \int_0^R r^{n-1-2(s+1)} dr < \infty \end{aligned}$$

genau für  $n - 1 - 2(s + 1) > -1$ , d.h.  $s < \frac{n}{2} - 1$ .

Um die im Lemma behauptete  $H^{1,2}$ -Inklusion im Fall der zweiten Definition nachzuweisen, arbeitet man am besten mit einer *Abschneidefunktion*  $\tau_R \in C^\infty(\mathbb{R}^n)$ ,  $\tau = 1$  in  $\mathbb{R}^n \setminus B_{2R}$ ,  $\tau_R = 0$  in  $B_R$ ,  $|\nabla \tau_R| \leq KR^{-1}$ . Es ist dann  $|x|^{-s} \tau_R \in C^\infty(\mathbb{R}^n)$  und aufgrund partieller Integration gilt für  $\varphi \in C_0^\infty(\Omega)$

$$\left( D_i(|x|^{-s} \tau_R), \varphi \right)_{L^2(\Omega)} = - \left( |x|^{-s} \tau_R, D_i \varphi \right)_{L^2(\Omega)} \rightarrow - \left( |x|^{-s}, D_i \varphi \right)_{L^2(\Omega)} \quad (R \rightarrow 0),$$

da  $\tau_R \rightarrow 1$ . Andererseits ist

$$\left( D_i(|x|^{-s} \tau_R), \varphi \right)_{L^2(\Omega)} = (-s x_i |x|^{-s-2} \tau_R, \varphi)_{L^2(\Omega)} + \left( |x|^{-s} D_i \tau_R, \varphi \right)_{L^2(\Omega)}.$$

Es gilt  $|x|^{-s} |D_i \tau_R| \leq KR^{-s} R^{-1}$  und somit  $\int_{B_{2R}} |x|^{-s} |D_i \tau_R| \leq c_n K R^{n/2+\delta} \rightarrow 0$  ( $R \rightarrow 0$ ), wobei  $\delta = \frac{n}{2} - 1 - s$ . Somit erhält man die Identität ( $R \rightarrow 0$ )

$$-s(x_i |x|^{-s-2}, \varphi)_{L^2(\Omega)} = \left( D_i(|x|^{-s}), \varphi \right)_{L^2(\Omega)} = - \left( |x|^{-s}, D_i \varphi \right)_{L^2(\Omega)},$$

d.h. die verallgemeinerte Ableitung  $D_i(|x|^{-s})$  stimmt fast überall mit der klassischen Ableitung  $-s x_i |x|^{-s-2}$  überein, und diese gehört zu  $L^2$  unter den obigen Voraussetzungen. ■

Für  $n = 3, 4$  ist also z. B.  $\frac{1}{|x|} \notin H^{1,2}$ , für  $n \geq 5$  gilt die Inklusion jedoch. Für  $n \geq 3$  ist  $|x|^{-1/4}$  eine unbeschränkte  $H^{1,2}$ -Funktion. Für  $n = 2$  liefert Lemma 1.19 keine unbeschränkte  $H^{1,2}$ -Funktion. Allerdings ist

$$\log \left| \log \frac{1}{|x|} \right|$$

eine unbeschränkte  $H^{1,2}$ -Funktion. Im Fall  $n = 1$  sind  $H^{1,2}$ -Funktionen Hölderstetig zum Exponenten  $\frac{1}{2}$ .

- Leider können  $H^{1,2}$ -Funktionen auf einer dichten Menge Singularitäten haben, z.B.

$$u(x) := \sum_{j=0}^{\infty} 2^{-j} |x - x^{(j)}|^{-s}, \quad s \text{ wie im Lemma}$$

$x^{(j)}$  durchläuft alle Punkte aus  $\Omega$  mit rationalen Koordinaten.

## 2.2 Der Projektionssatz

Ist  $H$  ein unendlich-dimensionaler Hilbertraum und  $V$  ein linearer Teilraum, so gibt es im Gegensatz zum endlich-dimensionalen Fall *nicht notwendig* eine Zerlegung der Gestalt

$$H = V \oplus V^\perp. \quad (2.1)$$

Hierbei ist das „**Orthogonalkomplement von  $V$** “ definiert durch

$$V^\perp := \{ x \in H \mid (x, v) = 0 \quad \forall v \in V \}.$$

Die Gleichung (2.1) bedeutet: Zu jedem  $u \in H$  gibt es eine eindeutige Zerlegung

$$u = v + w, \quad v \in V, \quad w \in V^\perp.$$

Ein Gegenbeispiel zu (2.1) ist der Hilbertraum  $L^2(I)$ ,  $I = (a, b)$  und der Unterraum  $V = C_0^\infty(I)$ . In der Tat: Sei  $u \in L^2(I)$ ,  $u \notin C_0^\infty(I)$ . Angenommen, es wäre  $u = v + w$  mit  $v \in C_0^\infty(I)$  und  $w \in C_0^\infty(I)^\perp$ , so folgt aus

$$\int_I w \varphi \, dx = 0 \quad \text{für alle } \varphi \in C_0^\infty(I)$$

$w = 0$ . (Diese Implikation wurde im Beweis von Lemma 1.14 bewiesen.) Somit ist  $u = v \in C_0^\infty(I)$ , was ein Widerspruch ist.

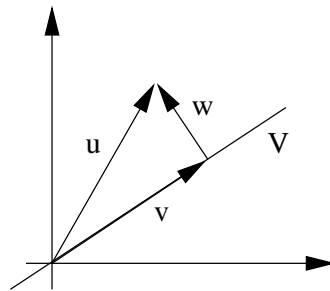
Die Hilbertraum-Theorie startet mit dem Satz, dass die Zerlegung (2.1) für *abgeschlossene* lineare Unterräume  $V \subset H$  funktioniert.

**2.2 Satz (Projektionssatz).** *Sei  $V$  ein abgeschlossener Teilraum des Hilbertraumes  $H$ . Dann gibt es zu jedem  $u \in H$  eine eindeutige Zerlegung*

$$u = v + w, \quad v \in V, \quad w \in V^\perp.$$

*Man schreibt:  $H = V \oplus V^\perp$ .*

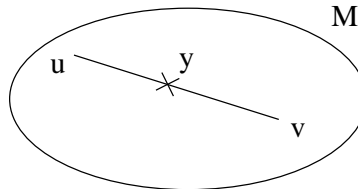
Der Beweis beruht auf der Tatsache, dass das Element  $v$  in Satz 2.2 dem Element  $u$  „am nächsten“ ist, d.h.  $\|u - v\| = \inf \{ \|u - y\| \mid y \in V \}$ .



Zur Konstruktion eines solchen Elementes benötigen wir einen entsprechenden Satz über die *Annahme des minimalen Abstandes*, den wir gleich allgemein für *konvexe* Mengen  $M$  formulieren.

**2.3 Definition.** Eine Menge  $M \subset H$  heisst **konvex**, wenn für alle  $u, v \in M$  und für alle  $\alpha \in [0, 1]$  gilt

$$y = \alpha u + (1 - \alpha)v \in M.$$



**2.4 Satz.** Es sei  $M \subset H$  eine nichtleere, abgeschlossene, konvexe Menge und  $u \in H$ . Dann gibt es ein eindeutiges Element  $v \in M$  mit

$$\|u - v\| = \inf \{ \|u - w\| \mid w \in M \} =: d.$$

BEWEIS : Wenn  $u \in M$ , gibt es nichts zu beweisen ( $\inf = 0$ ). Sei also  $u \notin M$ . Sei  $(w_j)$ ,  $w_j \in M$ , eine Minimalfolge, d.h.  $\|u - w_j\| \rightarrow d$  ( $j \rightarrow \infty$ ). Wir benutzen die Parallelogrammgleichung

$$\|a + b\|^2 + \|a - b\|^2 = 2\|a\|^2 + 2\|b\|^2$$

mit  $a = \frac{1}{2}u - \frac{1}{2}w_j$ ,  $b = \frac{1}{2}u - \frac{1}{2}w_k$ . Dann gilt

$$\|u - \frac{1}{2}(w_j + w_k)\|^2 + \|\frac{1}{2}(w_k - w_j)\|^2 = \frac{1}{2}\|u - w_j\|^2 + \frac{1}{2}\|u - w_k\|^2 \rightarrow d^2$$

für  $j, k \rightarrow \infty$ . Da  $M$  konvex ist und somit  $\frac{1}{2}(w_j + w_k) \in M$ , gilt  $\|u - \frac{1}{2}(w_j + w_k)\|^2 \geq d^2$ . Damit folgt

$$d^2 + \|\frac{1}{2}(w_k - w_j)\|^2 \leq \|u - \frac{1}{2}(w_j + w_k)\|^2 + \|\frac{1}{2}(w_k - w_j)\|^2 = d^2 + o(1),$$

und somit  $\|\frac{1}{2}(w_k - w_j)\| = o(1)$ , d.h.  $(w_j)$  ist eine Cauchyfolge und hat einen Limes  $v \in M$ . (Beachte, dass  $M$  abgeschlossen ist.) Da  $w_j \rightarrow v$ , folgt  $\|u - v\|^2 = d^2$ , und der Satz ist - bis auf die Eindeutigkeit - bewiesen. Das Element  $v \in M$  ist das den minimalen Abstand realisierende Element.

Die Eindeutigkeit folgt aus der folgenden Überlegung: Sei  $v, v' \in M$  und  $\|u - v\| = d$ ,  $\|u - v'\| = d$ . Dann gilt im reellen Fall

$$\begin{aligned} \|u - \frac{1}{2}(v + v')\|^2 &= \|u\|^2 - (u, v) - (u, v') + \frac{1}{4}\|v\|^2 + \frac{1}{2}(v, v') + \frac{1}{4}\|v'\|^2 \\ &= \frac{1}{2}\|u - v\|^2 + \frac{1}{2}\|u - v'\|^2 - \frac{1}{4}\|v\|^2 + \frac{1}{2}(v, v') - \frac{1}{4}\|v'\|^2 \\ &= d^2 - \frac{1}{4}\|v - v'\|^2 < d^2, \quad \text{falls } v \neq v'. \end{aligned}$$

(Im komplexen Fall ersetzt man in der ersten Zeile  $(u, v)$  und  $(u, v')$  durch  $\operatorname{Re}(u, v)$  bzw.  $\operatorname{Re}(u, v')$  sowie in der ersten und zweiten Zeile  $(v, v')$  durch  $\operatorname{Re}(v, v')$ .) Falls  $v \neq v'$ , kann somit  $\|u - v\|$  und  $\|u - v'\|$  nicht der minimale Abstand von  $u$  zu  $M$  sein. ■

BEWEIS von Satz 2.2: Aus Satz 2.4 mit  $M = V$  folgt die Existenz eines eindeutigen Elementes  $v \in V$  mit  $\|u - v\|^2 = \inf \{ \|u - w\|^2 \mid w \in V \} =: d^2$ . Es gilt

$$\|u - v\|^2 \leq \|u - v + t\varphi\|^2, \quad t \in \mathbb{R} \text{ bzw. } \mathbb{C}, \quad \varphi \in V.$$

Daraus folgt, zunächst im *reellen Fall*:

$$0 \leq 2(u - v, t\varphi) + t^2 \|\varphi\|^2.$$

Kürzen mit  $t > 0$  und Grenzübergang  $t \rightarrow 0$  ergibt  $0 \leq 2(u - v, \varphi)$  für alle  $\varphi \in V$ . Ersetzt man in dieser Ungleichung  $\varphi$  durch  $-\varphi$ , erhält man  $0 \geq 2(u - v, \varphi)$ , also insgesamt

$$0 = 2(u - v, \varphi) \quad \text{für alle } \varphi \in V$$

Man setze  $w = u - v$ . Dann gilt  $w \perp V$  und  $u = u - v + v = v + w$ ,  $v \in V$ ,  $w \perp V$ .

Im *komplexen Fall* erhält man

$$0 \leq 2 \operatorname{Re}(u - v, t\varphi) + |t|^2 \|\varphi\|^2 \tag{2.5}$$

und daraus wie eben

$$0 = \operatorname{Re}(u - v, \varphi)$$

Um auch  $\operatorname{Im}(u - v, \varphi) = 0$  zu erhalten, wählt man in (2.5)  $t = i\tau$  ( $i = \sqrt{-1}$ ,  $\tau \in \mathbb{R}$ ,  $\tau > 0$ ) und erhält

$$0 \leq 2 \operatorname{Re}(-i\tau(u - v, \varphi)) - \tau^2 \|\varphi\|^2$$

Kürzen mit  $\tau$ , und Grenzübergang  $\tau \rightarrow 0$  ergibt

$$0 \leq \operatorname{Re}(-i(u - v, \varphi)) = \operatorname{Im}(u - v, \varphi)$$

Durch Übergang zu  $-\varphi$  folgt wieder insgesamt

$$\operatorname{Im}(u - v, \varphi) = 0$$

Die Eindeutigkeit der Zerlegung ergibt sich folgendermassen: Sei  $u = v + w = v' + w'$ ,  $w, w' \perp V$ ,  $v, v' \in V$ . Daraus folgt  $v - v' = w' - w$ . Skalare Multiplikation mit  $v - v'$  ergibt

$$\|v - v'\|^2 = (w' - w, v - v') = 0.$$

■

## 2.3 Beschränkte lineare Funktionale und der Rieszsche Darstellungssatz

**3.1 Definition.** Ein *beschränktes lineares Funktional*  $F$  auf einem Hilbertraum  $H$  ist eine lineare Abbildung  $H \rightarrow \mathbb{R}$  bzw.  $\mathbb{C}$  mit der Eigenschaft, dass eine Konstante  $K$  existiert, so dass

$$|F(u)| \leq K\|u\| \quad \text{für alle } u \in H.$$

Die Menge der beschränkten linearen Funktionale wird auch mit  $H^*$  bezeichnet und heisst *Dualraum* des Hilbertraumes  $H$ .

**3.2 Satz.** Ein lineares Funktional auf einem Hilbertraum ist genau dann beschränkt, wenn es stetig ist.

BEWEIS :

(i) *Beschränktheit impliziert Stetigkeit.* Dies ist trivial: Gilt  $u_m \rightarrow u$ , so folgt

$$|F(u_m) - F(u)| = |F(u_m - u)| \leq K \|u_m - u\| \rightarrow 0.$$

(ii) *Stetigkeit impliziert Beschränktheit.*  $F$  sei stetig. Angenommen,  $F$  wäre nicht beschränkt. Dann gibt es eine Folge  $(u_m)$ ,  $u_m \neq 0$  mit  $\frac{|F(u_m)|}{\|u_m\|} := \alpha_m \rightarrow \infty$  ( $m \rightarrow \infty$ ). Setze

$$v_m := \frac{1}{\alpha_m} \frac{u_m}{\|u_m\|}$$

Dann gilt einerseits  $\|v_m\| \rightarrow 0$ , andererseits  $|F(v_m)| = \frac{1}{\alpha_m} \frac{|F(u_m)|}{\|u_m\|} = 1$ , d.h.  $F$  wäre nicht stetig. ■

**3.3 Definition.** Die **Norm** eines beschränkten linearen Funktionals  $F : H \rightarrow \mathbb{R}$  ist definiert durch

$$\|F\| = \inf \{ K; |F(u)| \leq K \|u\| \text{ für alle } u \in H \}.$$

• Äquivalent sind die Definitionen

$$\|F\| = \sup \left\{ \frac{|F(u)|}{\|u\|}; u \in H, u \neq 0 \right\} = \sup \{ |F(u)|; \|u\| \leq 1 \}.$$

(Nachweis: Übungsaufgabe Blatt 4)

**3.4 Lemma.** Die oben definierte Abbildung  $\|\cdot\| : H \rightarrow \mathbb{R}$  ist eine Norm auf  $H^*$ .

BEWEIS : Übungsaufgabe Blatt 4 ■

**3.5 Satz (Rieszscher Darstellungssatz).** Sei  $H$  ein Hilbertraum. Zu jedem  $F \in H^*$  gibt es eindeutig bestimmtes Element  $f \in H$  mit

$$F(u) = (u, f).$$

Man sagt daher:  $H$  lässt sich mit  $H^*$  identifizieren und schreibt - unpräziserweise -  $H = H^*$ .

BEWEIS : Sei  $N = \{x \in H \mid F(x) = 0\}$  der Kern von  $F$ .  $N$  ist abgeschlossener, linearer Teilraum von  $H$ . Ist nämlich  $u_m \in N$ ,  $u_m \rightarrow u$ , so folgt  $0 = F(u_m) \rightarrow F(u)$ , d.h.  $u \in N$ . O.B.d.A. dürfen wir  $N \neq H$  annehmen. (Andernfalls ist  $f = 0$ .) Es gibt daher ein  $w_0 \in H$ , das nicht in  $N$  ist. Nach dem Projektionssatz ist

$$w_0 = v + w \quad \text{mit } v \in N, w \in N^\perp, w \neq 0.$$

Wir beachten, dass  $F(w) \neq 0$ , da  $w \notin N$ , und  $F(u - \frac{F(u)}{F(w)}w) = 0$ , also  $u - \frac{F(u)}{F(w)}w \in N$ . Da  $w \in N^\perp$ , folgt

$$\left(u - \frac{F(u)}{F(w)}w, w\right) = 0$$

und

$$(u, w) = F(u) \frac{\|w\|^2}{F(w)},$$

somit

$$F(u) = \left(u, \frac{\overline{F(w)}}{\|w\|^2} w\right).$$

Damit ist das darstellende Element konstruiert. ■

Eine wichtige Anwendung des Riesz'schen Darstellungssatzes ist der Nachweis eines Existenzsatzes für elliptische Randwertprobleme der Gestalt

$$-\sum_{i,k=1}^n D_i(a_{ik}(x)D_k u) + c(x)u = f \text{ in } \Omega \quad (3.6)$$

mit Randbedingungen für  $u$  auf  $\partial\Omega$ . Hierbei ist  $\Omega$  ein Gebiet des  $\mathbb{R}^n$  mit  $\partial\Omega \in C^{0,1}$ ,  $f \in L^2(\Omega)$ ,  $a_{ik} \in L^\infty(\Omega)$ ,  $c \in L^\infty(\Omega)$ , und es gilt die Elliptizitätsbedingung

$$\sum_{i,k=1}^n a_{ik} \xi_k \xi_i \geq \lambda_0 |\xi|^2 \quad \text{für alle } \xi \in \mathbb{R}^n. \quad (3.7)$$

mit einer positiven Konstante  $\lambda_0$  („Elliptizitätskonstante“). Ferner gelte

$$a_{ik} = a_{ki} \quad (3.8)$$

und <sup>2</sup>

$$c(x) \geq \lambda_1 > 0 \quad (3.9)$$

Wir versehen nun den Raum  $H^{1,2}(\Omega)$  mit einem *neuen* Skalarprodukt, nämlich

$$(u, v)_A \equiv \sum_{i,k=1}^n \int_{\Omega} a_{ik} D_k u D_i v \, dx + \int_{\Omega} c u v \, dx$$

und beschränken uns auf den reellen Fall. Die Bedingungen (3.7) und (3.8) stellen sicher, dass  $(\cdot, \cdot)$  ein Skalarprodukt ist. Diese Vorgehensweise ist typisch für die Anwendungen der Funktionalanalysis - *man passt das Skalarprodukt dem Problem an.*

Aufgrund von (3.8) und (3.9) erkennt man, dass die Skalarprodukte  $(\cdot, \cdot)_{H^{1,2}}$  und  $(\cdot, \cdot)_A$  äquivalent sind, d.h. es gibt Konstanten  $K_1, K_2$  mit

$$K_1 \|u\|_{H^{1,2}} \leq \|u\|_A \leq K_2 \|u\|_{H^{1,2}}. \quad (3.10)$$

Äquivalente Skalarprodukte erzeugen den gleichen Konvergenzbegriff,  $H^{1,2}(\Omega)$  ist also auch Hilbertraum bzgl.  $(\cdot, \cdot)_A$  und wir können den Riesz'schen Darstellungssatz anwenden, indem wir das beschränkte lineare Funktional  $\varphi$ ,

$$\varphi(v) := (f, v)$$

durch ein Element  $u \in H^{1,2}(\Omega)$  darstellen:

$$(u, v)_A = (f, v) \quad \text{für alle } v \in H^{1,2}(\Omega)$$

---

<sup>2</sup>Für Nullrandbedingungen könnte man  $\lambda_1 = 0$  zulassen.



Ausgeschrieben bedeutet dies

$$\sum_{i,k=1}^n \int_{\Omega} a_{ik} D_k u D_i v \, dx + \int_{\Omega} c u v \, dx = \int_{\Omega} f v \, dx, \quad \text{für alle } v \in H^{1,2}(\Omega) \quad (3.11)$$

und man sagt, dass  $u$  eine **schwache Lösung** der Differentialgleichung (3.6) ist. Um (3.11) zu interpretieren nimmt man an, dass  $u$  hinreichend glatt ist, z.B.  $u \in C^2(\overline{\Omega})$ , und macht man folgende Überlegungen. Offensichtlich ist  $C_0^\infty(\Omega) \subset H^{1,2}(\Omega)$  und somit erhalten wir nach partieller Integration

$$- \sum_{i,k=1}^n \int_{\Omega} D_i (a_{ik} D_k u) v \, dx + \int_{\Omega} c u v \, dx = \int_{\Omega} f v \, dx, \quad \text{für alle } v \in C_0^\infty(\Omega),$$

was sofort (3.6) fast überall in  $\Omega$  liefert. Ausserdem folgt aus (3.11) noch die **natürliche Randbedingung** (man wähle  $v \in C^\infty(\overline{\Omega})$ , integriere partiell und benutze (3.6))

$$\sum_{i,k=1}^n a_{ik} D_k u \nu_i = 0 \quad \text{auf } \partial\Omega, \quad \nu = \text{äusserer Normalenvektor.} \quad (3.12)$$

Somit sehen wir, dass die schwache Lösung  $u$  von (3.11) eine Lösung des „verallgemeinerten Neumann Problems“ (3.6), (3.12) ist. Statt  $H^{1,2}(\Omega)$  lässt sich auch  $H_0^{1,2}(\Omega)$  als Grundraum nehmen, (3.12) entfällt dann und die schwache Lösung  $u \in H_0^{1,2}(\Omega)$  von (3.11) mit „ $v \in H^{1,2}(\Omega)$ “ ersetzt durch „ $v \in H_0^{1,2}(\Omega)$ “ ist eine Lösung des „verallgemeinerten Dirichlet Problems“ (3.6) mit  $u = 0$  auf  $\partial\Omega$ .

## 2.4 Beschränkte lineare Abbildungen

**4.1 Definition.** Seien  $H_1, H_2$  Hilberträume. Eine lineare Abbildung  $A : H_1 \rightarrow H_2$  heisst **beschränkt** genau dann, wenn eine Zahl  $K \in \mathbb{R}$  existiert mit

$$\|Au\| \leq K\|u\| \quad \text{für alle } u \in H_1. \quad (4.2)$$

• Für  $w \in H_1, z \in H_2$  wird das gleiche Symbol für die Norm verwendet:  $\|w\|, \|z\|$ . Wenn Verwechslungsgefahr besteht, kann man einen Index verwenden, z.B.  $\|w\|_1, \|z\|_2$  etc.

**4.3 Satz.** Eine Abbildung  $A : H_1 \rightarrow H_2$  ist genau dann beschränkt, wenn sie stetig ist.

BEWEIS :

„ $\Rightarrow$ “ Sei  $A$  beschränkt und  $u_m \rightarrow u$  in  $H_1$ , d.h.  $\|u_m - u\| \rightarrow 0$  ( $m \rightarrow \infty$ ). Dann folgt aus (4.2)

$$\|Au_m - Au\| \rightarrow 0,$$

d.h.  $Au_m \rightarrow Au$  in  $H_2$ .

„ $\Leftarrow$ “ Sei  $A$  stetig. Angenommen,  $A$  wäre nicht beschränkt. Dann existiert eine Folge  $(u_m)$ ,  $u_m \in H_1$ ,  $u_m \neq 0$  mit

$$\frac{\|Au_m\|}{\|u_m\|} =: \alpha_m \rightarrow \infty.$$

Wir setzen  $v_m := \alpha_m^{-1} \frac{u_m}{\|u_m\|}$ . Es gilt  $\|Av_m\| = 1$  und  $\|v_m\| \rightarrow 0$ . Dies widerspricht der Stetigkeit von  $A$ . ■

• Satz und Beweis gelten in *beliebigen normierten Räumen*.

### Beispiele beschränkter linearer Abbildungen in Hilberträumen:

a)  $H_1 = H_2 = L^2(\Omega)$ ,  $\Omega$  offene Teilmenge des  $\mathbb{R}^n$ . Sei  $A : H_1 \rightarrow H_2$  gegeben durch

$$(Au)(x) := g(x)u(x), \quad u \in L^2(\Omega),$$

mit einer festen Funktion  $g \in L^\infty(\Omega)$ .  $A$  ist beschränkt, da  $\|Au\|_{L^2} \leq \|g\|_{L^\infty} \|u\|_{L^2}$ .

b)  $H_1 = H^{1,2}(\Omega)$ ,  $H_2 = L^2(\Omega)$ .

$$(Au)(x) := D_1 u(x),$$

wobei  $D_1$  die verallgemeinerte partielle Ableitung in Richtung  $\mathbf{e}_1$  ist.  $A$  ist beschränkt, da

$$\|D_1 u\|_{L^2} \leq (\|u\|_{L^2}^2 + \|\nabla u\|_{L^2}^2)^{1/2} = \|u\|_{H^{1,2}}.$$

c) Sei  $H_1 = H_2 = L^2(\Omega)$  und  $K : \Omega \times \Omega \rightarrow \mathbb{R}$ ,  $K \in L^2(\Omega \times \Omega)$ , d.h.  $\int_{\Omega} \int_{\Omega} |K(x,y)|^2 dx dy =: C < \infty$ . Dann ist die durch

$$(Au)(x) := \int_{\Omega} K(x,y) u(y) dy \tag{4.4}$$

definierte Abbildung beschränkt. In der Tat haben wir

$$\begin{aligned} \|Au\|_{L^2}^2 &= \int_{\Omega} \left| \int_{\Omega} K(x,y) u(y) dy \right|^2 dx \\ &\leq \int_{\Omega} \int_{\Omega} |K(x,y)|^2 dy \int_{\Omega} |u(y)|^2 dy dx \\ &= \int_{\Omega} \int_{\Omega} |K(x,y)|^2 dy dx \int_{\Omega} |u(y)|^2 dy = C \|u\|^2. \end{aligned}$$

Die Abbildung in (4.4) heisst **Integraloperator mit Kern**  $K$ . Es gibt Kerne, die so singular sind, dass  $\int_{\Omega} \int_{\Omega} |K(x,y)|^2 dx dy = \infty$ , sie aber dennoch einen beschränkten Integraloperator erzeugen.

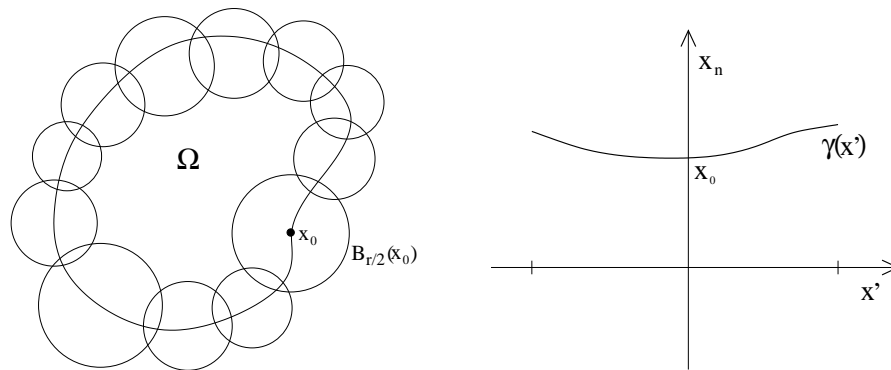
d) Sei  $\Omega$  ein Gebiet des  $\mathbb{R}^n$  mit  $\partial\Omega \in C^1$ ,  $H_1 = H^{1,2}(\Omega)$  und  $H_2 = L^2(\partial\Omega)$ . Jede Funktion  $u \in C^1(\bar{\Omega})$  besitzt eine Restriktion  $\tilde{R}u$  auf  $\partial\Omega$ , die definiert wird durch

$$(\tilde{R}u)(z) := u(z), \quad z \in \partial\Omega.$$

**4.5 Lemma.** *Es gibt eine Konstante  $K$ , die nur von  $\Omega$  abhängt so, dass für alle  $u \in C^1(\bar{\Omega})$  gilt:*

$$\|\tilde{R}u\|_{L^2(\partial\Omega)} \leq K \|u\|_{H^{1,2}(\Omega)} \quad \forall u \in C^1(\bar{\Omega}).$$

BEWEIS : Da  $\partial\Omega \in C^1$  kann man den Rand durch Bälle  $B_{r/2}(x_0) \subset \mathbb{R}^n$  überdecken, so dass er lokal durch  $C^1$ -Funktionen  $\gamma(x')$  beschrieben werden kann.



a) Sei  $x_0 \in \partial\Omega$  und sei der Rand nahe  $x_0$  flach, d.h. es gibt ein  $r > 0$  so, dass  $\partial\Omega \cap B_r(x_0) = \{(x', x_n); x_n = 0\}$ . Sei  $B_s^+ =: B_r(x_0) \cap \{(x', x_n); x_n \geq 0\}$ ,  $0 < s < r$  und sei  $\tau$  eine Abschneidefunktion, d.h.  $\tau \in C_0^\infty(B_r(x_0))$ ,  $\tau = 1$  in  $B_{r/2}(x_0)$ . Dann gilt:

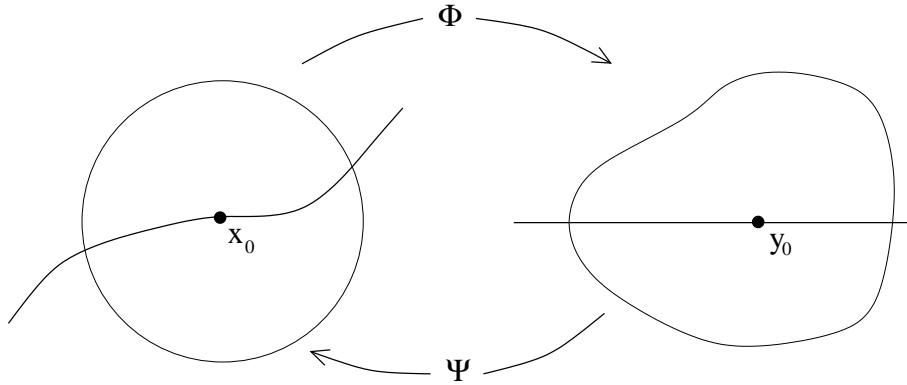
$$\begin{aligned} \int_{B_{r/2}(x_0) \cap \partial\Omega} |u|^2 dx' &\leq \int_{\{x_n=0\}} |u|^2 \tau dx' = - \int_{B_r^+(x_0)} D_{x_n}(|u|^2 \tau) dx \\ &= - \int_{B_r^+(x_0)} |u|^2 D_{x_n}(\tau) + 2|u| \operatorname{sign}(u) D_{x_n}(u) \tau dx \\ &\leq c \int_{B_r^+(x_0)} |u|^2 + |\nabla u|^2 dx, \end{aligned}$$

wobei wir in der letzten Zeile die Young-Ungleichung benutzt haben.

b) Sei nun der Rand nahe  $x_0$  nicht flach und durch  $\gamma(x')$ ,  $\gamma \in C^1$  beschrieben. Durch die Transformation  $x \mapsto y = \Phi(x)$ , wobei

$$\begin{aligned} \Phi_i(x) &= y_i = x_i, \quad i = 1, \dots, n-1, \\ \Phi_n(x) &= y_n = x_n - \gamma(x'), \end{aligned}$$

hat das Bild  $\Phi(B_r(x_0) \cap \partial\Omega)$  einen flachen Rand nahe  $y_0$  bzgl. der  $y$ -Koordinaten.



Sei  $x = \Psi(y)$  die Umkehrabbildung. Wir haben  $\det(\nabla_x \Phi) = \det(\nabla_y \Psi) = 1$ .  
Somit gilt für  $\tilde{u}(y) = u(\Psi(y))$  unter Berücksichtigung von a) und  $\gamma \in C^1$

$$\begin{aligned}
 \int_{B_{r/2}(x_0) \cap \partial\Omega} |u|^2 dS &= \int_{\Phi(B_{r/2}(x_0) \cap \partial\Omega)} |\tilde{u}|^2 \sqrt{1 + |\nabla_{y'} \gamma(y')|^2} dy' \\
 &\leq c \int_{\Phi(B_r(x_0) \cap \Omega)} |\tilde{u}|^2 + |\nabla_y \tilde{u}|^2 dy \\
 &\leq c \int_{B_r(x_0) \cap \Omega} |u|^2 + |\nabla_x u|^2 |\nabla_y \Psi(\Phi(x))|^2 dx \\
 &\leq c \int_{B_r(x_0) \cap \Omega} |u|^2 + |\nabla_x u|^2 dx.
 \end{aligned}$$

c) Da  $\partial\Omega$  kompakt ist existieren endlich viele Punkte  $x_0^i \in \partial\Omega$ ,  $i = 1, \dots, N$  so, dass  $\partial\Omega = \cup_{i=1}^N \partial\Omega \cap B_{r^i/2}(x_0^i)$ , wobei für jedes  $i = 1, \dots, N$  der Rand nahe  $x_0^i$  durch eine Funktion  $\gamma^i$  (siehe b)) beschrieben ist. Eine wiederholte Anwendung von b) liefert

$$\begin{aligned}
 \int_{\partial\Omega} |u|^2 dS &\leq \sum_{i=1}^N \int_{B_{r^i/2}(x_0^i) \cap \partial\Omega} |u|^2 dx \\
 &\leq c \sum_{i=1}^N \int_{B_{r^i/2}(x_0^i) \cap \Omega} |u|^2 + |\nabla u|^2 dx \\
 &\leq c \int_{\Omega} |u|^2 + |\nabla u|^2 dx,
 \end{aligned}$$

was gerade die Behauptung von Lemma 4.5 ist. ■

<sup>3</sup>Wir schreiben  $x = (x_1, \dots, x_{n-1}, x_n) =: (x', x_n)$ .

$\tilde{R}$  lässt sich durch Abschliessung auf ganz  $H^{1,2}(\Omega)$  zu einer beschränkten linearen Abbildung

$$R : H^{1,2}(\Omega) \rightarrow L^2(\partial\Omega)$$

fortsetzen.  $R$  heisst **Restriktions- oder Spuroperator**. Die Abbildung  $R$  wird verwendet, um Elementen von  $H^{1,2}$ , die eigentlich Klassen von Funktionen sind, die bis auf eine Menge vom Maß Null erklärt sind, dennoch eine Restriktion auf die Menge  $\partial\Omega$ , welche Mass Null hat, zuzuordnen. An Stelle der Voraussetzung, dass  $\Omega$  einen Rand  $\partial\Omega \in C^1$  hat, reicht es, zu verlangen, dass  $\partial\Omega \in C^{0,1}$ , d.h.  $\Omega$  hat einen Lipschitz-stetigen Rand. Die Konstruktion von  $R$  geschieht folgendermaßen:

Zu  $u \in H^{1,2}(\Omega)$  existieren Funktionen  $u_j \in C^1(\bar{\Omega})$  mit  $u_j \rightarrow u$  in  $H^{1,2}(\Omega)$  ( $j \rightarrow \infty$ ). (Beweisbedürftig, da  $u_j \in C^1(\bar{\Omega})$  und nicht bloß  $u_j \in C^1(\Omega)$  verlangt wird - die Glattheit von  $\partial\Omega$  wird verwendet. Siehe Evans S. 127 ff.) Wegen Lemma 4.5 gilt  $\|\tilde{R}u_j - \tilde{R}u_k\|_{L^2(\partial\Omega)} \rightarrow 0$  ( $j, k \rightarrow \infty$ ), d.h.  $\tilde{R}u_k$  ist eine Cauchyfolge mit Limes  $v \in L^2(\partial\Omega)$ . Man definiert nun  $Ru = v$  und erhält

$$\|Ru\|_{L^2(\partial\Omega)} \leq K\|u\|_{H^{1,2}(\Omega)} \quad \forall u \in H^{1,2}(\Omega).$$

Es folgen nun einige Lösbarkeitssätze für lineare Gleichungen  $Au = f$  mit beschränkten linearen Abbildungen  $A : H \rightarrow H$ ,  $H$  ein Hilbertraum.

**4.6 Satz.** Sei  $A$  eine beschränkte lineare Abbildung eines Hilbertraumes  $H$  in sich. Es gelte die **Koerzitivitätsbedingung**

$$\operatorname{Re}(Au, u) \geq c_0\|u\|^2 \quad \text{für alle } u \in H, \quad (4.7)$$

mit einer Konstanten  $c_0 > 0$ . Dann ist die Gleichung  $Au = f$  für alle  $f \in H$  eindeutig lösbar und die inverse Abbildung  $A^{-1}$  ist beschränkt.

BEWEIS : Zunächst bemerken wir, dass der Bildbereich  $A(H)$  abgeschlossen ist. Aus (4.7) folgt  $\|Au\| \geq c_0\|u\|$ . Ist nun  $Au_j \rightarrow z$ , so folgt  $c_0\|u_j - u_k\| \leq \|Au_j - Au_k\| \rightarrow 0$  ( $j, k \rightarrow \infty$ ), d.h.  $u_j$  ist eine Cauchyfolge mit Limes  $u$ . Aus der Stetigkeit von  $A$  und der Eindeutigkeit der Grenzwerte folgt  $z = Au$ . Wäre nun  $A(H)$  ein echter Teilraum von  $H$ , so gäbe es ein Element  $w \neq 0$ ,  $w \perp A(H)$ . Aus (4.7) folgt der Widerspruch

$$0 = \operatorname{Re}(Aw, w) \geq c_0\|w\|^2 > 0,$$

d.h.  $A(H) = H$ . Die Eindeutigkeit und die Stetigkeit der Inversen folgen ebenfalls aus der Ungleichung  $\|Au - Au'\| \geq c_0\|u - u'\|$ . ■

**Beispiel:**  $H = L^2(\Omega)$ ,  $Au := u - \int_{\Omega} K(x, y) u(y) dy$  unter der Voraussetzung

$$\int_{\Omega} \int_{\Omega} |K(x, y)|^2 dx dy \leq q < 1.$$

Mit Hilfe der Cauchy-Schwarzschen Ungleichung folgt

$$\left| \left( u, \int_{\Omega} K(\cdot, y) u(y) dy \right) \right| \leq \|u\|_{L^2}^2 \int_{\Omega} \int_{\Omega} |K|^2 dx dy.$$

Die Konstante  $c_0$  in (4.7) ist dann  $1 - q$ .

Eine der wichtigsten Anwendungen von Satz 4.6 liegt jedoch in der Theorie der elliptischen partiellen Differentialgleichungen. Hierzu benötigen wir den Begriff der beschränkten Bilinearform (bzw. Sesquilinearform) und das Lemma von Lax-Milgram.

**4.8 Definition.** Eine **beschränkte Bilinearform (Sesquilinearform)** ist eine Abbildung  $[\cdot, \cdot]$  von  $H \times H$  nach  $\mathbb{R}$  bzw.  $\mathbb{C}$  mit den Eigenschaften

- a)  $[u, v]$  ist linear in  $u$  und  $v$  (bzw. linear in  $u$  und antilinear in  $v$ ).
- b) Es existiert eine Konstante  $K$  mit

$$|[u, v]| \leq K \|u\| \|v\| \quad \text{für alle } u, v \in H.$$

**4.9 Definition.** Eine Bilinearform (Sesquilinearform) heisst **koerziv**, wenn eine Konstante  $c_0 > 0$  existiert mit  $\mathcal{R}e[u, u] \geq c_0 \|u\|^2$  für alle  $u \in H$ . Man beachte, dass wir keine Symmetrie für  $[u, v]$  vorausgesetzt haben.

**4.10 Lemma (Lax-Milgram).** Sei  $H$  ein Hilbertraum und  $[\cdot, \cdot]$  eine koerzive, beschränkte Bilinearform auf  $H$ . Dann gibt es zu jedem beschränkten, linearen Funktional  $f \in H^*$  ein eindeutiges  $u \in H$  mit

$$[v, u] = f(v) \quad \text{für alle } v \in H. \quad (4.11)$$

BEWEIS : Für jedes  $u \in H$  definiert die Abbildung  $\varphi_u(v) = [v, u]$  ein beschränktes, lineares Funktional auf  $H$ . Nach dem Rieszschen Darstellungssatz gibt es daher ein eindeutiges Element  $A(u) \in H$  mit

$$[v, u] = (v, A(u)).$$

Die Abbildung  $A$  ist linear, denn es gilt  $(v, A(\lambda x + \beta y) - \lambda A(x) - \beta A(y)) = 0$  für alle  $v \in H$  und somit  $A(\lambda x + \beta y) - \lambda A(x) - \beta A(y) = 0$ . Wir schreiben daher  $A(u) = Au$ . Da  $[\cdot, \cdot]$  eine beschränkte Bilinearform (Sesquilinearform) ist, gilt  $\|Au\|^2 = (Au, Au) = [Au, u] \leq K \|Au\| \|u\|$ , woraus die Beschränktheit von  $A$  folgt. Schließlich ist  $\mathcal{R}e(Au, u) = \mathcal{R}e[u, u] \geq c_0 \|u\|^2$ , d.h.  $A$  ist koerziv. Das lineare Funktional  $f$  stellen wir nach dem Rieszschen Darstellungssatz durch ein Element  $f_0 \in H$  dar:

$$f(v) = (v, f_0).$$

Die Gleichung (4.11) übersetzt sich daher in die Form

$$(v, Au) = (v, f_0) \quad \text{für alle } v \in H,$$

oder, äquivalent,

$$Au = f_0.$$

Da wir gezeigt haben, dass  $A$  beschränkt und koerziv ist, folgt mit Satz 4.6 die eindeutige Lösbarkeit von (4.11). ■

**Beispiel:** Sei  $a_{ik} \in L^\infty(\Omega)$ ,  $i, k = 1, \dots, n$ , und es gelte die Elliptizitätsbedingung

$$\sum_{i,k=1}^n a_{ik}(\xi) \xi_i \xi_k \geq \lambda_0 |\xi|^2 \quad \text{für alle } \xi \in \mathbb{R}^n, \lambda_0 > 0. \quad (4.12)$$

Ferner sei  $c \in L^\infty(\Omega)$ ,  $c \geq \lambda_1 > 0$ . Sei  $H = H^{1,2}(\Omega)$  und die Bilinearform  $[\cdot, \cdot]$  auf  $H \times H$  definiert durch

$$[u, v] := \sum_{i,k=1}^n (a_{ik} D_k u, D_i v)_{L^2} + (cu, v)_{L^2}.$$

Man überlegt sich für  $f \in L^2(\Omega)$  leicht, dass  $[\cdot, \cdot]$  die Voraussetzungen des Lemmas von Lax-Milgram erfüllt. Ferner ist das durch  $f(v) = (f, v)_{L^2}$  definierte Funktional auch auf  $H^{1,2}(\Omega)$  beschränkt (es ist sogar als Funktional auf  $L^2(\Omega)$  beschränkt). Nach dem Lemma von Lax-Milgram gibt es daher eine eindeutig bestimmte Funktion  $u \in H^{1,2}(\Omega)$  mit

$$\sum_{i,k=1}^n (a_{ik} D_k u, D_i v)_{L^2} + (cu, v)_{L^2} = (f, v)_{L^2} \quad \forall v \in H^{1,2}(\Omega),$$

d.h. es gibt eine schwache Lösung der elliptischen partiellen Differentialgleichung

$$-\sum_{i,k=1}^n D_i (a_{ik}(x) D_k u) + c(x) u = f(x)$$

mit natürlichen Randbedingungen. Eine andere Randbedingung, nämlich Nullrandbedingungen, erhält man, wenn man als Grundraum  $H_0^{1,2}(\Omega)$  wählt.

**4.13 Definition.** Die **Norm** einer beschränkten linearen Abbildung  $A : H_1 \rightarrow H_2$ ,  $H_1, H_2$  Hilberträume, ist die Grösse

$$\|A\| = \sup \{ \|Au\| \mid \|u\| = 1 \}$$

oder äquivalent

$$\|A\| = \inf \{ K \in \mathbb{R} \mid \|Au\| \leq K\|u\| \quad \text{für alle } u \in H_1 \}.$$

Man beweist leicht

$$\|A + B\| \leq \|A\| + \|B\|, \quad \|\lambda A\| = |\lambda| \|A\|$$

sowie für  $A : H_1 \rightarrow H_2$ ,  $C : H_2 \rightarrow H_3$

$$\|CA\| \leq \|C\| \|A\|.$$

**4.14 Satz (Neumann Reihe).** Sei  $H$  ein Hilbertraum und  $A : H \rightarrow H$  linear und beschränkt mit

$$\|A\| \leq q < 1. \quad (4.15)$$

Dann besitzt die Abbildung  $I - A$  eine auf ganz  $H$  definierte, beschränkte Inverse.

BEWEIS : Wir wollen für beliebige  $f \in H$  die Gleichung  $(I - A)u = f$  lösen. Mit der Bedingung (4.15) zeigt man leicht, dass die Reihe  $Sf = f + Af + A^2f + A^3f + \dots$  in  $H$  konvergiert. („Die Neumann Reihe  $S = \sum_{j=0}^{\infty} A^j$  konvergiert, wenn  $\|A\| < 1$ “) und ähnlich wie bei der Berechnung der geometrischen Reihe erhält man für alle  $f \in H$

$$(I - A)Sf = f,$$

und somit  $(I - A)^{-1} = \sum_{j=0}^{\infty} A^j$ , wobei  $A^0 := I$ . Die Eindeutigkeit und Stetigkeit der Inversen folgt aus

$$\|u - Au\| \geq \|u\| - \|Au\| \geq \|u\| (1 - \|A\|) \geq (1 - q) \|u\|.$$

■

## 2.5 Adjungierte Abbildungen

Es seien  $H_1, H_2$  Hilberträume und  $A : H_1 \rightarrow H_2$  eine beschränkte lineare Abbildung. Für jedes  $u \in H_2$  wird durch  $\varphi_u(v) := (Av, u)$  ein beschränktes lineares Funktional  $\varphi_u : H_1 \rightarrow \mathbb{R}$  bzw.  $\mathbb{C}$  definiert. Nach dem Rieszschen Darstellungssatz lässt sich  $\varphi_u$  durch ein Element, welches wir mit  $A^*u$  bezeichnen, darstellen:

$$\varphi_u(v) = (v, A^*u). \quad (5.1)$$

**5.2 Definition.** Sei  $A : H_1 \rightarrow H_2$  linear und beschränkt. Die vermöge (5.1) definierte Abbildung  $A^* : H_2 \rightarrow H_1$  heisst die zu  $A$  **adjungierte Abbildung**.

**5.3 Satz.**  $A^*$  ist linear und beschränkt.

BEWEIS : Wir haben für alle  $u, w \in H_2$ ,  $\alpha, \beta \in \mathbb{R}$  bzw.  $\mathbb{C}$

$$\begin{aligned} (v, A^*(\alpha u + \beta w)) &= (Av, \alpha u + \beta w) = \bar{\alpha}(Av, u) + \bar{\beta}(Av, w) \\ &= \bar{\alpha}(v, A^*u) + \bar{\beta}(v, A^*w) = (v, \alpha A^*u + \beta A^*w) \quad \text{für alle } v \in H_1 \end{aligned}$$

und also ist  $A^*(\alpha u + \beta w) = \alpha A^*u + \beta A^*w$ , d.h.  $A^*$  ist linear. Um die Beschränktheit zu beweisen benutzt man:

$$|(v, A^*u)| = |(Av, u)| \leq \|u\| \|A\| \|v\|.$$

Da für alle Elemente eines Hilbertraumes gilt  $\|w\| = \sup_{\|z\|=1} |(z, w)|$ , erhalten wir

$$\|A^*u\| = \sup_{\|v\|=1} |(v, A^*u)| \leq \|A\| \|u\|,$$

d.h.  $A^*$  ist beschränkt. ■

**5.4 Satz.** Falls  $A : H_1 \rightarrow H_2$  beschränkt ist gilt  $A^{**} = A$ . Ferner ist  $\|A\| = \|A^*\|$ .



BEWEIS : Übungsaufgabe ■

Die adjungierte Abbildung ist von grosser Bedeutung für die lineare Funktionalanalysis. Ein Beispiel ist der folgende *Alternativsatz*, den man aus der Linearen Algebra endlich-dimensionaler Räume kennt.

**5.5 Satz.** *Es seien  $H_1, H_2$  Hilberträume und  $A : H_1 \rightarrow H_2$  eine lineare beschränkte Abbildung mit abgeschlossenem Bildbereich  $A(H_1)$ . Dann ist*

$$H_2 = A(H_1) \oplus N(A^*).$$

Hierbei ist  $N(A^*)$  der Kern von  $A^*$ .

**5.6 Folgerung (Fredholmsche Alternative).** *Unter der Voraussetzung von Satz 5.5 gilt: Die Gleichung  $Au = f$  ist genau dann lösbar, wenn  $f \perp N(A^*)$ .*

BEWEIS (Satz 5.5): Da  $A(H_1)$  abgeschlossen ist, gibt es nach dem Projektionssatz einen Teilraum  $V \subset H_2$  mit

$$H_2 = A(H_1) \oplus V,$$

d.h. jedes  $z \in H_2$  besitzt eine eindeutige Darstellung

$$z = w + v, \quad w \in A(H_1), \quad v \in V, \quad w \perp v.$$

Behauptung:  $N(A^*) = V$ .

„ $\subset$ “: Sei  $y \in N(A^*)$ . Dann ist  $(Au, y) = (u, A^*y) = 0$  für alle  $u \in H_1$ , d.h.  $y \perp A(H_1)$ . Aus der Zerlegung  $y = y_1 + y_2$ ,  $y_1 \in A(H_1)$ ,  $y_2 \in V$ ,  $y_1 \perp y_2$ , folgt  $y - y_2 = y_1$  und  $\|y - y_2\|^2 = (y_1, y - y_2) = 0$ , d.h.  $y = y_2 \in V$ . Damit folgt  $N(A^*) \subset V$ .

„ $\supset$ “: Sei umgekehrt  $y \in V$ . Dann ist  $y \perp A(H_1)$  und  $0 = (Av, y) = (v, A^*y)$  für alle  $v \in H_1$ . Daraus folgt  $A^*y = 0$  und  $y \in N(A^*)$ , also  $V \subset N(A^*)$ .

Damit ist  $N(A^*) = V$  und Satz 5.5 bewiesen. ■

BEWEIS (Folgerung 5.6):

„ $\Leftarrow$ “: Sei  $f \perp N(A^*)$ . Nach Satz 5.5 haben wir eine Zerlegung  $f = f_1 + f_2$ ,  $f_1 \in A(H_1)$ ,  $f_2 \in N(A^*)$ ,  $f_1 \perp f_2$ . Somit ist  $\|f - f_1\|^2 = (f_2, f - f_1) = 0$ , d.h.  $f = f_1 \in A(H_1)$ , d.h. es gibt ein  $u \in H_1$  mit  $Au = f$ .

„ $\Rightarrow$ “: Sei  $f = Au$  und  $v \in N(A^*)$ . Dann haben wir  $(f, v) = (Au, v) = (u, A^*v) = 0$ , d.h.  $f \perp N(A^*)$ . ■

**Beispiele von adjungierten Operatoren:**

a)  $H_1 = H_2 = L^2(\Omega)$ ,  $\Omega$  Gebiet im  $\mathbb{R}^n$ ,

$$(Au)(x) := \int_{\Omega} K(x, y) u(y) dy$$

mit einem Kern  $K \in L^2(\Omega \times \Omega)$ . Es gilt

$$(A^*v)(y) = \int_{\Omega} \overline{K(x, y)} v(x) dx.$$

- b)  $H_1 = H_2 = L^2(\Omega)$ ,  $(Au)(x) := g(x)u(x)$  mit  $g \in L^\infty(\Omega)$ .  $A^*v(x) = \overline{g(x)}v(x)$ .
- c)  $H_1 = H^{1,2}(I)$ ,  $I = [a, b] \subset \mathbb{R}$ ,  $H_2 = L^2(I)$ ,  $Au = u'$ . Nach Definition ist  $(u, A^*v)_{H_1} = (u', v)_{L^2}$  für alle  $u \in H_1$ , d.h.

$$(u', (A^*v)')_{L^2} + (u, A^*v)_{L^2} = (u', v).$$

Wir verschieben die Inhomogenität der rechten Seite durch die Substitution  $y(t) = A^*v(t) - \int_a^t v(s) ds$  und erhalten

$$(u', y')_{L^2} + (u, y + \int_a^t v(s) ds)_{L^2} = 0,$$

d.h.  $y$  ist schwache Lösung der Differentialgleichung

$$-y'' + y + \int_a^t v(s) ds = 0 \quad (5.7)$$

mit den natürlichen Randbedingungen  $y'(a) = y'(b) = 0$ . Diese Differentialgleichung lässt sich für gegebenes  $v$  in  $y$  lösen.  $A^*v(t) = y(t) + \int_a^t v(s) ds$  ist dann die gewünschte Darstellung der adjungierten Abbildung. Offensichtlich ist bei dieser Wahl der Grundräume *nicht*  $A^*v = -v'$ , sondern aufgrund der verschiedenen Skalarprodukte ergibt sich ein komplizierterer Ausdruck, der linear in  $v$  ist.

### 5.8 Fredholmsche Alternative bei gewöhnlichen Differentialgleichungen.

Sei  $H_1 = H_2 = H$ ,  $H = H^{1,2}(I)$  mit  $I = [a, b]$ . Sei  $A$  die vermöge des Rieszschen Darstellungssatzes definierte Abbildung  $A : H \rightarrow H$ , die definiert ist durch

$$(Au, v)_{H^{1,2}} = (u', v')_{L^2} + (cu, v)_{L^2} \quad \text{mit } c \in L^\infty(I).$$

Sei  $f \in L^2(I)$ . Die Gleichung <sup>4</sup>

$$(Au, v)_{H^{1,2}} = (f, v)_{L^2} = (f_0, v)_{H^{1,2}} \quad \text{für alle } v \in H^{1,2}(I)$$

besagt, dass  $u$  eine schwache Lösung der gewöhnlichen Differentialgleichung

$$-u'' + cu = f \quad \text{f.ü. in } I$$

ist mit den natürlichen Randbedingungen  $u'(a) = u'(b) = 0$ . Da der Bildraum von  $A$  abgeschlossen ist (nicht trivial, beweisbedürftig! siehe Abschnitt 2.8), ist nach der Fredholmschen Alternative die Gleichung  $Au = f_0$  genau dann lösbar, wenn

$$(f_0, v)_{H^{1,2}} = (f, v)_{L^2} = 0$$

für alle  $v$  mit

$$(u, A^*v)_{H^{1,2}} = (Au, v)_{H^{1,2}} = 0 \quad \text{für alle } u \in H^{1,2}(I).$$

---

<sup>4</sup> $f_0 \in H^{1,2}(I)$  ist das nach dem Rieszschen Darstellungssatz repräsentierende Element des Funktionals  $v \rightarrow (f, v)$ ,  $v \in H^{1,2}(I)$

Dies bedeutet aber, dass  $v$  eine schwache Lösung der Differentialgleichung

$$-v'' + c(x)v = 0 \quad \text{f.ü. in } I$$

mit den natürlichen Randbedingungen  $v'(a) = v'(b) = 0$  ist. Wir stellen fest: Die Gleichung

$$-u'' + cu = f \quad \text{in } [a, b], \quad u'(a) = u'(b) = 0$$

ist genau dann lösbar, wenn  $f$  orthogonal bezüglich des  $L^2$ -Skalarproduktes zu den Lösungen von

$$-v'' + cv = 0, \quad v'(a) = v'(b) = 0$$

ist. Völlig analog behandelt man allgemeinere Differentialgleichungen oder auch Probleme mit Nullrandbedingungen. Man arbeitet dann mit  $H_0^{1,2}$  anstelle von  $H^{1,2}$ .

**5.9 Adjungierte Abbildungen unbeschränkter Operatoren.** Fasst man, wie eben, die Differentiation  $Au = u'$  als beschränkte Abbildung von  $H^{1,2}$  nach  $L^2$  auf, so ergibt sich  $A^*$  als Lösung einer Differentialgleichung zweiter Ordnung, ist also relativ kompliziert. Eigentlich hätte man ja ganz gerne eine Theorie, in der  $A^*u = -u'$  ist. Dies erreicht man, indem man lineare, nicht notwendig beschränkte Abbildungen betrachtet, die nicht auf dem gesamten zugrundeliegenden Hilbertraum erklärt sind.

**Beispiel:** a)  $H = L^2(0, 2\pi)$ ,  $Au = u'$ ,  $D(A) = \text{Definitionsbereich von } A = C^\infty(0, 2\pi)$ .  $A$  ist ein unbeschränkter Operator, denn für  $u_n(x) = \sin nx \in C^\infty(0, \pi)$  gilt

$$\|u_n\|_{L^2}^2 = \pi, \quad \|u_n'\|_{L^2}^2 = n^2\pi.$$

b)  $H = L^2(\mathbb{R})$ ,  $Au = u'$ ,  $D(A) = C_0^\infty(\mathbb{R})$ .  $A$  ist ein unbeschränkter Operator, denn für

$$u_n(x) = n^{\frac{1}{2}}J(nx),$$

wobei  $J$  der Glättungskern aus Abschnitt 1.2.4 ist, gilt

$$\|u_n\|_{L^2}^2 = \|J\|_{L^2}^2 \leq c, \quad \|u_n'\|_{L^2}^2 = n^2 \|J'\|_{L^2}^2 \rightarrow \infty \quad (n \rightarrow \infty).$$

In den Anwendungen der Funktionalanalysis ist der Definitionsbereich  $D(A)$  zumeist nur dicht in  $H$ . Dann definiert man die adjungierte Abbildung wie folgt:

**5.10 Definition.** Sei  $A : D(A) \rightarrow H$  eine lineare Abbildung, wobei  $D(A)$  ein linearer dichter Teilraum von  $H$  ist. Dann hat die **adjungierte Abbildung**  $A^* : D(A^*) \rightarrow H$  den Definitionsbereich

$$D(A^*) := \{v \in H; \varphi_v(u) = (Au, v) \text{ ist ein beschränktes lineares Funktional auf } D(A)\}.$$

$A^*v$  ist dasjenige durch den Riesz'schen Darstellungssatz gegebene Element von  $H$ , für welches

$$(u, A^*v) = (Au, v) \quad \text{für alle } u \in D(A)$$

gilt.

Zur Definition von  $\varphi_v$  in der  $D(A^*)$  definierenden Gleichung muss beachtet werden: Zunächst ist  $\varphi_v$  nur auf  $D(A)$  erklärt. Aber da für  $v \in D(A^*)$ ,  $(u_j) \in D(A)$  gilt

$$|\varphi_v(u_j - u_k)| = (A(u_j - u_k), v) \leq K \|u_j - u_k\|, \quad (5.11)$$

lässt sich  $\varphi_v$  auf ganz  $H$  als beschränktes lineares Funktional fortsetzen. Ist  $u \in H$  und nicht in  $D(A)$ , so gibt es, da  $D(A)$  dicht in  $H$  ist, eine Folge  $(u_j)$  mit  $u_j \rightarrow u$  in  $H$  und  $(u_j)$  ist Cauchyfolge. Wegen (5.11) konvergiert dann  $\varphi_v(u_j)$  und der Limes wird als  $\varphi_v(u)$  definiert. Die Definition ist unabhängig von der Folge  $(u_j)$ , die gegen  $u$  geht. Auf das so *fortgesetzte lineare Funktional* wird der Rieszsche Darstellungssatz angewandt.

**Beispiel:**  $H = L^2(\mathbb{R})$ ,  $D(A) = C_0^\infty(\mathbb{R})$  (oder  $H^{1,2}(\mathbb{R})$ ),  $Au = u'$ . Sei  $v \in D(A^*)$ , d.h.  $v \in L^2(\mathbb{R})$  und

$$|(u', v)| \leq K_v \|u\|_{L^2} \quad \text{für alle } u \in D(A). \quad (5.12)$$

Aufgrund der Definition der verallgemeinerten Abbildung  $v'$  bedeutet (5.12)

$$|(u, v')| \leq K \|u\|_{L^2} \quad \forall u \in C_0^\infty(\mathbb{R}),$$

dies impliziert aber sofort  $v' \in L^2(\mathbb{R})$ , denn

$$\|v'\|_{L^2} = \sup_{\substack{\|u\|_{L^2} = 1 \\ u \in C_0^\infty(\mathbb{R})}} |(u, v')| \leq K,$$

also ist  $v \in H^{1,2}(\mathbb{R})$ . Umgekehrt für  $v \in H^{1,2}(\mathbb{R})$  gilt (5.12) trivialerweise. Daher gilt

$$D(A^*) = H^{1,2}(\mathbb{R}),$$

und wir erhalten

$$(u', v) = -(u, v'), \quad (5.13)$$

d.h.  $A^*v = -v'$ . (5.13) stimmt mit der üblichen Formel für partielle Integration überein, da die auftretenden Randterme wegfallen wegen  $u \in D(A)$ ,  $v \in D(A^*)$ .

**5.14 Definition.** Eine lineare Abbildung  $A : D(A) \subseteq H \rightarrow H$ , wobei  $D(A)$  ein linearer dichter Teilraum von  $H$  ist, heisst **selbstadjungiert**, wenn

- i)  $D(A) = D(A^*)$  und
- ii)  $Au = A^*u$  für alle  $u \in D(A)$ .

**Beispiel:**

- a) Die Abbildung  $A : D(A) \rightarrow L^2(\mathbb{R})$  definiert durch

$$Au = \sqrt{-1} u', \quad D(A) = C_0^\infty(\mathbb{R})$$

ist *nicht* selbstadjungiert, da  $D(A)$  zu klein ist. Wenn man jedoch  $D(A) = H^{1,2}(\mathbb{R})$  wählt, dann ist  $D(A) = D(A^*)$  (siehe vorheriges Beispiel) und es gilt:

$$(Au, v) = \int_{\mathbb{R}} \sqrt{-1} u' \bar{v} dt = - \int_{\mathbb{R}} \sqrt{-1} u \bar{v}' dt = \int_{\mathbb{R}} u \overline{\sqrt{-1} v'} dt = (u, Av).$$

Um zu zeigen, dass keine Randterme bei der partiellen Integration auftreten, beweist man obige Identität erst für  $u, v \in C_0^\infty(\mathbb{R})$  und benutzt dann dass  $C_0^\infty(\mathbb{R})$  in  $H^{1,2}(\mathbb{R})$  dicht ist.<sup>5</sup> Damit ist  $Au = \sqrt{-1}u'$  mit  $D(A) = H^{1,2}(\mathbb{R})$  selbstadjungiert.

Die Abbildung  $A$  mit  $D(A) = H^{1,2}(\mathbb{R})$  wird auch als *Abschließung* der Abbildung  $A$  mit  $D(A) = C_0^\infty(\mathbb{R})$  bezeichnet.

### b) Der Laplace-Operator

$$Au = -\Delta u \equiv -\sum_{i=1}^n D_i^2 u = -\sum_{i=1}^n \frac{\partial^2 u}{\partial x_i^2} \quad (5.15)$$

ist bei geeignetem  $D(A)$  als Operator  $A : D(A) \rightarrow L^2(\Omega)$  selbstadjungiert. Zum Beweis benötigen wir einige Hilfsmittel.

**5.16 Satz (Poincaré-Ungleichung).** *Sei  $\Omega$  ein beschränktes Gebiet des  $\mathbb{R}^n$ . Dann gibt es eine Konstante  $K$ , die nur von  $n$  und dem Durchmesser von  $\Omega$  abhängt, so dass für alle  $u \in H_0^{1,2}(\Omega)$  gilt:*

$$\|u\|_{H^{1,2}(\Omega)}^2 \leq K \sum_{i=1}^n \|D_i u\|_{L^2(\Omega)}^2, \quad (5.17)$$

d.h.  $\|\nabla u\|_{L^2(\Omega)}$  ist eine äquivalente Norm auf  $H_0^{1,2}(\Omega)$ .

BEWEIS : Da  $C_0^\infty(\Omega)$ -Funktionen dicht in  $H_0^{1,2}(\Omega)$  und  $L^2(\Omega)$  sind reicht es (5.17) für solche Funktionen zu zeigen. Sei also  $\varphi \in C_0^\infty(\Omega)$ .

$$\begin{aligned} \|\varphi\|_{L^2}^2 &= \frac{1}{n} \sum_{i=1}^n \int_{\Omega} |\varphi|^2 1 \, dx = -\frac{1}{n} \sum_{i=1}^n \int_{\Omega} D_i(\varphi \varphi) x_i \, dx \\ &= -\frac{2}{n} \sum_{i=1}^n \int_{\Omega} \varphi D_i(\varphi) x_i \, dx. \end{aligned}$$

Sei  $\Omega \subseteq \{x; |x_i| \leq d, i = 1, \dots, n\}$ . Dann erhalten wir mit der Hölder-Ungleichung

$$\|\varphi\|_{L^2}^2 \leq \frac{2d}{n} \|\varphi\|_{L^2} \sum_{i=1}^n \|D_i \varphi\|_{L^2(\Omega)},$$

d.h.

$$\|\varphi\|_{L^2} \leq \frac{2d}{\sqrt{n}} \left( \sum_{i=1}^n \int_{\Omega} |D_i \varphi|^2 \, dx \right)^{\frac{1}{2}}. \quad (5.18)$$

Wir erhalten also für  $\varphi \in C_0^\infty(\Omega)$

$$\|\varphi\|_{H^{1,2}}^2 = \|\varphi\|_{L^2}^2 + \sum_{i=1}^n \|D_i \varphi\|_{L^2}^2 \leq \left( \frac{4d^2}{n} + 1 \right) \sum_{i=1}^n \|D_i \varphi\|_{L^2}^2$$

<sup>5</sup>Dies bedeutet, dass  $H_0^{1,2}(\mathbb{R}) = H^{1,2}(\mathbb{R})$ , was nur im Fall der Integration über die gesamte reelle Achse gilt.

Für beliebige  $u \in H_0^{1,2}(\Omega)$  nimmt man dann eine Folge  $u_n \in C_0^\infty(\Omega)$  mit  $u_n \rightarrow u$  in  $H_0^{1,2}(\Omega)$  und geht in der letzten Ungleichung zur Grenze  $n \rightarrow \infty$ . Dies beweist (5.17). ■

- Oft wird die Ungleichung (5.18), die die  $L^2$ -Norm einer Funktion nur durch die  $L^2$ -Norm des Gradienten abschätzt, als **Poincaré-Ungleichung** bezeichnet. Man kann auch noch folgende Version der Poincaré-Ungleichung beweisen. Allerdings benötigt man hierzu Hilfsmittel, die wir erst in Abschnitt 2.8 zur Verfügung haben.

**5.19 Satz.** Sei  $\Omega$  ein beschränktes Gebiet des  $\mathbb{R}^n$  mit  $\partial\Omega \in C^{0,1}$ . Dann existiert eine Konstante  $K$ , so dass für alle  $u \in H^{1,2}(\Omega)$  gilt:

$$\int_{\Omega} |u - u_{\Omega}|^2 dx \leq K \int_{\Omega} |\nabla u|^2 dx, \quad (5.20)$$

wobei

$$u_{\Omega} := \int_{\Omega} u dx := \frac{1}{|\Omega|} \int_{\Omega} u dx$$

der Mittelwert von  $u$  über  $\Omega$  ist.

- Insbesondere ist also  $\|\nabla u\|_{L^2}$  eine äquivalente Norm auf

$$\dot{H}^{1,2}(\Omega) \equiv \{u \in H^{1,2}(\Omega); u_{\Omega} = 0\}.$$

**5.21 Die Räume  $H^{k,2}(\Omega)$ .** Sei  $\Omega$  ein beschränktes Gebiet des  $\mathbb{R}^n$ . Sei  $k \in \mathbb{N}$  und  $u \in L^2(\Omega)$ . Man sagt  $u$  besitzt eine **verallgemeinerte Ableitungen  $k$ -ter Ordnung**, falls es Funktionen  $D^\alpha u \in L^2(\Omega)$ ,  $\alpha = (\alpha_1, \dots, \alpha_n)$ ,  $\alpha_i \in \mathbb{N}_0$ ,  $|\alpha| = k$  gibt, so dass für alle  $\varphi \in C_0^\infty(\Omega)$  gilt:

$$(u, D^\alpha \varphi) = (-1)^k (D^\alpha u, \varphi). \quad (5.22)$$

Man definiert dann  $H^{k,2}(\Omega)$  als

$$H^{k,2}(\Omega) := \{u \in L^2(\Omega); \quad u \text{ besitzt verallgemeinerte Ableitungen } D^\alpha u, |\alpha| \leq k\}$$

und definiert ein Skalarprodukt durch

$$(u, v)_{H^{k,2}(\Omega)} \equiv \sum_{|\alpha| \leq k} (D^\alpha u, D^\alpha v)_{L^2}. \quad (5.23)$$

Oftmals sind Überlegungen einfacher, wenn man mit periodischen Funktionen arbeitet. Sei  $Q = (-K, K)^n$  ein Würfel des  $\mathbb{R}^n$  und sei  $u : \mathbb{R}^n \rightarrow \mathbb{R}$  eine **periodische Funktion** mit Periode  $2K$ , d.h.  $u(x + 2K\mathbf{e}_i) = u(x)$ ,  $i = 1, \dots, n$ , wobei  $\mathbf{e}_i$  die Standardbasisvektoren sind. Wir bezeichnen mit  $H_{\text{per}}^{k,2}(Q)$  den Teilraum von  $H^{k,2}(Q)$  der periodischen Funktionen versehen mit dem  $H^{k,2}(Q)$ -Skalarprodukt. Hierbei wird die

Konvention benutzt, dass  $H^{0,2}(Q) \equiv L^2(Q)$  ist. Mit  $\dot{H}_{\text{per}}^{k,2}(Q)$  bezeichnen wir den Teilraum von  $H_{\text{per}}^{k,2}(Q)$  bestehend aus Funktionen mit Mittelwert Null. Aufgrund von Satz 5.19 ist  $\|\nabla u\|_{L^2}$  eine äquivalente Norm auf  $\dot{H}_{\text{per}}^{1,2}(Q)$ .

• Wir betrachten nun folgendes Problem für den Laplace-Operator: Sei  $f \in L^2_{\text{per}}(Q)$  gegeben. Wir suchen  $u \in \dot{H}_{\text{per}}^{1,2}(Q) \cap H_{\text{per}}^{2,2}(Q)$  mit

$$-\Delta u = f \quad \text{fast überall in } Q. \quad (5.24)$$

**5.25 Lemma.** Für alle  $f \in L^2_{\text{per}}(Q)$  existiert genau eine Lösung  $u \in \dot{H}_{\text{per}}^{1,2}(Q) \cap H_{\text{per}}^{2,2}(Q)$  des Problems (5.24). Diese Lösung erfüllt die Abschätzung

$$\|u\|_{H^{2,2}} \leq K \|f\|_{L^2}. \quad (5.26)$$

BEWEIS : Wir setzen

$$[u, v] \equiv \sum_{i=1}^n \int_Q D_i u D_i v \, dx, \quad F(v) \equiv \int_Q f v \, dx. \quad (5.27)$$

Dann ist  $F$  ein beschränktes lineares Funktional auf  $\dot{H}_{\text{per}}^{1,2}(Q)$  und es gilt

$$[u, u] \geq \sum_{i=1}^n \|D_i u\|_{L^2}^2 \geq c \|u\|_{H^{1,2}}^2$$

aufgrund der Bemerkungen vor dem Lemma über die Äquivalenz von Normen. Das Lemma von Max-Milgram (Lemma 4.10) liefert die Existenz einer eindeutigen schwachen Lösung  $u \in \dot{H}_{\text{per}}^{1,2}(Q)$  des Problems

$$[u, v] = F(v) \quad \forall v \in \dot{H}_{\text{per}}^{1,2}(Q). \quad (5.28)$$

Um zu zeigen, dass die Lösung  $u$  auch in  $H_{\text{per}}^{2,2}(Q)$  liegt benutzt man die Methode der **Differenzenquotienten**. Für  $0 < h < \frac{K}{2}$ , wobei  $e_j$   $j = 1, \dots, n$  die Standardbasisvektoren des  $\mathbb{R}^n$  sind, definieren wir durch

$$T_h(x) = T_{j,h}(x) \equiv x + h e_j \quad (5.29)$$

einen *Translationsoperator*. Wir setzen

$$w(x) \equiv u(T_{j,h}(x)) - u(x),$$

welches wiederum eine Funktion aus  $\dot{H}_{\text{per}}^{1,2}(Q)$  ist. Somit können wir in (5.28) als Testfunktion  $v = h^{-2}w(x)$  wählen und erhalten

$$\begin{aligned} & \frac{1}{h^2} \int_Q \sum_{i=1}^n D_i(u(x)) D_i(u(T_h(x)) - u(x)) \, dx \\ &= \frac{1}{h^2} \int_Q f(x)(u(T_h(x)) - u(x)) \, dx. \end{aligned}$$

Wenn man  $h^{-2}(u(x) - u(T_{-h}(x)))$  als Testfunktion wählt und  $y = x - he_j$  substituiert erhält man

$$\begin{aligned} & \frac{1}{h^2} \int_Q \sum_{i=1}^n D_i(u(T_h(x))) D_i(u(T_h(x)) - u(x)) dx \\ &= \frac{1}{h^2} \int_Q f(T_h(x))(u(T_h(x)) - u(x)) dx, \end{aligned}$$

wobei man noch ausnutzt, das aufgrund der Periodizität die Integrale über  $Q$  bzw. über  $\{x + he_j; x \in Q\}$  gleich sind. Diese beiden Gleichungen voneinander abgezogen liefern

$$\begin{aligned} \sum_{i=1}^n \int_Q \left| D_i \frac{w(x)}{h} \right|^2 dx &= \frac{1}{h^2} \int_Q (f(T_h(x)) - f(x))(u(T_h(x)) - u(x)) dx \\ &= \frac{-1}{h^2} \int_Q f(x)(u(T_h(x)) - 2u(x) + u(T_{-h}(x))) dx. \end{aligned} \quad (5.30)$$

Mit der Bezeichnung  $\varphi(x) = h^{-1}(u(T_h(x)) - u(x))$  kann man den letzten Term schreiben als

$$-\frac{1}{h} \int_Q f(x)(\varphi(x) - \varphi(T_{-h}(x))) dx \leq \|f\|_{L^2} \|h^{-1}(\varphi - \varphi \circ T_{-h})\|_{L^2}.$$

Unter Benutzung von Lemma 5.31 a) und der Young-Ungleichung kann man die rechte Seite durch

$$c \|f\|_{L^2}^2 + \frac{1}{2} \left\| \nabla \frac{w}{h} \right\|_{L^2}^2$$

abschätzen. Dies zusammen mit (5.30) ergibt

$$\left\| \nabla \frac{w}{h} \right\|_{L^2}^2 \leq 2c \|f\|_{L^2}^2.$$

Lemma 5.31 b) und die Periodizität von  $u$  liefern nun<sup>6</sup>

$$\sum_{j=1}^n \|D_j \nabla u\|_{L^2}^2 \leq 2c \|f\|_{L^2}^2,$$

d.h.  $u \in H_{\text{per}}^{2,2}(Q)$ . Dies zusammen mit (5.27) für  $v \in C_{\text{per}}^\infty(\Omega)$  liefert

$$-\sum_{i=1}^n \int_Q D_i^2 u v dx = \int_Q f v dx,$$

<sup>6</sup>Eine Argumentationsmöglichkeit die Abschätzung auf ganz  $Q$  zu bekommen ist, die obige Rechnung auf den Würfel  $\tilde{Q} = (-3K, 3K)^n$  durchzuführen.



woraus man sofort (5.24) fast überall folgert. ■

Das folgende Lemma zeigt den Zusammenhang zwischen dem klassischen Differenzenquotienten und den verallgemeinerten Ableitungen auf. Für  $\Omega \subseteq \mathbb{R}^n$  mit  $\partial\Omega \in C^{0,1}$  bezeichnen wir

$$\Omega_h \equiv \{x \in \Omega; \text{dist}(x, \partial\Omega) > h\}.$$

**5.31 Lemma.** Sei  $\Omega \subseteq \mathbb{R}^n$  ein Gebiet mit  $\partial\Omega \in C^{0,1}$  und  $u \in L^2(\Omega)$ .

a) Für  $u \in H^{1,2}(\Omega)$  und  $|h| \leq h_0$  gilt:

$$\|h^{-1}(u \circ T_{j,h} - u)\|_{L^2(\Omega_h)} \leq \|D_j u\|_{L^2(\Omega)}, \quad j = 1, \dots, n.$$

b) Seien  $h^{-1}(u \circ T_{j,h} - u) \in L^2(\Omega_{h_0})$  für alle  $h_0 > 0$ ,  $0 < h < h_0$  und gelte

$$\|h^{-1}(u \circ T_{j,h} - u)\|_{L^2(\Omega_h)} \leq C_1, \quad j = 1, \dots, n$$

dann existiert die verallgemeinerte Ableitung  $D_j u$  und es gilt

$$\|D_j u\|_{L^2(\Omega)} \leq C_1.$$

BEWEIS : a) Da  $C^1(\bar{\Omega})$  Funktionen dicht in  $H^{1,2}(\Omega)$  sind reicht es die Ungleichung für solche Funktionen zu beweisen. Sei also  $u \in C^1(\bar{\Omega})$ , dann haben wir mit der Bezeichnung  $\Delta_j^h u(x) \equiv h^{-1}(u(T_{j,h}(x)) - u(x))$

$$\begin{aligned} \int_{\Omega_h} |\Delta_j^h u(x)|^2 dx &= \int_{\Omega_h} \left| \int_0^1 D_j(x + the_j) dt \right|^2 dx \\ &\leq \int_{\Omega_h} \int_0^1 |D_j(x + the_j)|^2 dt dx \\ &\leq \int_0^1 dt \int_{\Omega} |D_j u(y)|^2 dy, \end{aligned}$$

was die Behauptung ist.

b) Hierfür fehlen uns noch einige Hilfsmittel (siehe Abschnitt 2.7). ■

Man kann auch das Analogon zu Lemma 5.25 für das Dirichlet Problem für den Laplace Operator beweisen. Sei  $\Omega \subseteq \mathbb{R}^n$  ein Gebiet mit  $\partial\Omega \in C^{1,1}$ . Für  $f \in L^2(\Omega)$  suchen wir eine Lösung von

$$\begin{aligned} -\Delta u &= f && \text{in } \Omega, \\ u &= 0 && \text{auf } \partial\Omega. \end{aligned} \tag{5.32}$$

Man erhält folgende Aussage: Für alle  $f \in L^2(\Omega)$  existiert genau eine Lösung  $u \in H_0^{1,2}(\Omega) \cap H^{2,2}(\Omega)$  von (5.32) mit

$$\|u\|_{H^{2,2}(\Omega)} \leq c \|f\|_{L^2(\Omega)}. \tag{5.33}$$

Der Beweis der Existenz einer schwachen Lösung in  $H_0^{1,2}(\Omega)$  läuft völlig analog. Zum Beweis der  $H^{2,2}(\Omega)$  Regularität geht man auch analog vor. Allerdings muss man eine Testfunktion nehmen, die in  $H_0^{1,2}(\Omega)$  liegt, was im Allgemeinen für  $u(T_h(x)) - u(x)$  nicht gilt (die Funktion hat keine Nullrandwerte). Also muss man die Argumentation lokalisieren, d.h. sei  $\Omega' \subseteq\subseteq \Omega$  und  $\tau \in C_0^\infty(\Omega)$  mit  $\tau = 1$  in  $\Omega'$  und  $|\nabla\tau| \leq c \operatorname{dist}(\Omega', \mathbb{R}^n \setminus \Omega)$ . Man wählt dann  $v(x) = h^{-2}(u(T(x)) - u(x))\tau^2$  und erhält bei etwas mehr technischem Aufwand:

$$\sum_{i=1}^n \int_{\Omega'} |\Delta_j^h \nabla u|^2 dx \leq c(\tau) \|f\|_{L^2(\Omega)}^2$$

was sofort mit Lemma 5.31 b)  $\nabla^2 u \in L^2(\Omega')$  liefert. Um eine analoge Aussage nahe dem Rand  $\partial\Omega$  zu erhalten kann man mit lokalen Koordinaten, die dem Rand angepasst sind argumentieren (siehe Alt S. 398). Dies ist etwas aufwendiger, funktioniert aber im Prinzip ähnlich wie die Aussage im Innern. Eine Kombination beider Abschätzungen liefert dann (5.33).

**5.34 Satz.** *Der Laplace-Operator  $Au = -\Delta u$  mit  $D(A) = \dot{H}_{\text{per}}^{2,2}(Q)$  ist selbstadjungiert.*

BEWEIS : Für glatte Funktionen  $u, v$  gilt

$$-\int_Q \Delta uv dx = \sum_{i=1}^h \int_Q D_i u D_i v dx = -\int_Q u \Delta v dx,$$

wobei die Randterme aufgrund der Periodizität wegfallen. Somit gilt für Funktionen  $u, v \in C_{\text{per}}^\infty(Q)$

$$(Au, v) = (u, Av), \quad (5.35)$$

d.h.  $A = A^*$ . Da  $C_{\text{per}}^\infty$ -Funktionen dicht in  $\dot{H}_{\text{per}}^{2,2}(Q)$  sind und alle auftretenden Integrale endlich sind gilt (5.35) für alle  $u, v \in \dot{H}_{\text{per}}^{2,2}(Q)$ . Sei nun  $v \in \dot{H}_{\text{per}}^{2,2}(Q)$ , dann gilt aufgrund von (5.35) für alle  $u \in \dot{H}_{\text{per}}^{2,2}(Q)$

$$|(Au, v)| \leq \|u\|_{L^2} \|v\|_{H^{2,2}},$$

d.h. nach der Definition von  $A^*$  ist  $v \in D(A^*)$  und es gilt  $(Au, v) = (u, Av)$ . Dies zusammen mit (5.35) liefert  $A^*v = Av$ , d.h.  $A = A^*$  auf  $\dot{H}_{\text{per}}^{2,2}(Q)$  und  $\dot{H}_{\text{per}}^{2,2}(Q) = D(A) \subseteq D(A^*)$ . Es bleibt zu zeigen, dass  $D(A^*) = D(A)$ . Sei  $v \in D(A^*)$ . Dann existiert nach Definition ein  $A^*v = f \in L_{\text{per}}^2(Q)$  mit

$$(Au, v) = (u, f) \quad \forall u \in D(A).$$

Nach Satz 5.25 existiert für  $f \in L_{\text{per}}^2(Q)$  eine eindeutige Lösung  $\tilde{v} \in \dot{H}_{\text{per}}^{2,2}(Q)$  der Gleichung  $A\tilde{v} = f$ . Wir wollen zeigen, dass  $v = \tilde{v}$ . Sei  $g \in L_{\text{per}}^2(Q)$  beliebig. Aufgrund von Satz 5.25 existiert ein  $\tilde{u} \in \dot{H}_{\text{per}}^{2,2}(Q)$  mit  $A\tilde{u} = g$ . Wir haben

$$\begin{aligned} (g, v - \tilde{v}) &= (A\tilde{u}, v) - (A\tilde{u}, \tilde{v}) = (\tilde{u}, f) - (\tilde{u}, A\tilde{v}) \\ &= (\tilde{u}, f) - (\tilde{u}, f) = 0, \end{aligned}$$

wobei wir (5.35) und  $A\tilde{v} = f$  benutzt haben. Somit ist  $v = \tilde{v}$ ,  $v \in D(A)$  und  $f = Av$ , d.h.  $D(A^*) \subseteq D(A)$ . ■

Im Falle von Nullrandbedingungen geht man analog vor.

**5.36 Satz.** Sei  $\Omega$  ein Gebiet des  $\mathbb{R}^n$  mit  $\partial\Omega \in C^{1,1}$ . Der Laplace-Operator  $Au = -\Delta u$  mit  $D(A) = H_0^{1,2}(\Omega) \cap H^{2,2}(\Omega)$  ist selbstadjungiert.

BEWEIS : Der Beweis läuft völlig analog zum Beweis von Satz 5.34. Man muss nur beim Beweis von (5.35) beachten, dass die Randterme wegfallen, da  $u, v \in H_0^{1,2}(\Omega)$ . ■

## 2.6 Separable Hilberträume und Orthogonalsysteme

Wir erinnern:

**6.1 Definition.** Eine Teilmenge  $M$  eines Hilbertraumes  $H$  heißt **dicht** in  $H$ , wenn es zu jedem  $u \in H$  eine Folge  $(u_j)$ ,  $u_j \in M$  gibt mit  $u_j \rightarrow u$ .

**6.2 Definition.** Ein Hilbertraum  $H$  heißt **separabel**, wenn es eine abzählbare, dichte Teilmenge in  $H$  gibt.

Die meisten Hilberträume, die aus den Anwendungen kommen, sind separabel, z.B.  $L^2(\Omega)$ ,  $H^{1,2}(\Omega)$ .

**6.3 Definition.** Ein **Orthogonalsystem** ist eine Menge  $\{\varphi_j \in H \mid j \in \mathbb{N}, \varphi_j \neq 0\}$  mit

$$(\varphi_k, \varphi_j) = 0 \quad \forall j \neq k.$$

Das Orthogonalsystem heißt **Orthonormalsystem**, wenn zusätzlich  $\|\varphi_j\| = 1$ ,  $j \in \mathbb{N}$ . Ein Orthogonalsystem heißt **vollständig** in  $H$ , wenn  $\text{span}\{\varphi_j \mid j \in \mathbb{N}\}$  dicht in  $H$  ist. Hierbei ist  $\text{span}\{\varphi_j \mid j \in \mathbb{N}\}$  die Menge der endlichen Linearkombinationen von beliebigen Elementen aus dem System, d.h.

$$\text{span}\{\varphi_j \mid j \in \mathbb{N}\} \equiv \left\{ \sum_{j=1}^n \alpha_j \varphi_j \mid \alpha_j \in \mathbb{R} \text{ bzw. } \mathbb{C}, j = 1, \dots, n \right\}.$$

**6.4 Lemma.** Es sei  $M = \{v_j \in H \mid j \in \mathbb{N}\}$ . Dann gibt es Orthonormalsysteme  $\{\varphi_j \in H, j \in \mathbb{N} \mid \varphi_j \neq 0\}$  mit  $\text{span } M = \text{span}\{\varphi_j \mid j \in \mathbb{N}\}$ .

BEWEIS : Man führt einen Orthogonalisierungsprozess durch. O.B.d.A. dürfen wir annehmen, dass für jedes  $N$  die Elemente  $v_1, v_2, \dots, v_N$  linear unabhängig sind. Setze

$$\varphi_1 = \frac{v_1}{\|v_1\|}$$

Sei  $\varphi_j$  schon konstruiert. Man setze

$$\tilde{\varphi}_{j+1} = v_{j+1} - \sum_{l=1}^j (v_{j+1}, \varphi_l) \varphi_l, \quad (6.5)$$

$$\varphi_{j+1} = \frac{\tilde{\varphi}_{j+1}}{\|\tilde{\varphi}_{j+1}\|}. \quad (6.6)$$

Offensichtlich ist  $\varphi_{j+1} \perp \varphi_l$ ,  $1 \leq l \leq j$ . Damit sind die  $\varphi_j$  durch vollständige Induktion definiert. Durch die Konstruktion erreicht man

$$\text{span}\{\varphi_j \mid j \in \mathbb{N}\} = \text{span}\{M\}.$$

■

Aus den Definitionen und Lemma 6.4 folgt offensichtlich

**6.7 Satz.** *Jeder separable Hilbertraum  $H$  besitzt ein vollständiges Orthonormalsystem.*

**6.8 Lemma.** *Sei  $\{\varphi_j\}_{j \in \mathbb{N}}$  ein Orthonormalsystem in einem Hilbertraum  $H$ . Dann ist für jedes  $w \in H$ ,  $N \in \mathbb{N}$*

$$\left\| \sum_{j=1}^N (w, \varphi_j) \varphi_j \right\|^2 = \sum_{j=1}^N |(w, \varphi_j)|^2 \leq \|w\|^2.$$

Somit gilt auch  $\sum_{j=1}^{\infty} |(w, \varphi_j)|^2 \leq \|w\|^2 < \infty$ .

BEWEIS : Es gilt

$$\begin{aligned} 0 &\leq \left\| w - \sum_{j=1}^N (w, \varphi_j) \varphi_j \right\|^2 \\ &= \|w\|^2 + \sum_{j=1}^N |(w, \varphi_j)|^2 - \sum_{j=1}^N \left\{ (w, (w, \varphi_j) \varphi_j) + ((w, \varphi_j) \varphi_j, w) \right\} \\ &= \|w\|^2 + \sum_{j=1}^N |(w, \varphi_j)|^2 - \sum_{j=1}^N \left\{ \overline{(w, \varphi_j)} (w, \varphi_j) + (w, \varphi_j) \overline{(w, \varphi_j)} \right\} \\ &= \|w\|^2 - \sum_{j=1}^N |(w, \varphi_j)|^2. \end{aligned}$$

Hieraus folgt die Aussage des Lemmas. ■

**6.9 Satz.** *Sei  $\{\varphi_j\}$  ein vollständiges Orthonormalsystem in einem Hilbertraum  $H$ . Dann lässt sich jedes Element  $u \in H$  als konvergente, **verallgemeinerte Fourierreihe***

$$u = \sum_{j=1}^{\infty} (u, \varphi_j) \varphi_j \tag{6.10}$$

darstellen und es gilt

$$\|u\|^2 = \sum_{j=1}^{\infty} |(u, \varphi_j)|^2. \tag{6.11}$$

BEWEIS : Nach Definition 6.3 gibt es zu  $u \in H$  eine Folge  $u^N = \sum_{j=1}^N \mu_j^N \varphi_j \rightarrow u \in H$ . Es ist  $\mu_j^N = (u^N, \varphi_j)$ , wie man sich durch Multiplikation im Sinne des Skalarproduktes mit  $\varphi_j$  überlegt. Es gilt wegen Lemma 6.8

$$\left\| \sum_{j=1}^N (u, \varphi_j) \varphi_j - u^N \right\|^2 = \left\| \sum_{j=1}^N (u - u^N, \varphi_j) \varphi_j \right\|^2 \leq \|u - u^N\|^2.$$

Da  $u^N \rightarrow u$ , folgt somit

$$\left\| \sum_{j=1}^N (u, \varphi_j) \varphi_j - u \right\|^2 \rightarrow 0 \quad (N \rightarrow \infty), \quad (6.12)$$

was (6.10) beweist. Aus (6.12) folgt

$$\sum_{j=1}^N |(u, \varphi_j)|^2 - \sum_{j=1}^N (u, \varphi_j) (\varphi_j, u) - \sum_{j=1}^N (u, \varphi_j) \overline{(u, \varphi_j)} + \|u\|^2 \rightarrow 0,$$

d.h.

$$\|u\|^2 - \sum_{j=1}^N |(u, \varphi_j)|^2 \rightarrow 0 \quad (N \rightarrow \infty).$$

Damit ist (6.11) bewiesen. ■

Man kann sich überlegen, dass die Vollständigkeit eines Orthonormalsystems äquivalent zu Relation (6.11) ist.

**6.13 Lemma.** *Sei  $\{\varphi_j\}$  ein Orthonormalsystem in einem Hilbertraum  $H$ . Wenn für alle  $u \in H$  gilt:*

$$\|u\|^2 = \sum_{j=1}^{\infty} |(u, \varphi_j)|^2,$$

dann ist  $\{\varphi_j\}$  ein vollständiges Orthonormalsystem.

BEWEIS : Sei  $u \in H$ . Im Beweis von Lemma 6.8 haben wir gezeigt:

$$\left\| u - \sum_{j=1}^n (u, \varphi_j) \varphi_j \right\|^2 = \|u\|^2 - \sum_{j=1}^n |(u, \varphi_j)|^2.$$

Nach Voraussetzung konvergiert die rechte Seite gegen Null, d.h.

$$\lim_{n \rightarrow \infty} \sum_{j=1}^n (u, \varphi_j) \varphi_j = u$$

und somit ist  $\text{span}\{\varphi_j \mid j \in \mathbb{N}\}$  dicht in  $H$ . ■

Die **verallgemeinerten Fourierkoeffizienten**  $(u, \varphi_j)$  besitzen eine *Minimaleigenschaft*. Die Lösung des Minimumproblems „Suche  $\mu_1, \dots, \mu_N \in \mathbb{R}$  bzw.  $\mathbb{C}$ , so dass

$$\left\| u - \sum_{j=1}^N \mu_j \varphi_j \right\|^2$$

minimal ist“ lautet  $\mu_j = (u, \varphi_j)$ . In der Tat haben wir für  $\mu_j \in \mathbb{R}$  bzw.  $\mathbb{C}$ ,  $j = 1, \dots, N$

$$\begin{aligned} \left\| u - \sum_{j=1}^N \mu_j \varphi_j \right\|^2 &= \|u\|^2 + \sum_{j=1}^N |\mu_j|^2 - \sum_{j=1}^N \overline{\mu_j} (u, \varphi_j) - \sum_{j=1}^N \mu_j \overline{(u, \varphi_j)} \\ &\geq \|u\|^2 - \sum_{j=1}^N |(u, \varphi_j)|^2, \end{aligned}$$

da  $\bar{a}b + a\bar{b} \leq |a|^2 + |b|^2$ . Somit ist die rechte Seite eine untere Schranke, die unabhängig von  $\mu_j$  ist und für  $\mu_j = (u, \varphi_j)$  angenommen wird, d.h. für  $\mu_j = (u, \varphi_j)$  wird das Minimum von  $\|u - \sum_{j=1}^N \mu_j \varphi_j\|^2$  angenommen.

**Beispiel:** Das System

$$\{(2\pi)^{-\frac{1}{2}} e^{ikx}\}_{k \in \mathbb{Z}} \quad (6.14)$$

ist ein Orthonormalsystem in  $L^2(-\pi, \pi)$ . In der Tat haben wir

$$\int_{-\pi}^{\pi} e^{imx} dx = \frac{e^{im\pi} - e^{-im\pi}}{im} = 0 \quad m \in \mathbb{Z} \setminus \{0\}.$$

Aufgrund der Relation

$$\cos mx = \frac{e^{imx} + e^{-imx}}{2}, \quad \sin mx = \frac{e^{imx} - e^{-imx}}{2i} \quad (6.15)$$

ist auch das System

$$\{(2\pi)^{-\frac{1}{2}}, \pi^{-\frac{1}{2}} \cos nx, \pi^{-\frac{1}{2}} \sin nx\}_{k \in \mathbb{N}} \quad (6.16)$$

ein Orthonormalsystem in  $L^2(-\pi, \pi)$ .

**6.17 Satz.** Die Systeme (6.14), (6.16) sind vollständige Orthonormalsysteme im  $L^2_{\text{per}}(-\pi, \pi)$ .

**BEWEIS :** Aufgrund der Beziehung (6.15) genügt es dies für eines der Systeme zu zeigen. Wir nehmen das System (6.14). Wir wissen bereits, dass das System (6.14) ein Orthonormalsystem ist. Da  $C^\infty[-\pi, \pi]$  dicht in  $L^2_{\text{per}}(-\pi, \pi)$  ist, reicht es zu zeigen, dass sich glatte Funktionen durch endliche Linearkombinationen von Funktionen aus (6.14), die wir mit  $l_k := (2\pi)^{-1/2} e^{-ikx}$  bezeichnen, approximieren lassen. Sei also  $f \in C^1[-\pi, \pi]$  und

$$P_n(f) \equiv \sum_{|k| \leq n} (f, l_k) l_k.$$

Nach Lemma 6.8 gilt

$$\sum_{k \in \mathbb{Z}} |(f, l_k)|^2 \leq \|f\|_{L^2}^2 < \infty,$$

und also für  $m > n$

$$\|P_m(f) - P_n(f)\|_{L^2}^2 \leq \sum_{|k| > n} |(f, l_k)|^2 \rightarrow 0 \quad (n \rightarrow \infty).$$

Somit ist  $(P_n(f))$  eine Cauchyfolge in  $L^2_{\text{per}}(-\pi, \pi)$  und es existiert ein  $\tilde{f} \in L^2_{\text{per}}(-\pi, \pi)$  so, dass  $P_n(f) \rightarrow \tilde{f}$  in  $L^2_{\text{per}}(-\pi, \pi)$  und für eine Teilfolge  $P_{n_k}(f)(x) \rightarrow \tilde{f}(x)$  fast überall. Aufgrund des folgenden Lemmas erhalten wir  $f = \tilde{f}$ , was die Dichtheit von  $\text{span}\{l_k, k \in \mathbb{Z}\}$  in  $L^2_{\text{per}}(-\pi, \pi)$  beweist. ■

**6.18 Lemma.** Sei  $f \in L^2_{\text{per}}(-\pi, \pi)$  und existieren  $M, \alpha > 0$  so, dass

$$|h^{-1}(f(x+h) - f(x))| \leq M \quad \forall |h| < \alpha.$$

Dann konvergieren die  $n$ -ten Partialsummen  $P_n(f)$  fast überall gegen  $f$ .

- Die Voraussetzungen des Lemmas sind insbesondere für Lipschitz-stetige Funktionen erfüllt.

BEWEIS : Wir haben

$$\begin{aligned} P_n(f)(x) - f(x) &= \sum_{|k| \leq n} (f - f(x), l_k) l_k(x) \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} (f(y) - f(x)) \sum_{|k| \leq n} e^{-iky} dy e^{ikx} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} (f(x-z) - f(x)) \sum_{|k| \leq n} e^{ikz} dz \\ &= \frac{1}{\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (f(x-2y) - f(x)) \sum_{|k| \leq n} e^{ik2y} dy, \end{aligned}$$

wobei wir erst  $z = x - y$  und dann  $z = 2y$  substituiert haben. Mit Hilfe der Summenformel der geometrischen Reihe erhält man

$$\begin{aligned} \sum_{|k| \leq n} e^{ikx} &= e^{-inx} \sum_{k=0}^{2n} e^{ikx} = e^{-inx} \frac{1 - e^{ix(2n+1)}}{1 - e^{ix}} \\ &= \frac{e^{ix(n+\frac{1}{2})} - e^{-ix(n+\frac{1}{2})}}{e^{\frac{ix}{2}} - e^{-\frac{ix}{2}}} = \frac{\sin((n+\frac{1}{2})x)}{\sin \frac{x}{2}}, \end{aligned}$$

und somit

$$P_n(f)(x) - f(x) = \frac{1}{\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{f(x-2y) - f(y)}{2y} \frac{2y}{\sin y} \sin((2n+1)y) dy,$$

d.h.  $P_n(f)(x) - f(x)$  ist ein Fourierkoeffizient bzgl. des Systems (6.16) der Funktion

$$g(x) = \frac{1}{\sqrt{\pi}} \chi_{[-\frac{\pi}{2}, \frac{\pi}{2}]} \frac{f(x-2y) - f(y)}{2y} \frac{2y}{\sin y},$$

welche aufgrund der Voraussetzungen eine  $L^2$ -Funktion ist. Lemma 6.8 liefert also  $P_n(f)(x) - f(x) \rightarrow 0$ , ( $n \rightarrow \infty$ ), was die Behauptung ist. ■

Mit Hilfe eines vollständigen Orthonormalsystems in  $L^2_{\text{per}}(-\pi, \pi)$  kann man vollständige Orthonormalsysteme in  $L^2_{\text{per}}(Q)$ ,  $Q = (-L, L)^n \subseteq \mathbb{R}^n$  konstruieren. Das System

$$l_k(x) = (2L)^{-\frac{1}{2}} e^{i\frac{\pi}{L}kx}, \quad k \in \mathbb{Z} \quad (6.19)$$

ist ein vollständiges Orthonormalsystem in  $L^2_{\text{per}}(-L, L)$ .

**6.20 Folgerung.** *Das System*

$$(2L)^{-\frac{n}{2}} e^{i\frac{\pi}{L}kx}, \quad (6.21)$$

wobei  $k = (k_1, \dots, k_n)$ ,  $k_i \in \mathbb{Z}$ ,  $i = 1, \dots, n$ , ein Multiindex ist, und  $x = (x_1, \dots, x_n)$  ein Vektor im  $\mathbb{R}^n$ , ist ein vollständiges Orthonormalsystem in  $L^2_{\text{per}}(Q)$ , wobei  $Q = (-L, L)^n$ .

BEWEIS : Wir beweisen die Folgerung nur für  $n = 2$ , der allgemeine Fall folgt analog. Das System (6.21) kann als Produkt von Funktionen des Systems (6.19) geschrieben werden. Wir haben

$$(2L)^{-1} e^{i\frac{\pi}{L}(k_1x_1+k_2x_2)} = l_{k_1}(x_1) \cdot l_{k_2}(x_2). \quad (6.22)$$

Man rechnet leicht nach, dass diese Funktionen ein Orthonormalsystem bilden. Aufgrund von Satz 6.9 und Lemma 6.13 ist das Orthonormalsystem (6.22) vollständig genau dann, wenn für alle  $u \in L^2_{\text{per}}(Q)$ , mit  $Q = (-L, L)^2$  gilt:

$$\int_{-L}^L \int_{-L}^L |u(x, y)|^2 dx dy = \sum_{j, k \in \mathbb{Z}} \left( \int_{-L}^L \int_{-L}^L u(x, y) l_j(x) l_k(y) dx dy \right)^2. \quad (6.23)$$

Aus  $u \in L^2_{\text{per}}(Q)$  folgt mit dem Satz von Fubini  $u(x) \in L^2_{\text{per}}(-L, L)$  für fast alle  $x \in (-L, L)$ . Da  $\{l_k(y)\}$  ein vollständiges Orthonormalsystem in  $L^2_{\text{per}}(-L, L)$  ist gilt für fast alle  $x \in (-L, L)$

$$\begin{aligned} \int_{-L}^L |u(x, y)|^2 dy &= \sum_{j \in \mathbb{Z}} \left( \int_{-L}^L u(x, y) l_j(y) dy \right)^2 \\ &=: \sum_{j \in \mathbb{Z}} |g_j(x)|^2. \end{aligned} \quad (6.24)$$

Da die  $|g_j(x)|^2$  nichtnegative messbare Funktionen sind (Man wende den Satz von Fubini auf  $u(x, y) l_j(y) \in L^2_{\text{per}}(Q)$  an) kann man, nach Satz 3.11, (6.24) über  $(-L, L)$  integrieren. Somit ist

$$\begin{aligned} \int_{-L}^L \int_{-L}^L |u|^2 dx &= \sum_{j \in \mathbb{Z}} \int_{-L}^L |g_j(x)|^2 dx \\ &= \sum_{j, k \in \mathbb{Z}} \left( \int_{-L}^L g_j(x) l_k(x) dx \right), \end{aligned}$$



wobei die Vollständigkeitsrelation (6.11) auf jedes der  $g_j$  angewendet wurde. Somit ist (6.23) bewiesen und das System (6.22) vollständig. ■

**6.25 Satz.** *Jeder separable Hilbertraum ist isometrisch-isomorph zum Hilbertraum  $l^2$ .*

BEWEIS : Sei  $\{\varphi_j\}$  ein vollständiges Orthonormalsystem in  $H$ . Jedes  $u \in H$  besitzt eine Darstellung

$$u = \sum_{j=1}^{\infty} (u, \varphi_j) \varphi_j .$$

Dem Element  $u$  ordnen wir den Vektor

$$Tu = ((u, \varphi_1), (u, \varphi_2), (u, \varphi_3), \dots)$$

zu. Offensichtlich ist  $T$  eine umkehrbar eindeutige Abbildung von  $H$  auf  $l^2$ . Insbesondere ist

$$(u, v) = \sum_{j=1}^{\infty} (u, \varphi_j) \overline{(v, \varphi_j)} = (Tu, Tv)_{l^2} .$$

■

Satz 6.25 ist nützlich, da gewisse Sätze über Hilberträume in dem recht anschaulichen Raum  $l^2$  bewiesen werden können. Ein Beispiel werden wir in dem nächsten Abschnitt sehen.

## 2.7 Schwache Konvergenz und schwache Kompaktheit

Neben der Normkonvergenz gibt es in Hilberträumen einen weiteren, sehr wichtigen Konvergenzbegriff, nämlich die *schwache* Konvergenz.

**7.1 Definition.** *Eine Folge  $(u_j)_{j \in \mathbb{N}}$ ,  $u_j \in H$  in einem Hilbertraum  $H$  **konvergiert schwach** gegen ein Element  $u$ , in Zeichen*

$$u_j \rightharpoonup u \quad \text{oder} \quad u_j \rightharpoonup u \text{ schwach in } H \quad (j \rightarrow \infty)$$

genau dann, wenn für alle  $v \in H$

$$(u_j, v) \rightarrow (u, v) \quad (j \rightarrow \infty) .$$

Im Unterschied hierzu bezeichnet man die Konvergenz bezüglich der Norm als **starke Konvergenz**.

**7.2 Lemma.** *Es gilt:*

- a) Falls  $u_n \rightarrow u$  stark in  $H$ , dann konvergiert  $u_n \rightharpoonup u$  schwach in  $H$ .
- b) Aus  $u_n \rightharpoonup u$  schwach in  $H$  folgt

$$\|u\| \leq \liminf_{n \rightarrow \infty} \|u_n\| . \tag{7.3}$$

BEWEIS : a) Für  $u_j \rightarrow u$  folgt für alle  $v \in H$

$$|(u_j, v) - (u, v)| \leq \|u_j - u\| \|v\| \rightarrow 0 \quad (j \rightarrow \infty),$$

d.h.  $u_j \rightarrow u$  schwach in  $H$ .

b) Sei  $u_j$  eine schwach konvergente Folge mit Grenzwert  $u$ . Für alle  $v \in H$  gilt:

$$|(u_j, v)| \leq \|u_j\| \|v\| .$$

Aufgrund der schwachen Konvergenz von  $u_j$  erhalten wir

$$|(u, v)| = \liminf_{j \rightarrow \infty} |(u_j, v)| \leq \liminf_{j \rightarrow \infty} \|u_j\| \|v\| ,$$

was sofort die Behauptung (7.3) liefert, wenn man  $\|u\| = \sup_{\|v\|=1} |(u, v)|$  beachtet. ■

**7.4 Lemma.** Sei  $\{\varphi_j\}$  ein vollständiges Orthonormalsystem in  $H$ . Dann konvergiert

$$\varphi_j \rightarrow 0 \quad \text{schwach in } H .$$

BEWEIS : Sei  $v \in H$ . Es gilt

$$v = \sum_{j=1}^{\infty} (v, \varphi_j) \varphi_j .$$

Da nach Satz 6.9  $\sum_{j=1}^{\infty} |(v, \varphi_j)|^2$  konvergiert, gilt  $(v, \varphi_j) \rightarrow 0, (j \rightarrow \infty)$ . ■

Da  $\|\varphi_j\| = 1$ , kann  $(\varphi_j)$  nicht stark (d.h. bezüglich der Normkonvergenz) gegen Null konvergieren. Man sieht also, dass die schwache Konvergenz verschieden („schwächer“) von der starken Konvergenz ist.

Der folgende Satz ist ein Analogon des Satzes von Bolzano-Weierstraß.

**7.5 Satz (Schwache Kompaktheit beschränkter Mengen im Hilbertraum).**

Sei  $(u_m)_{m \in \mathbb{N}}, u_m \in H$  eine beschränkte Folge in einem Hilbertraum  $H$ . Dann gibt es eine schwach konvergente Teilfolge  $(u_m)_{m \in \Lambda}, \Lambda \subseteq \mathbb{N}$ .

BEWEIS : Wir führen das Problem auf den Fall des separablen Hilbertraumes zurück, indem wir die Abschließung  $H_0$  von  $\text{span}\{u_m \mid m \in \mathbb{N}\}$  betrachten.  $H_0$  ist separabel, da die Menge der endlichen Linearkombinationen von Elementen  $u_m$  mit rationalen Koeffizienten einerseits abzählbar, andererseits dicht in  $H_0$  ist.  $H_0$  ist isometrisch isomorph zu  $l^2$  (siehe Satz 6.13), und Satz 7.5 übersetzt sich in die Aussage:

„Sei  $c^l = (c_1^l, c_2^l, \dots) \in l^2, \sum_{j=1}^{\infty} |c_j^l|^2 \leq K, l \in \mathbb{N}$ . Dann gibt es eine Teilfolge  $l \in \Lambda, \Lambda \subseteq \mathbb{N}$  und ein Element  $c = (c_1, c_2, \dots) \in l^2$  mit

$$\sum_{j=1}^{\infty} c_j^l \bar{\mu}_j \rightarrow \sum_{j=1}^{\infty} c_j \bar{\mu}_j \quad (l \rightarrow \infty, l \in \Lambda)$$

für alle  $\mu = (\mu_1, \mu_2, \dots) \in l^2$ . “

Der Beweis geschieht folgendermassen: Nach dem Weierstraßschen Doppelfolgensatz gibt es eine Teilfolge  $l \in \Lambda, \Lambda \subseteq \mathbb{N}$ , so dass

$$c_j^l \rightarrow c_j \quad (l \rightarrow \infty, l \in \Lambda)$$

für alle  $j \in \mathbb{N}$ . (Man setzt das bekannte Diagonalverfahren ein.) Diese Folge  $l \in \Lambda$  ist bereits unsere gesuchte Teilfolge, denn es gilt zunächst

$$\sum_{j=1}^N |c_j|^2 \leq \lim_{l \rightarrow \infty} \sum_{j=1}^N |c_j^l|^2 \leq K,$$

und somit

$$\sum_{j=1}^N |c_j|^2 \leq K,$$

d.h.  $c \in l^2$ . Weiterhin gilt

$$\left| \sum_{j=1}^{\infty} (c_j - c_j^l) \bar{\mu}_j \right| \leq \underbrace{\left| \sum_{j=1}^N (c_j - c_j^l) \bar{\mu}_j \right|}_{=: A} + \underbrace{\left| \sum_{j=N+1}^{\infty} (c_j - c_j^l) \bar{\mu}_j \right|}_{=: B}.$$

Der Term  $A$  geht gegen Null für  $l \rightarrow \infty, l \in \Lambda, N$  fest, da  $c_j^l \rightarrow c_j, l \rightarrow \infty, l \in \Lambda$ , und  $B$  lässt sich abschätzen durch

$$\begin{aligned} B &\leq \left( \sum_{j=N+1}^{\infty} |c_j - c_j^l|^2 \right)^{1/2} \left( \sum_{j=N+1}^{\infty} |\bar{\mu}_j|^2 \right)^{1/2} \leq (\|c\|_{l^2} + \|c_j^l\|_{l^2}) \left( \sum_{j=N+1}^{\infty} |\bar{\mu}_j|^2 \right)^{1/2} \\ &\leq 2K \left( \sum_{j=N+1}^{\infty} |\bar{\mu}_j|^2 \right)^{1/2} \end{aligned}$$

Da  $\sum_{j=1}^{\infty} |\mu_j|^2$  konvergiert, lässt sich für  $N \geq N(\varepsilon), N$  fest, erreichen, dass  $B < \varepsilon$  ist. Damit erhält man

$$|(c, \mu) - (c^l, \mu)| < o(1) + \varepsilon \quad (l \rightarrow \infty, l \in \Lambda),$$

und die schwache Konvergenz  $c^l \rightarrow c, l \in \Lambda$ , ist bewiesen.

Damit haben wir für beliebige  $v \in H_0$  die Konvergenz  $(u_m, v) \rightarrow (u, v)$  ( $m \rightarrow \infty, m \in \Lambda$ ),  $u = \sum_{j=1}^{\infty} c_j \varphi_j$ , wobei  $(\varphi_j)$  ein Orthonormalsystem von  $H_0$  ist, und damit auch  $(u_m, w) \rightarrow (u, w), w = w_1 + w_2, w_1 \in H_0, w_2 \in H_0^\perp$ . ■

Nun können wir Lemma 5.31 b) beweisen.

**5.31 Lemma.** Sei  $\Omega \subseteq \mathbb{R}^n$  ein Gebiet mit  $\partial\Omega \in C^{0,1}$  und  $u \in L^2(\Omega)$ .

a) Für  $u \in H^{1,2}(\Omega)$  und  $|h| \leq h_0$  gilt:

$$\|h^{-1}(u \circ T_{j,h} - u)\|_{L^2(\Omega_h)} \leq \|D_j u\|_{L^2(\Omega)}, \quad j = 1, \dots, n.$$

b) Seien  $h^{-1}(u \circ T_{j,h} - u) \in L^2(\Omega_{h_0})$  für alle  $h_0 > 0$ ,  $0 < h < h_0$  und gelte

$$\|h^{-1}(u \circ T_{j,h} - u)\|_{L^2(\Omega_h)} \leq C_1, \quad j = 1, \dots, n$$

dann existiert die verallgemeinerte Ableitung  $D_j u$  und es gilt

$$\|D_j u\|_{L^2(\Omega)} \leq C_1.$$

BEWEIS : Für  $\varphi \in C_0^\infty(\Omega)$  und  $|h|$  klein genug haben wir

$$\begin{aligned} \int_{\Omega} \Delta_j^h u \varphi \, dx &= \int_{\Omega} h^{-1}(u(x + he_j) - u(x)) \varphi(x) \, dx \\ &= - \int_{\Omega} u(x) (-h)^{-1} (\varphi(x - he_j) - \varphi(x)) \, dx \quad (*) \\ &= - \int_{\Omega} u \Delta_j^{-h} \varphi \, dx. \end{aligned}$$

Mit dem Diagonalverfahren wählen wir eine Folge  $h_k \rightarrow 0$  so, dass  $\Delta_j^{h_k} u \rightharpoonup w$  schwach in  $L^2(\Omega')$  für alle  $\Omega' \subseteq \subseteq \Omega$ , was aufgrund der Voraussetzungen möglich ist. Für diese Folge  $h_k$  konvergiert die linke Seite in (\*) gegen  $\int_{\Omega} w \varphi \, dx$  und die rechte Seite gegen  $-\int_{\Omega} u D_j \varphi \, dx$ . Nach der Definition der verallgemeinerten Ableitung ist also  $w = D_j u$ . Die Abschätzung für  $D_j u$  folgt sofort aus

$$\begin{aligned} \|D_j u\|_{L^2(\Omega)} &= \sup_{\substack{\varphi \in C_0^\infty(\Omega) \\ \|\varphi\|_{L^2(\Omega)} \leq 1}} \left| \int_{\Omega} D_j u \varphi \, dx \right| = \sup_{\substack{\varphi \in C_0^\infty(\Omega) \\ \|\varphi\|_{L^2(\Omega)} \leq 1}} \liminf_{h_k \rightarrow 0} \left| \int_{\Omega} \Delta_j^{h_k} u \varphi \, dx \right| \\ &\leq \liminf_{h_k \rightarrow 0} \|\Delta_j^{h_k} u\|_{L^2(\Omega_{h_k})} \leq C_1, \end{aligned}$$

wobei wir die Cauchy-Schwarz Ungleichung und die Kompaktheit des Trägers von  $\varphi$  benutzt haben. ■

Es folgen weitere Betrachtungen zur schwachen Konvergenz sowie ihrem Verhältnis zur starken Konvergenz.

**7.6 Satz.** Sei  $H$  ein Hilbertraum und  $(u_m)$ ,  $u_m \in H$ , schwach konvergent. Dann ist  $(u_m)$  beschränkt.

BEWEIS : Dies ist eine Folge des Prinzips der gleichmäßigen Beschränktheit, ein grundlegender Satz, welcher besagt, dass in vollständigen Räumen eine „punktweise“ beschränkte Folge linearer, beschränkter Abbildungen gleichmäßig beschränkt ist. Dieser Satz wird im zweiten Teil der Vorlesung bewiesen. Im obigen Fall sind die linearen, beschränkten Abbildungen lineare Funktionale, nämlich  $\varphi_m(v) := (v, u_m)$ . Wegen der schwachen Konvergenz der  $u_m$  ist  $\varphi_m(v)$  beschränkt, d.h.  $|\varphi_m(v)| \leq K_v$  für alle  $m \in \mathbb{N}$ , also „punktweise“ beschränkt. Das Prinzip der gleichmäßigen Beschränktheit ergibt  $\sup_{\|v\|=1} |\varphi_m(v)| \leq K$ , woraus  $\|u_m\| \leq K$  folgt. ■

**7.7 Satz.** Sei  $A : H_1 \rightarrow H_2$  beschränkt und linear,  $H_1, H_2$  Hilberträume. Dann ist  $A$  stetig bezüglich der schwachen Konvergenz von  $H_1$  und  $H_2$ , d.h.  $u_m \rightharpoonup u$  schwach in  $H_1$  impliziert  $Au_m \rightharpoonup Au$  schwach in  $H_2$ .

BEWEIS : Sei  $u_m \rightharpoonup u$  schwach in  $H_1$ . Dann gilt  $(Au_m, v) = (u_m, A^*v) \rightarrow (u, A^*v) = (Au, v)$ . ■

Die Umkehrung von Satz 7.6 können wir im Augenblick nicht beweisen.

**7.8 Die Umgebungsbasis der schwachen Topologie.** Die Umgebungsbasis eines Elementes  $u \in H$  in der **starken Topologie** sind die  $\varepsilon$ -Umgebungen  $U_\varepsilon(u) = \{x \in H; \|u - x\| < \varepsilon\}$ ,  $\varepsilon > 0$ . In der **schwachen Topologie** sind es die Mengen

$$U_{\varepsilon, \varphi_1, \varphi_2, \dots, \varphi_m}(u) = \{x \in H \mid |(\varphi_j, u - x)| < \varepsilon, j = 1, \dots, m\}.$$

Hierbei sind alle  $\varepsilon > 0$  und alle möglichen Kombinationen  $\varphi_m \in H, m \in \mathbb{N}$  zu verwenden. Im separablen Fall könnte man sich auf abzählbar viele  $\varphi_m$  beschränken. Dann wäre das sogenannte erste Abzählbarkeitsaxiom erfüllt, d.h. jedes Element hat eine abzählbare Umgebungsbasis ( $\varepsilon$  lässt man die rationalen Werte durchlaufen). Wenn das erste Abzählbarkeitsaxiom nicht erfüllt ist, gibt es bei einigen topologischen Begriffen Probleme, z.B. muss man dann zwischen folgenkompakt und überdeckungskompakt unterscheiden.

Ein weiterer, grundlegender, für die Variationsrechnung und Optimierung sehr nützlicher Satz ist

**7.9 Satz (Banach-Saks).** Es sei  $H$  ein Hilbertraum und  $(u_m), u_m \in H$ , eine schwach konvergente Folge mit Grenzwert  $u \in H$ . Dann gibt es eine Teilfolge  $\Lambda \subseteq \mathbb{N}$ , so dass die arithmetischen Mittel  $\frac{1}{N} \sum_{j=1}^N u_{m_j}, N \rightarrow \infty$ , stark gegen  $u$  konvergieren.

BEWEIS : O.B.d.A. nehmen wir an, dass  $u = 0$  gilt (Übergang von  $u_m$  zu  $u_m - u$ ). Wegen der schwachen Konvergenz  $u_m \rightharpoonup 0$  wählen wir sukzessive Indizes  $m_j \in \mathbb{N}$  aus, so dass

$$|(u_{m_j}, u_{m_k})| \leq \frac{1}{j^2} \quad \text{für } j < k, \quad j \in \mathbb{N}.$$

In der Tat: Ist  $u_{m_1}$  bis  $u_{m_k}$  bereits konstruiert und damit fest, nutzen wir

$$(u_m, u_{m_l}) \rightarrow 0 \quad (m \rightarrow \infty, l = 1, \dots, k)$$

aus und erhalten so den Index  $m_{k+1}$ . Die Indexfolge  $(m_j)$  ist die gewünschte, denn es gilt

$$\begin{aligned} \left\| \frac{1}{N} \sum_{j=1}^N u_{m_j} \right\|^2 &= \frac{1}{N^2} \sum_{j=1}^N \|u_{m_j}\|^2 + \frac{2}{N^2} \sum_{k=1}^N \sum_{j=1}^{k-1} \operatorname{Re}(u_{m_j}, u_{m_k}) \\ &\leq \frac{N}{N^2} K + \frac{2}{N^2} \sum_{k=1}^N \sum_{j=1}^{k-1} \frac{1}{j^2} \leq \frac{K}{N} + \frac{2}{N^2} N \frac{\pi^2}{6} \rightarrow 0 \quad (N \rightarrow \infty). \end{aligned}$$

■

Wir geben zwei Anwendungen des Satzes von Banach-Saks:

**7.10 Satz.** Sei  $(u_m)$  eine schwach konvergente Folge vektorwertiger Funktionen aus  $L^2(\Omega; \mathbb{R}^N)$ ,  $\Omega \subseteq \mathbb{R}^n$  offen, und  $C$  eine abgeschlossene, konvexe Menge des  $\mathbb{R}^N$ , so dass  $u_m(x) \in C$  für alle  $x \in \Omega$  mit Ausnahme einer Menge vom Mass Null. Dann gilt für den schwachen Limes  $u$  der  $u_m$ , dass  $u(x) \in C$  für alle  $x \in \Omega$  mit Ausnahme einer Menge vom Mass Null.

BEWEIS : Ist  $u_m(x) \in C$ , so gilt auch für die arithmetischen Mittel

$$\frac{1}{M} \sum_{j=1}^M u_{m_j}(x) \in C.$$

Da es sich durch Auswahl von Teilfolgen erreichen lässt, dass  $\frac{1}{M} \sum_{j=1}^M u_{m_j} \rightarrow u$  in  $L^2(\Omega; \mathbb{R}^N)$  konvergiert ( $M \rightarrow \infty$ ), erhält man durch Auswahl einer weiteren Teilfolge, dass  $\frac{1}{M} \sum_{j=1}^M u_{m_j}$  fast überall konvergiert (Blatt 2; Aufgabe 4). Daraus folgt für den Limes  $u$ , dass  $u(x) \in C$  für alle  $x \in \Omega \setminus E$ ,  $\mu(E) = 0$  gilt. ■

Wenn  $C$  nicht konvex ist, ist die entsprechende Aussage im allgemeinen falsch. Allgemeiner zeigt man mit der gleichen Methode

**7.11 Satz.** Es sei  $C$  eine konvexe, abgeschlossene Menge in einem Hilbertraum  $H$  und  $u_m \in C$ ,  $u_m \rightharpoonup u$  schwach in  $H$  ( $m \rightarrow \infty$ ). Dann gilt  $u \in C$ .

Wir präsentieren noch eine Anwendung des Satzes von Banach-Saks in der Optimierung:

**7.12 Satz.** Es sei  $C$  eine nichtleere, abgeschlossene, konvexe Menge eines Hilbertraumes  $H$  und  $f : C \rightarrow \overline{\mathbb{R}}$  stetig in der starken Topologie von  $H$ . Ferner sei  $f$  konvex und koerziv auf  $C$ , d.h.  $f(u) \rightarrow \infty$  für  $\|u\| \rightarrow \infty$ ,  $u \in C$ . Dann existiert ein Minimum von  $f$  auf  $C$ .

BEWEIS : O.B.d.A. sei  $f$  nicht konstant gleich  $\infty$  auf  $C$ . Sei  $(u_m)$ ,  $u_m \in C$  eine Minimalfolge, d.h.

$$f(u_m) \rightarrow \inf_C f =: I < \infty, \quad |f(u_m) - I| \leq 2^{-m}.$$

Wegen der Koerzivitat von  $f$  ist  $\|u_m\| \leq K$  und es gibt eine Teilfolge, so dass, nach Ummummerierung,

$$u_m \rightharpoonup u \text{ schwach} \quad \text{und} \quad \frac{1}{N} \sum_{j=1}^N u_j \rightarrow u \text{ stark.}$$

Es ist  $\frac{1}{N} \sum_{j=1}^N u_j \in C$ , also auch  $u \in C$ . Ferner gilt  $\frac{1}{N} \sum_{j=1}^N f(u_j) \rightarrow \inf_C f$ ,  $N \rightarrow \infty$ .

In der Tat, sei  $\varepsilon > 0$ , dann gilt fur geeignetes  $N_0$  und  $N > N_0$

$$\begin{aligned} \left| \frac{1}{N} \sum_{j=1}^N f(u_j) - I \right| &\leq \left| \frac{1}{N} \sum_{j=1}^N (f(u_j) - I) \right| \\ &\leq \frac{1}{N} \sum_{j=1}^{N_0} |f(u_j) - I| + \frac{1}{N} \sum_{j=N_0}^{\infty} |f(u_j) - I| \\ &\leq \frac{N_0}{N} + \varepsilon \leq 2\varepsilon \quad \text{fur } N \text{ gro genug.} \end{aligned}$$

Somit erhalten wir für  $N$  groß genug:

$$f(u) \leq \varepsilon + f\left(\frac{1}{N} \sum_{j=1}^N u_j\right) \leq \varepsilon + \frac{1}{N} \sum_{j=1}^N f(u_j) \leq 2\varepsilon + \inf_C f.$$

Anschließend der Grenzübergang  $\varepsilon \rightarrow 0$  ergibt die Behauptung. ■

## 2.8 Kompakte lineare Abbildungen

**8.1 Definition.** Eine Menge  $M$  eines Hilbertraumes  $H$  heißt **relativ folgenkompakt** (bzw. **relativ schwach folgenkompakt**), wenn jede Folge  $(u_m)$ ,  $u_m \in M$  eine Teilfolge besitzt, welche in  $H$  stark (bzw. schwach) konvergiert.

Unpräziserweise sagt man auch, dass  $M$  relativ kompakt (bzw. relativ schwach kompakt) ist, meint aber relativ folgenkompakt (bzw. relativ schwach folgenkompakt).

**8.2 Definition.** Eine Menge  $M$  eines Hilbertraumes  $H$  heißt **folgenkompakt** (bzw. **schwach folgenkompakt**), wenn sie relativ folgenkompakt (bzw. relativ schwach folgenkompakt) ist und alle Häufungspunkte bzgl. der starken (bzw. schwachen) Topologie in  $M$  liegen.

Üblicherweise definiert man in topologischen Räumen diejenigen Mengen  $M$  als **kompakt**, welche die Eigenschaft haben, dass jede Überdeckung von  $M$  durch offene Mengen  $O_i$ ,  $i \in I$ , eine endliche Teilüberdeckung besitzt, d.h. es gibt eine *endliche* Indexmenge  $I_1 \subset I$  mit  $M \subset \bigcup_{i \in I_1} O_i$ . Im Unterschied zu folgenkompakt spricht man auch von überdeckungskompakt. Für die an Anwendungen in der Analysis orientierte Funktionalanalysis wird der Begriff der „Überdeckungskompaktheit“ als zu umständlich empfunden. Außerdem dient es bei Beweisen manchmal dem Vorstellungsvermögen, bei einer Folge an einen zeitlichen Ablauf zu denken.

Wie wir in Abschnitt 2.7 bereits diskutiert haben, sind beschränkte Mengen im Hilbertraum nicht notwendig relativ folgenkompakt - allerdings relativ schwach folgenkompakt.

Fundamental für die Anwendungen der Funktionalanalysis ist der Begriff des kompakten Operators (Abbildung).

**8.3 Definition.** Eine Abbildung  $A : H_1 \rightarrow H_2$  mit Hilberträumen  $H_1, H_2$  heißt **kompakt**, wenn  $A$  beschränkte Mengen in relativ folgenkompakte überführt. Die Abbildung  $A : H_1 \rightarrow H_2$  heißt **vollstetig**, wenn sie schwach konvergente Folgen in stark konvergente Folgen überführt.

Bei diesen Definitionen muss  $A$  nicht notwendig linear sein.

**8.4 Lemma.** Seien  $H_1, H_2$  Hilberträume. Die lineare Abbildung  $A : H_1 \rightarrow H_2$  ist genau dann kompakt, wenn sie vollstetig ist. Insbesondere ist  $A$  beschränkt.

BEWEIS :

„ $\Leftarrow$ “ Sei  $A$  vollstetig und  $M \subseteq H_1$  eine beschränkte Menge. Sei  $(u_m)$  eine Folge mit  $u_m \in M$ . Zu zeigen ist, dass  $(Au_m)$  eine konvergente Teilfolge hat. Da beschränkte Mengen schwach folgenkompakt sind, gibt es eine Teilfolge  $\Lambda \subset \mathbb{N}$ , so dass  $u_m \rightharpoonup u$  ( $m \in \Lambda$ ,  $m \rightarrow \infty$ ) schwach in  $H_1$ . Da  $A$  vollstetig ist, konvergiert  $Au_m$ ,  $m \in \Lambda$  stark in  $H_2$ .

„ $\Rightarrow$ “ Sei  $A$  kompakt und  $u_m \rightharpoonup u$  schwach in  $H_1$ . Wir zeigen, dass  $A$  beschränkt ist. Wäre  $A$  unbeschränkt, gäbe es eine Folge  $v_m$  mit  $\|v_m\| = 1$  und  $\|Av_m\| \rightarrow \infty$ . Dies widerspricht der Forderung, dass  $(Av_m)$  relativ kompakt ist.

Nach Satz 7.6 sind beschränkte lineare Abbildungen stetig von der schwachen Topologie in die schwache, d.h. aus  $u_m \rightharpoonup u$  schwach folgt  $Au_m \rightharpoonup Au$  schwach. Angenommen,  $Au_m$  konvergiere nicht stark gegen  $Au$ . Dann gibt es ein  $\varepsilon > 0$  und eine Teilfolge  $\Lambda$  mit  $\|Au_m - Au\| > \varepsilon$ ,  $m \in \Lambda$ . Da  $(Au_m)_{m \in \Lambda}$  relativ kompakt ist, gibt es eine Teilfolge  $\Lambda_1 \subset \Lambda$ , so dass  $Au_m \rightarrow f$  stark in  $H_2$ . Es folgt  $\|f - Au\| > \varepsilon$ . Andererseits gilt  $(Au_m, v) = (u_m, A^*v) \rightarrow (u, A^*v) = (Au, v)$  und  $(Au_m, v) \rightarrow (f, v)$ . Damit folgt  $Au = f$ . Es muss daher die Konvergenz  $Au_m \rightarrow Au$  stattfinden. ■

### Beispiele kompakter linearer Abbildungen:

Typische Beispiele sind Integraloperatoren: Sei  $\Omega \subseteq \mathbb{R}^n$  beschränkt und offen,  $K \in L^2(\Omega \times \Omega)$  und

$$Au(x) := \int_{\Omega} K(x, y) u(y) dy \quad (8.5)$$

**8.6 Lemma.** Die durch (8.5) definierte Abbildung  $A : L^2(\Omega) \rightarrow L^2(\Omega)$  ist vollstetig.

BEWEIS : Wir beweisen die Aussage zunächst für Kerne  $K \in C(\overline{\Omega} \times \overline{\Omega})$ . Es gilt für  $u_m \in L^2(\Omega)$ ,  $\|u_m\| \leq c$

$$\begin{aligned} |Au_m(x) - Au_m(z)| &= \left| \int_{\Omega} (K(x, y) - K(z, y)) u_m(y) dy \right| \\ &\leq \sup_{y \in \overline{\Omega}} |K(x, y) - K(z, y)| \int_{\Omega} |u_m(y)| dx \\ &\leq \sup_{y \in \overline{\Omega}} |K(x, y) - K(z, y)| |\Omega|^{1/2} \|u_m\|_{L^2(\Omega)}. \end{aligned}$$

Da  $K$  gleichmäßig stetig ist, gibt es für  $\varepsilon > 0$  ein  $\delta > 0$ , so dass

$$\sup_{y \in \overline{\Omega}} |K(x, y) - K(z, y)| < \varepsilon \quad \text{für } |x - z| < \delta.$$



Die Funktionen  $Au_m$  sind daher *gleichgradig stetig*. Aus

$$\begin{aligned} \sup_{x \in \overline{\Omega}} |Au_m(x)| &\leq \sup_{x \in \overline{\Omega}} \int_{\Omega} |K(x, y)| |u_m(y)| dy \\ &\leq \sup_{x, y \in \overline{\Omega}} |K(x, y)| |\Omega|^{\frac{1}{2}} \|u_m\|_{L^2(\Omega)} \end{aligned}$$

folgt, dass  $Au_m$  auch *gleichgradig beschränkt* ist. Nach dem Satz von Arzela-Ascoli gibt es daher eine gleichmäßig konvergente Teilfolge. Konvergenz in  $L^\infty(\Omega)$  impliziert Konvergenz in  $L^2(\Omega)$  bei beschränktem Gebiet, denn

$$\|Au_m - Au\|_{L^2(\Omega)} \leq |\Omega|^{\frac{1}{2}} \|Au_m - Au\|_{L^\infty(\Omega)} .$$

Die Abbildung  $A$  ist daher bei gleichmäßig stetigem Kern kompakt.

Im allgemeinen Fall  $K \in L^2(\Omega \times \Omega)$  verwenden wir ein Approximationsargument.

Da  $C_0^\infty(\Omega)$  dicht in  $L^2(\Omega)$  ist, gibt es eine Folge von Funktionen  $K_m \in C(\overline{\Omega} \times \overline{\Omega})$  mit

$$\int_{\Omega} \int_{\Omega} |K - K_m|^2 dx dy \rightarrow 0 \quad (m \rightarrow \infty) .$$

Wir setzen  $A_m u(x) := \int_{\Omega} K_m(x, y) u(y) dy$  und erhalten

$$\begin{aligned} \|A - A_m\| &= \sup_{\|u\|_{L^2}=1} \int_{\Omega} \left| \int_{\Omega} (K(x, y) - K_m(x, y)) u(y) dy \right|^2 dx \\ &\leq \int_{\Omega} \int_{\Omega} |K(x, y) - K_m(x, y)|^2 dx dy \rightarrow 0 \end{aligned}$$

Da  $A_m$  vollstetig ist, impliziert der folgende Satz die Aussage von Lemma 8.6. ■

**8.7 Satz.** *Seien  $H_1, H_2$  Hilberträume und  $A_m : H_1 \rightarrow H_2$  linear und vollstetig. Sei*

$$\|A - A_m\| \rightarrow 0 \quad (m \rightarrow \infty) .$$

*Dann ist  $A$  vollstetig.*

Interessanterweise führt also die Konvergenz bezüglich der Operatoren nicht aus der Menge der vollstetigen Abbildungen heraus.

BEWEIS : Sei  $u_j \rightharpoonup u$  schwach in  $H_1$ . Es gilt  $\|u_j\| \leq K$  für alle  $j \in \mathbb{N}$  und somit  $\|u\| \leq K$ . Sei  $\varepsilon > 0$  vorgegeben. Dann gilt

$$\|A - A_m\| < \frac{\varepsilon}{4K}$$

für mindestens ein  $m$ . Es konvergiert  $A_m u_j \rightarrow A_m u$  ( $j \rightarrow \infty$ ) stark in  $H_2$ . Daraus folgt

$$\begin{aligned} \|Au_j - Au\| &= \|A_m u_j - A_m u + (A - A_m) u_j - (A - A_m) u\| \\ &\leq \|A_m u_j - A_m u\| + \|A - A_m\| \|u_j\| + \|A - A_m\| \|u\| \\ &< o(1) + \frac{\varepsilon}{2} < \varepsilon . \end{aligned}$$

■

Ein weiterer, äußerst wichtiger kompakter Operator ist die *Einbettung*  $J$  von  $H^{1,2}$  nach  $L^2$ . Es sei  $H_1 = H^{1,2}(I)$ ,  $I = [a, b]$ , oder  $H_1 = H_0^{1,2}(\Omega)$ ,  $\Omega$  offen und beschränkt. (Im Fall  $H_1 = H^{1,2}(\Omega)$  würde man noch geringfügige Regularität an  $\partial\Omega$  benötigen, z.B.  $\partial\Omega \in C^{0,1}$  reicht völlig.)

**8.8 Satz.** *Sei  $J : H_1 \rightarrow L^2$  die Abbildung, welche einer Funktion  $u \in H_1$  ( $H_1$  wie oben) die Funktion  $u$ , aufgefasst als Element von  $L^2$ , zuordnet. Dann ist  $J$  vollstetig, d.h. aus  $u_m \rightharpoonup u$  schwach in  $H_1$  folgt  $u_m \rightarrow u$  stark in  $L^2$ .*

BEWEIS : (i) Der eindimensionale Fall  $H_1 = H^{1,2}(I)$

Sei  $u_m \in H^{1,2}(I)$ ,  $\|u_m\|_{H^{1,2}} \leq K$ . Dann ist (vgl. Definition von  $H^{1,2}(I)$  nach Tonelli)

$$|u_m(\eta) - u_m(\xi)| = \left| \int_{\xi}^{\eta} u'_m ds \right| \leq |\eta - \xi|^{1/2} \left( \int_I |u'_m|^2 dt \right)^{1/2} \leq K |\eta - \xi|^{1/2},$$

d.h. die  $u_m$  sind gleichgradig stetig. Die Beschränktheit folgt, da

$$\left| \int_I u_m(\eta) d\eta - u_m(\xi) \right| \leq \int_I |u_m(\eta) - u_m(\xi)| d\eta \leq K \int_I |\eta - \xi|^{1/2} d\eta \leq \tilde{K},$$

wobei  $\int_I f(s) ds \equiv \frac{1}{|I|} \int_I f(s) ds$  das *Mittelwertintegral* ist, und

$$\left| \int_I u_m(\eta) d\eta \right| \leq |I|^{1/2} \|u_m\|_{L^2}.$$

Nach dem Satz von Arzela-Ascoli gibt es daher eine Teilfolge, die gleichmässig und damit auch stark in  $L^2(I)$  konvergiert.

(ii) Der Fall  $H_1 = H_0^{1,2}(\Omega)$

Da  $H_0^{1,2}(\Omega)$  der Abschluss von  $C_0^\infty(\Omega)$  in der  $H^{1,2}(\Omega)$ -Norm ist, können wir  $H_0^{1,2}(\Omega)$  als Teilraum von  $H_0^{1,2}(Q)$  mit einem Würfel  $Q = [-\pi L, \pi L]^n \supseteq \supseteq \Omega$  betrachten.<sup>7</sup> Die  $H_0^{1,2}(Q)$ -Funktionen werden periodisch auf ganz  $\mathbb{R}^n$  fortgesetzt mit  $Q$  als Periodizitätswürfel. Sei nun  $u_m \in H_0^{1,2}(Q)$  beschränkt. Dann besitzt  $u_m \in L^2(Q)$  nach Folgerung 6.20 und Satz 6.9 eine Fourierentwicklung

$$u_m = \sum_k c_k^m e^{ikx/L},$$

wobei die Summation über die Multiindizes  $k = (k_1, k_2, \dots, k_n)$ ,  $k_j \in \mathbb{Z}$ , läuft. Die Koeffizienten  $c_k$  sind gegeben durch

$$c_k^m = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_Q u_m(x) e^{-\frac{ikx}{L}} dx.$$

<sup>7</sup>Die Funktionen aus  $H_0^{1,2}(\Omega)$  werden durch 0 auf  $Q \setminus \Omega$  fortgesetzt.

Da auch  $D_j u^m \in L^2(Q)$ ,  $j = 1, \dots, n$ , haben wir auch

$$D_j u^m = \sum_k d_k^{m,j} e^{\frac{ikx}{L}},$$

wobei

$$\begin{aligned} d_k^{m,j} &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_Q D_j u^m(x) e^{-\frac{ikx}{L}} dx \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_Q u^m(x) e^{-\frac{ikx}{L}} \frac{ik_j}{L} dx \\ &= \frac{ik_j}{L} c_k^m. \end{aligned}$$

Die Beschränktheit der Folge  $\|u_m\|_{H^{1,2}}$  zieht nach Satz 6.9 die Beschränktheit von

$$\sum_k |k|^2 |c_k^m|^2 \leq K, \quad |c_{0,0,\dots,0}^m|^2 \leq K \quad (8.9)$$

nach sich. Aus (8.9) folgt die Kompaktheit der Folge  $c^m = (\dots, c_l^m, \dots)$  in  $l^2$ . Seine Elemente sind hier Folgen, deren Glieder mit Multi-Indizes versehen sind. Wählt man nämlich nach dem Doppelfolgensatz eine Teilfolge  $\Lambda \subset \mathbb{N}$  aus mit  $c^m \rightarrow c$  *komponentenweise*, so folgt aus (8.9)  $\sum_k |k|^2 |c_k|^2 \leq K$  und

$$\begin{aligned} \sum_k |c_k^m - c_l|^2 &\leq \sum_{|k| \leq N} |c_k^m - c_k|^2 + \frac{1}{N^2} \sum_{|k| \geq N} |k|^2 |c_k^m - c_k^2| \\ &\leq o(1) + \frac{2K}{N^2} < \varepsilon \quad \text{für } m \geq m(\varepsilon). \end{aligned}$$

Auf die Fouriersche Reihe übertragen, impliziert die Konvergenz  $c^m \rightarrow c$  in  $l^2$  die Konvergenz  $u^m \rightarrow u$  stark in  $L^2(Q)$ , wobei  $c \in l^2$  der Vektor der Fourierkoeffizienten der  $u$  definierenden Reihe ist. ■

Im Falle  $H_1 = H^{1,2}(\Omega)$ ,  $\partial\Omega \in C^{0,1}$  läuft der Beweis wie im Falle  $H_0^{1,2}(\Omega)$ . Allerdings muss man folgende Aussage benutzen:

Es existiert ein *Fortsetzungsoperator*  $C : H^{1,2}(\Omega) \rightarrow H_0^{1,2}(Q)$ , d.h.

$$(Cu)|_\Omega = u \quad \forall u \in H^{1,2}(\Omega)$$

und

$$\|Cu\|_{H_0^{1,2}(Q)} \leq k \|u\|_{H^{1,2}(\Omega)}.$$

Zur Konstruktion dieses Operators wird die Regularität des Randes benutzt.

Satz 8.8 lässt sich auch folgendermaßen interpretieren:

**8.10 Lemma.** Sei  $J : H_1 \rightarrow H_1$ , wobei  $H_1$  wie in Satz 8.8 definiert ist, die durch den Rieszschen Darstellungssatz gegebene Abbildung mit

$$(Ju, v)_{H_1} = (u, v)_{L^2} \quad \text{für alle } u, v \in H_1. \quad (8.11)$$

Dann ist  $J$  kompakt.

BEWEIS : Sei  $\|u_m\|_{H_1} \leq K$ . Dann folgt aus (8.11) mit  $v = Ju_m$ ,  $u$  ersetzt durch  $u_m$

$$\|Ju_m\|_{H_1} \leq \|u_m\|_{L^2} \leq K.$$

Die  $Ju_m$  sind also beschränkt. Andererseits gibt es wegen Satz 8.8 eine Teilfolge  $\Lambda$ , so dass  $(u_m)_{m \in \Lambda}$  eine Cauchy-Folge in  $L^2$  ist. Daraus folgt

$$\sup_{\|v\|_{H_1}=1} (Ju_m - Ju_k, v)_{H_1} \leq \|u_m - u_k\|_{L^2} \rightarrow 0 \quad (m, k \rightarrow \infty, m, k \in \Lambda),$$

woraus  $\|Ju_m - Ju_k\|_{H_1} \rightarrow 0$  folgt. ■

Jetzt können wir Satz 5.19 beweisen.

**5.19 Satz.** Sei  $\Omega$  ein beschränktes Gebiet des  $\mathbb{R}^n$  mit  $\partial\Omega \in C^{0,1}$ . Dann existiert eine Konstante  $K$ , so dass für alle  $u \in H^{1,2}(\Omega)$  gilt:

$$\int_{\Omega} |u - u_{\Omega}|^2 dx \leq K \int_{\Omega} |\nabla u|^2 dx, \quad (5.20)$$

wobei

$$u_{\Omega} := \int_{\Omega} u dx := \frac{1}{|\Omega|} \int_{\Omega} u dx$$

der Mittelwert von  $u$  über  $\Omega$  ist.

BEWEIS : Da sich die linke Seite nicht verändert, wenn man zu  $u$  eine beliebige Konstante addiert können wir annehmen, dass  $u_{\Omega} = 0$ . Sei die Behauptung falsch. Dann gibt es eine Folge  $u_k \in H^{1,2}(\Omega)$  mit Mittelwert Null, so dass

$$\begin{aligned} \int_{\Omega} |u_k|^2 dx &= 1, \\ \int_{\Omega} |\nabla u_k|^2 dx &\leq \frac{1}{k}. \end{aligned} \quad (8.12)$$

Somit ist die Folge  $u_k$  in  $H^{1,2}(\Omega)$  beschränkt und nach den Bemerkungen zu Satz 8.8 gibt es eine Teilfolge  $u_{k_j}$  die in  $L^2(\Omega)$  stark gegen  $u \in L^2(\Omega)$  konvergiert, wobei aufgrund von (8.12)  $\|u\|_{L^2} = 1$ , und die in  $H^{1,2}(\Omega)$  schwach gegen  $u$  konvergiert. Nach (8.12) und Lemma 7.2 und Satz 1.3.15 gilt  $\nabla u = 0$ . Daraus folgt dass  $u$  in  $\Omega$  konstant ist. In der Tat, sei  $u_{\varepsilon} = J_{\varepsilon} * u$  die Regularisierung von  $u$ . Dann gilt für alle

$$x \in \Omega_\varepsilon = \{x \in \Omega; \text{dist}(x, \partial\Omega) > \varepsilon\}$$

$$\begin{aligned} \nabla u_\varepsilon(x) &= \nabla_x \int_{\mathbb{R}^n} J_\varepsilon(x-y)u(y) dy \\ &= \nabla_x \int_{\mathbb{R}^n} J_\varepsilon(y)u(x-y) dy \\ &= \int_{\mathbb{R}^n} J_\varepsilon(y)\nabla_x u(x-y) dy \\ &= J_\varepsilon * \nabla u = 0, \end{aligned}$$

da  $\nabla u \in L^2(\Omega)$  ist. Somit sind  $u_\varepsilon$  in  $\Omega_\varepsilon$  konstant. Nach Satz 1.5 konvergiert  $u_\varepsilon \rightarrow u$  in  $L^2(\Omega)$  und somit ist  $u$  auf ganz  $\Omega$  konstant. Der Grenzwert  $u$  hat auch Mittelwert Null, da

$$\left| \int_{\Omega} u_k dx - \int_{\Omega} u dx \right| \leq |\Omega|^{\frac{1}{2}} \|u_k - u\|_{L^2} \rightarrow 0 \quad (k \rightarrow \infty)$$

und die Mittelwerte von  $u_k$  Null sind. Damit ist  $u$  notwendig identisch Null, was ein Widerspruch zu  $\|u\|_{L^2} = 1$  ist. ■

Als weitere Anwendung der Begriffe werden wir in diesem Abschnitt die Endlichdimensionalität des Kernes von Operatoren der Gestalt  $A + K$ ,  $A$  koerziv und beschränkt,  $K$  kompakt, und verwandte Fragen behandeln. Zuvor benötigen wir noch

**8.13 Satz.** *Sei  $K : H_1 \rightarrow H_2$  eine lineare, vollstetige Abbildung eines Hilbertraumes  $H_1$  in einen Hilbertraum  $H_2$ . Dann ist die adjungierte Abbildung  $K^* : H_2 \rightarrow H_1$  vollstetig.*

BEWEIS : Sei  $v_m \rightharpoonup v$  schwach in  $H_2$ . Nach Definition des Begriffs „Supremum“ gibt es Elemente  $u_m$  mit  $\|u_m\| = 1$  und Zahlen  $\varepsilon_m \rightarrow 0$ , so dass

$$\|K^*v_m - K^*v\| = \sup_{\|u\|=1} |(u, K^*v_m - K^*v)| = |(u_m, K^*v_m - K^*v)| + \varepsilon_m.$$

O.B.d.A. kann man  $u_m$  so wählen, dass  $u_m \rightharpoonup u$  in  $H_1$  (vergleiche Satz 7.5). Wir haben also

$$\|K^*v_m - K^*v\| = |(Ku_m, v_m - v)| + \varepsilon_m, \quad (8.14)$$

wobei  $u_m \rightharpoonup u$  schwach in  $H_1$  und, da  $K$  vollstetig ist,  $Ku_m \rightarrow f$  stark in  $H_2$ . Da weiterhin  $v_m - v \rightharpoonup 0$  schwach konvergiert, gilt

$$(Ku_m, v_m - v) \rightarrow 0 \quad (m \rightarrow \infty).$$

In der Tat: Angenommen, es wäre  $|(Ku_m, v_m - v)| > \varepsilon$  für eine Teilfolge  $\Lambda$ . Man hat dann den Widerspruch

$$|(Ku_m, v_m - v)| = |(Ku_m - f, v_m - v) + o(1)| \leq \|v_m - v\| \|Ku_m - f\| + o(1) \rightarrow 0.$$

Mit (8.14) folgt daher die Behauptung. ■

**8.15 Satz.** *Es sei  $A : H_1 \rightarrow H_2$  eine lineare, beschränkte Abbildung zwischen den Hilberträumen  $H_1$  und  $H_2$ .  $A$  genüge der Koerzivitätsungleichung*

$$\|Au\| \geq c\|u\|, \quad u \in H_1$$

*mit einer positiven Konstanten  $c$ . Ferner sei  $K : H_1 \rightarrow H_2$  linear und vollstetig. Dann ist der Bildraum  $(A + K)(H_1)$  abgeschlossen und der Kern von  $A + K$  endlichdimensional.*

BEWEIS : (i) Angenommen, es wäre

$$\dim N(A + K) = \infty. \quad (8.16)$$

Dann gibt es nach Lemma 6.4 ein Orthonormalsystem  $\{\varphi_j\}$ ,  $\varphi_j \in N(A + K)$ . Da  $\varphi_j \in N(A + K)$ , gilt  $A\varphi_j = -K\varphi_j$ , d.h. die Folge  $A\varphi_j$  ist relativ kompakt, und es gibt eine stark konvergente Teilfolge  $(A\varphi_{j_l})$ . Da  $\varphi_{j_l} \perp \varphi_{j_k}$ ,  $l \neq k$ , folgt  $\|\varphi_{j_l} - \varphi_{j_k}\| = \sqrt{2}$ , und wegen der vorausgesetzten Koerzivitätsungleichung

$$\|A\varphi_{j_l} - A\varphi_{j_k}\| \geq c\sqrt{2}, \quad l \neq k.$$

Die Folge  $(A\varphi_{j_l})$  kann daher nicht konvergieren. Die Annahme (8.16) wurde daher zum Widerspruch geführt.

(ii) Es sei  $(A + K)u_m \rightarrow f$  in  $H_2$ . Wir müssen zeigen, dass ein  $u \in H_1$  existiert mit  $(A + K)u = f$ . Da  $A$  und  $K$  stetig sind, ist der Kern  $N(A + K)$  von  $A + K$  abgeschlossen. Jedes  $u_m$  hat somit eine eindeutige Zerlegung

$$u_m = v_m + w_m, \quad w_m \in N(A + K), \quad v_m \perp N(A + K).$$

Wir werden zeigen, dass die  $v_m$  konvergieren. Nach dem im Anschluß bewiesenen Lemma gibt es eine positive Konstante  $c_0$ , so dass

$$\|(A + K)v\| \geq c_0\|v\| \quad \text{für alle } v \in (N(A + K))^\perp. \quad (8.17)$$

Da  $(A + K)u_m = (A + K)v_m$ , gilt auch

$$(A + K)v_m \rightarrow f \quad (m \rightarrow \infty)$$

und  $((A + K)v_m)$  ist eine Cauchyfolge. Wegen (8.17) ist daher auch  $(v_m)$  eine Cauchyfolge (wende (8.17) mit  $v = v_m - v_k$  an). Daher folgt  $f = (A + K)v$ , und Satz 8.15 ist bewiesen. ■

**8.18 Lemma.** *Sei  $A : H_1 \rightarrow H_2$  linear und beschränkt,  $K : H_1 \rightarrow H_2$  linear und vollstetig, und es gelte*

$$\|Au\| \geq c\|u\| \quad \text{für alle } u \in H_1 \quad (8.19)$$

*mit einer positiven Konstanten  $c$ . Dann gibt es eine positive Konstante  $c_0$ , so dass*

$$\|Au + Ku\| \geq c_0\|u\| \quad \text{für alle } u \in (N(A + K))^\perp.$$

BEWEIS : Angenommen, die Behauptung wäre falsch. Dann gibt es eine Folge  $u_m \in (N(A + K))^\perp$  mit  $\|u_m\| = 1$ ,  $\|Au_m + Ku_m\| \rightarrow 0$ . Da  $K$  kompakt und  $(u_m)$  beschränkt, ist die Folge  $(Ku_m)$  relativ kompakt, somit auch die Folge  $Au_m$ . In der Tat, es gibt eine Teilfolge  $\Lambda \subset \mathbb{N}$ , so dass  $Ku_m - Ku_k \rightarrow 0$  ( $m, k \in \Lambda$ ,  $m, k \rightarrow \infty$ ). Da weiterhin  $Au_m + Ku_m \rightarrow 0$ , konvergiert auch

$$Au_m - Au_k + Ku_m - Ku_k \rightarrow 0 \quad (m, k \in \Lambda, m, k \rightarrow \infty),$$

und somit auch  $Au_m - Au_k \rightarrow 0$ . Aus (8.19), angewandt auf  $u_m - u_k$ , folgt, dass  $(u_m)_{m \in \Lambda}$  eine Cauchy-Folge ist. Somit gilt  $u_m \rightarrow u$  ( $m \rightarrow \infty$ ,  $m \in \Lambda$ ), und da  $\|u_m\| = 1$ , folgt  $\|u\| = 1$ , also  $u \neq 0$ . Da  $u_m \in N(A + K)^\perp$  und  $N(A + K)^\perp$  abgeschlossen ist, gilt auch  $u \in N(A + K)^\perp$ . (Orthogonalkomplemente sind immer abgeschlossen:  $0 = (u_m, v) \rightarrow (u, v)$ .) Damit gilt  $(A + K)u \neq 0$ , da  $(N(A + K))^\perp \cap N(A + K) = \{0\}$ , und es ergibt sich der Widerspruch

$$(A + K)u_m \rightarrow 0 \quad \text{und} \quad (A + K)u \neq 0.$$

■

Die Endlichdimensionalität des Defektraumes  $((A + K)(H_1))^\perp$  erhält man analog zu Satz 8.15 durch eine Voraussetzung an  $A^*$ .

**8.20 Satz.** *Es sei  $A : H_1 \rightarrow H_2$  eine lineare, beschränkte Abbildung zwischen den Hilberträumen  $H_1, H_2$ . Die adjungierte Abbildung  $A^*$  erfülle*

$$\|A^*v\| \geq c\|v\|, \quad v \in H_2$$

*mit einer positiven Konstanten  $c$ . Ferner sei  $K : H_1 \rightarrow H_2$  vollstetig und linear. Dann ist*

$$\dim ((A + K)H_1)^\perp < \infty.$$

BEWEIS : Nach Satz 8.15 und Satz 8.13, Satz 5.3 ist  $\dim N(A^* + K^*) < \infty$ . Sei nun  $v \in ((A + K)H_1)^\perp$ , d.h.  $((A + K)u, v) = 0$  für alle  $u \in H_1$ . Daraus folgt  $((A + K)H_1)^\perp \subset N(A^* + K^*)$ . Da  $\dim N(A^* + K^*) < \infty$  und  $((A + K)H_1)^\perp$  linear und abgeschlossen ist, folgt die Endlichkeit der Dimension von  $((A + K)H_1)^\perp$ . ■

Ist  $H_1 = H_2 = H$ , so folgen die Bedingungen  $\|Au\| \geq c\|u\|$ ,  $\|A^*u\| \geq c\|u\|$  beide aus der Bedingung

$$\operatorname{Re}(Au, u) \geq c\|u\|^2. \tag{8.21}$$

Im Folgenden bringen wir eine bedeutende Anwendung der Ergebnisse aus Abschnitt 2.5 und Abschnitt 2.8 auf Randwertprobleme elliptischer partieller Differentialgleichungen und auch gewöhnlicher Differentialgleichungen. Entscheidend ist, dass man für Randwertprobleme

$$-\sum_{i,k=1}^n D_i(a_{ik}(x) D_k u) + \sum_{i=1}^n b_i(x) D_i u + c(x) u = f \quad \text{in } \Omega$$

mit Randbedingungen auf  $\partial\Omega$ , sowie das eindimensionale Analogon

$$-(pu')' + ru' + qu = f \quad \text{in } [a, b] \text{ und Randbedingung}$$

die sogenannte *Gårdingsche Ungleichung* beweisen kann:

**8.22 Satz.** Sei  $\Omega$  ein beschränktes Gebiet des  $\mathbb{R}^n$  und sei  $a_{ik} \in L^\infty(\Omega)$ ,  $b_i \in L^\infty(\Omega)$ ,  $c \in L^\infty(\Omega)$ ,  $f \in L^2(\Omega)$ . Es gelte die Elliptizitätsbedingung

$$\sum_{i,k=1}^n a_{ik}(x) \xi_i \xi_k \geq \alpha |\xi|^2$$

mit einer positiven Konstanten  $\alpha > 0$ . Die beschränkte Bilinearform  $Q$  sei definiert durch

$$Q(u, v) := \int_{\Omega} \left[ \sum_{i,k=1}^n a_{ik} D_k u D_i v + \sum_{i=1}^n b_i D_i u v + c u v \right] dx.$$

Dann gilt die **Gårdingsche Ungleichung**

$$Q(u, u) \geq c_0 \|u\|_{H^{1,2}}^2 - \lambda_0 \|u\|_{L^2}^2 \quad \text{für alle } u \in H^{1,2}(\Omega)$$

mit einer positiven Konstanten  $c_0 > 0$  sowie einer Konstanten  $\lambda_0$ .

Für gewöhnliche Differentialgleichungen lassen wir eine noch etwas allgemeinere Bilinearform zu:

**8.23 Satz.** Sei  $I = [a, b]$ ,  $p \in L^\infty(I)$ ,  $p \geq c_1 = \text{const} > 0$ ,  $r \in L^\infty(I)$ ,  $q \in L^\infty(I)$ ,  $\alpha < \beta \in \mathbb{R}$  und

$$Q(u, v) := \int_a^b [p u' v' + r u' v + q u v] dx + \beta u(b) v(b) - \alpha u(a) v(a).$$

Dann gilt die **Gårdingsche Ungleichung**

$$Q(u, u) \geq c_0 \|u\|_{H^{1,2}}^2 - \lambda_0 \|u\|_{L^2}^2 \quad \text{für alle } u \in H^{1,2}(I),$$

mit einer positiven Konstanten  $c_0 > 0$  und einer Konstanten  $\lambda_0$ .

BEWEIS (Satz 8.22): Wegen der Elliptizitätsbedingung schätzt man

$$\int_{\Omega} \sum_{i,k=1}^n a_{ik} D_k u D_i u dx \geq \alpha \int_{\Omega} |\nabla u|^2 dx = \alpha \|u\|_{H^{1,2}}^2 - \alpha \|u\|_{L^2}^2$$

ab. Den zweiten Summand schätzt man nach unten ab,

$$\sum_{i=1}^n \int_{\Omega} b_i D_i u u dx \geq -\frac{\alpha}{2} \int_{\Omega} |\nabla u|^2 dx - \frac{\|b\|_{\infty}^2}{2\alpha} \int_{\Omega} |u|^2 dx,$$

woraus sich

$$\begin{aligned} Q(u, u) &\geq \alpha \int_{\Omega} |\nabla u|^2 dx - \frac{\alpha}{2} \int_{\Omega} |\nabla u|^2 dx - \frac{\|b\|_{\infty}^2}{2\alpha} \int_{\Omega} |u|^2 dx - \|c\|_{\infty} \int_{\Omega} |u|^2 dx \\ &= \frac{\alpha}{2} \left( \int_{\Omega} |\nabla u|^2 dx + \int_{\Omega} |u|^2 dx \right) - \left( \frac{\|b\|_{\infty}^2}{2\alpha} + \frac{\alpha}{2} + \|c\|_{\infty} \right) \int_{\Omega} |u|^2 dx \end{aligned}$$

ergibt. ■



BEWEIS (Satz 8.23): Wir gehen wie im Beweis von Satz 8.22 vor. Der zusätzliche Term wird wie folgt behandelt. Aus  $\beta > \alpha$  folgt

$$\begin{aligned} \beta |u(b)|^2 - \alpha |u(a)|^2 &\geq \beta (|u(b)|^2 - |u(a)|^2) + (\beta - \alpha) |u(a)|^2 \\ &\geq \beta (|u(b)|^2 - |u(a)|^2) . \end{aligned}$$

Weiter haben wir

$$|u(b)|^2 - |u(a)|^2 \leq \int_a^b (|u|^2)' dx \leq \varepsilon \int_a^b |u'|^2 dx + K_\varepsilon \int_a^b |u|^2 dx$$

und somit für  $\varepsilon = \frac{c_0}{4|\beta|}$

$$\begin{aligned} Q(u, u) &\geq c_1 \int |u'|^2 dt - \frac{c_1}{2} \int |u'|^2 dt - \frac{\|r\|_\infty^2}{2c_1} \int |u|^2 dt - \|q\|_\infty \int |u|^2 dt \\ &\quad - \frac{c_1}{4} \int |u'|^2 dt - K_\varepsilon \int |u|^2 dt \\ &\geq \frac{c_1}{4} \left( \int |u'|^2 dt + \int |u|^2 dt \right) - \left( \frac{\|r\|_\infty^2}{2c_1} - \|q\|_\infty - \frac{c_1}{4} - K_\varepsilon \right) \int_a^b |u|^2 dt . \end{aligned}$$

■

Um die Gårdingsche Ungleichung zu verwenden, setzen wir - nach dem Rieszschen Darstellungssatz -

$$(Au, v)_H = Q(u, v) \tag{8.24}$$

mit  $H = H_0^{1,2}$  oder  $H = H^{1,2}$  als Grundraum. Die Abbildung  $A$  ist dann eine beschränkte lineare Abbildung von  $H$  in sich und erfüllt

$$(Au, u)_H \geq c \|u\|_H^2 - \lambda_0 \|u\|_{L^2}^2 . \tag{8.25}$$

Es sei  $J$  die nach Lemma 8.10 *vollstetige*, lineare Abbildung von  $H$  nach  $H$ , für die gilt:

$$(Ju, v)_H = (u, v)_{L^2} .$$

Aus (8.24) folgt

$$((A + \lambda_0 J)u, u)_H \geq c \|u\|_H^2 , \tag{8.26}$$

und wir können Satz 8.20, Satz 8.15 und die Bemerkung im Zusammenhang mit (8.21) auf die Abbildung  $A + \lambda_0 J$  (anstelle von  $A$ ) und  $-\lambda_0 J$  (anstelle von  $K$ ) anwenden. Dies ergibt den *Hauptsatz zur sogenannten normalen Lösbarkeit elliptischer Randwertprobleme*.

**8.27 Satz.** *Unter den Voraussetzungen von Satz 8.22 oder Satz 8.23 ist das Randwertproblem:*

„gesucht ist  $u \in H_0^{1,2}$  bzw.  $H^{1,2}$ , so dass

$$Q(u, v) = (f, v)_{L^2} \quad \text{für alle } v \in H_0^{1,2} \text{ bzw. } H^{1,2} \text{ “}$$

lösbar genau dann, wenn

$$(f, v)_{L^2} = 0 \quad \text{für alle } v \in N(A^*),$$

$$N(A^*) = \{v \in H_0^{1,2} \text{ bzw. } H^{1,2} \mid Q(u, v) = 0 \text{ für alle } u \in H_0^{1,2} \text{ bzw. } H^{1,2}\}.$$

Es gilt  $\dim N(A^*) < \infty$  und  $\dim N(A) < \infty$  mit

$$N(A) = \{u \in H_0^{1,2} \text{ bzw. } H^{1,2} \mid Q(u, v) = 0 \text{ für alle } v \in H_0^{1,2} \text{ bzw. } H^{1,2}\}.$$

BEWEIS : Wir wollen die Fredholmsche Alternative (Folgerung 5.6) auf den Operator  $A : H \rightarrow H$  definiert in (8.24) anwenden. Mit den Bezeichnungen  $\tilde{A} := A + \lambda_0 J$  und  $K := -\lambda_0 J$  haben wir  $A = \tilde{A} + K$ . Aufgrund von (8.26) und der Bemerkung in Zusammenhang mit (8.21) sind alle Voraussetzungen von Satz 8.15 erfüllt und somit ist das Bild  $A(H)$  abgeschlossen und  $\dim N(A) < \infty$ . Satz 8.20 liefert  $\dim N(A^*) < \infty$ . Da das Bild  $A(H)$  abgeschlossen ist, ist die Aussage von Satz 8.25 nichts anderes als eine Umformulierung von Folgerung 5.6. ■

## 2.9 Das Eigenwertproblem für vollstetige lineare selbstadjungierte Operatoren

Aus der linearen Algebra wissen wir, dass jede hermitesche (=selbstadjungierte) lineare Abbildung eine vollständige Orthonormalbasis von Eigenwerten besitzt und die Eigenwerte reell sind. Eine völlig analoge Aussage gelingt für vollstetige, lineare, selbstadjungierte Abbildungen eines Hilbertraumes in sich. Darüberhinaus erhält man noch, dass es höchstens *abzählbar viele Eigenwerte* gibt und diese *gegen Null gehen* müssen, und die *Eigenräume* mit eventueller Ausnahme des Kernes *endlich-dimensional* sind. Ferner werden die Eigenvektoren über ein Maximum-Minimum-Prinzip erhalten. Wir werden diese Aussagen im Folgenden beweisen, zum Teil unter schwächeren Voraussetzungen.

Der Vollständigkeit halber erinnern wir an die

**9.1 Definition.** Sei  $A : H \rightarrow H$ . Ein Element  $u \in H$ ,  $u \neq 0$ , heißt **Eigenvektor** von  $A$  zum **Eigenwert**  $\lambda \in \mathbb{R}$  bzw.  $\mathbb{C}$ , falls

$$Au = \lambda u.$$

Allgemeiner lässt sich auch die Situation  $A : D(A) \rightarrow H$ ,  $D(A)$  dicht in  $H$ , betrachten. Dann ist  $0 \neq u \in D(A)$  ein Eigenvektor zum Eigenwert  $\lambda \in \mathbb{R}$  bzw.  $\mathbb{C}$ , falls  $Au = \lambda u$ .

**9.2 Lemma.** Sei  $H$  ein Hilbertraum über  $\mathbb{C}$  und  $D(A)$  ein dichter, linearer Teilraum von  $H$ . Die lineare Abbildung  $A : D(A) \rightarrow H$  sei selbstadjungiert. Dann ist jeder Eigenwert von  $A$  reell und Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal.

BEWEIS : Sei  $u \in D(A)$ ,  $u \neq 0$  und  $Au = \lambda u$ . Dann folgt  $(Au, u) = \lambda (u, u)$ , und, da  $A$  selbstadjungiert ist,  $A = A^*$  und  $u \in D(A^*)$ , d.h.

$$\lambda (u, u) = (Au, u) = (u, A^*u) = (u, Au) = \bar{\lambda} (u, u).$$

Daraus folgt

$$\lambda = \bar{\lambda}.$$

Gilt  $Au = \lambda u$ ,  $Av = \mu v$ , mit  $\lambda \neq \mu$ ,  $u, v \neq 0$ , so folgt

$$\lambda(u, v) = (Au, v) = (u, Av) = (u, \mu v) = \bar{\mu}(u, v) = \mu(u, v),$$

woraus  $(u, v) = 0$  folgt. ■

**9.3 Lemma.** *Sei  $H$  ein Hilbertraum und  $K : H \rightarrow H$  vollstetig und linear. Sei  $\lambda \in \mathbb{C}$  ein beliebiger, von Null verschiedener Eigenwert. Dann ist der Eigenraum  $V$  zu  $\lambda$*

$$V := \{u \in H \mid Ku = \lambda u\}$$

*endlich-dimensional.*

BEWEIS : Wenn  $V$  leer oder endlich-dimensional ist, ist nichts zu beweisen. Ist  $\dim V = \infty$ , so gibt es einen separablen Teilraum  $V_0$  mit  $\dim V_0 = \infty$ , und  $V_0$  besitzt ein vollständiges Orthonormalsystem  $\{\varphi_j \mid j \in \mathbb{N}\}$ . Da  $K$  vollstetig ist und damit kompakt ist und die Menge der  $\varphi_j$  beschränkt ist - es gilt ja  $\|\varphi_j\| = 1$  -, wird sie von  $K$  in eine relativ kompakte Menge überführt. Andererseits gilt  $K\varphi_j = \lambda\varphi_j$ , also ist auch die Menge der  $\varphi_j$  relativ kompakt (beachte, dass hier  $\lambda \neq 0$  benutzt wird). Ein abzählbar unendliches Orthonormalsystem kann jedoch nicht kompakt sein, da es wegen der Beziehung  $\|\varphi_j - \varphi_k\| = \sqrt{2}$ ,  $j \neq k$ , niemals eine Cauchyfolge bilden kann. Die Annahme  $\dim V = \infty$  ist daher nicht möglich. ■

Mit einer ähnlichen Methode beweist man

**9.4 Lemma.** *Es sei  $H$  ein Hilbertraum und  $K : H \rightarrow H$  vollstetig, linear und selbstadjungiert. Dann gibt es keinen von Null verschiedenen Häufungswert von paarweise verschiedenen Eigenwerten von  $K$ .*

BEWEIS : Sei  $Kv_j = \lambda_j v_j$  mit  $v_j \in H$ ,  $\|v_j\| = 1$ ,  $\lambda_j \neq \lambda_k \neq 0$  für  $j \neq k$ ,  $\lambda_j \in \mathbb{R}$ ,  $\lambda_j \rightarrow \lambda \neq 0$  ( $j \rightarrow \infty$ ),  $\lambda \in \overline{\mathbb{R}}$ . Da  $(v_j)$  beschränkt ist, ist auch die Folge  $(\lambda_j^{-1} v_j)$  beschränkt (hierbei wurde  $\lambda_j \rightarrow \lambda \neq 0$  benutzt) und somit ist die Folge  $\lambda_j^{-1} K v_j$  relativ kompakt. Daher ist auch die Folge  $(v_j)$  relativ kompakt. Dies ist jedoch nicht möglich, da  $v_j \perp v_k$  für  $j \neq k$  nach Lemma 9.3. Die  $v_j$  müssen daher ein abzählbar unendliches Orthonormalsystem bilden, welches nicht relativ kompakt sein kann - siehe Beweis von Lemma 9.3. Dies ist ein Widerspruch und somit ist Null der einzig mögliche Häufungspunkt. ■

- Aus Lemma 9.4 folgern wir sofort: Der Kern von  $K$  kann also eventuell unendlich-dimensional sein. Die Eigenwerte von  $K$  gehen gegen Null, sofern  $\dim N(K) < \infty$ , ( $\dim H = \infty$ ). Falls  $\dim N(K) = \infty$ , gibt es entweder abzählbar viele paarweise verschiedene Eigenwerte, die gegen Null gehen, oder aber nur endlich viele, von Null verschiedene Eigenwerte.

Es folgen nun Sätze, die die *Existenz* von Eigenwerten sicherstellen.

**9.5 Satz.** *Es sei  $H$  ein Hilbertraum und  $K : H \rightarrow H$  kompakt, linear und selbstadjungiert. Dann gibt es ein  $u \in H$  mit  $\|u\| = 1$  und*

$$|(Ku, u)| = \sup \{ |(Kv, v)| \mid \|v\| = 1 \} =: s. \quad (9.6)$$

*Dieses Element  $u$  ist Eigenvektor zum Eigenwert  $\lambda = (Ku, u)$ , und  $\lambda$  ist der betragsmäßig größte Eigenwert von  $K$ , wobei  $|\lambda| = s$ .*

BEWEIS : (i) Wir zeigen zunächst, dass das Supremum bei (9.6) durch einen Vektor  $u$  angenommen wird. Sei  $(u_m)$  eine *Maximalfolge*, d.h.

$$\|u_m\| = 1, \quad |(Ku_m, u_m)| \rightarrow s \quad (m \rightarrow \infty).$$

Wegen der schwachen Kompaktheit beschränkter Mengen im Hilbertraum und der vorausgesetzten Kompaktheit von  $K$  dürfen wir nach Ummumerierung annehmen, dass

$$u_m \rightharpoonup u \quad \text{schwach} \quad \text{und} \quad Ku_m \rightarrow Ku \quad \text{stark.}$$

Daraus folgt  $(Ku_m, u_m) \rightarrow (Ku, u)$ , denn

$$\begin{aligned} |(Ku_m, u_m) - (Ku, u)| &\leq |(Ku_m - Ku, u_m)| + |(Ku, u_m - u)| \\ &\leq \|Ku_m - Ku\| + o(1) = o(1). \end{aligned}$$

Es gilt daher

$$|(Ku, u)| = s.$$

Wir zeigen, dass  $u \neq 0$  gilt. (Aus der Tatsache, dass  $u_m \rightharpoonup u$  schwach und  $\|u_m\| = 1$ , können wir nämlich nicht notwendig schließen, dass  $\|u\| = 1$  ist - dies ist i.A. falsch.) Aus  $u_m \rightharpoonup u$  schwach,  $\|u_m\| = 1$  folgt nach Lemma 7.2  $\|u\| \leq 1$ . Angenommen, es wäre  $u = 0$ . Dann folgt  $|(Ku, u)| = 0$  und somit  $s = 0$ , und daher

$$(Kv, v) = 0 \quad \text{für alle } v \in H.$$

Aus dem im Anschluß bewiesenen Hilfssatz folgt damit  $K = 0$ , was wir als trivialen Fall von vornherein ausschließen dürfen. Wir dürfen damit  $u \neq 0$  annehmen. Wäre  $\|u\| = \theta < 1$ ,  $\theta > 0$ , so hätte  $\tilde{u} := \theta^{-1}u$  Norm 1, d.h.  $\|\tilde{u}\| = 1$ , und es ergäbe sich der Widerspruch

$$s \geq |(K\tilde{u}, \tilde{u})| = \theta^{-2} |(Ku, u)| = \theta^{-2} s > s.$$

Daher gilt  $\|u\| = 1$ , und die Maximaleigenschaft (9.6) von  $u$  ist gezeigt.

(ii) Wir zeigen, dass Maximal- bzw. Minimalstellen  $u$  von  $(Ku, u)$ ,  $\|u\| = 1$ , Eigenvektoren sind. Sei  $w(t) = u \cos t + v \sin t$ ,  $v \in H$ ,  $\|v\| = 1$ ,  $v \perp u$ . Wegen der Maximaleigenschaft (o.B.d.A.) von  $u$  gilt  $(Ku, u) \geq (Kw(t), w(t))$ , denn  $\|w(t)\| = 1$ , und somit

$$\left. \frac{d}{dt} (Kw(t), w(t)) \right|_{t=0} = 0.$$

Andererseits ist (da  $K$  selbstadjungiert)

$$\frac{d}{dt}(Kw(t), w(t)) = 2\mathcal{R}e(K\dot{v}(t), v(t)) = 2\mathcal{R}e(-Ku \sin t + Kv \cos t, u \cos t + v \sin t).$$

Damit gilt

$$0 = \frac{d}{dt}(Kw(t), w(t)) \Big|_{t=0} = 2\mathcal{R}e(Kv, u) = 2\mathcal{R}e(v, Ku) \quad \text{für alle } v \perp u. \quad (9.7)$$

Wegen des Projektionssatzes ist  $Ku = \lambda u + v_0$ ,  $v_0 \perp u$ , und wegen (9.7)

$$0 = \mathcal{R}e(v_0, Ku) = \|v_0\|^2,$$

d.h.  $v_0 = 0$  und  $Ku = \lambda u$ . Weiter gilt für das in (i) konstruierte  $u$

$$\lambda = \lambda(u, u) = (Ku, u)$$

und somit  $|\lambda| = s$ . Ist schließlich  $\mu \in \mathbb{R}$  ein weiterer Eigenwert von  $K$  zum normierten Eigenvektor  $u_0$ , so gilt

$$|\mu| = |(\mu u_0, u_0)| = |(Ku_0, u_0)| \leq \sup_{\|u\|=1} |(Ku, u)|.$$

■

**9.8 Lemma.** Sei  $H$  Hilbertraum und  $K : H \rightarrow H$  linear und selbstadjungiert. Es gelte  $(Ku, u) = 0$  für alle  $u \in H$ . Dann ist  $K = 0$ .

BEWEIS : Nach Voraussetzung gilt  $(K(u+v), u+v) = 0$ , somit

$$0 = (Ku, u) + (Kv, v) + 2\mathcal{R}e(Ku, v) = 2\mathcal{R}e(Ku, v) \quad \text{für alle } v \in H.$$

Aus  $(K(u+iv), u+iv) = 0$  folgt  $\mathcal{I}m(Ku, v) = 0$  und somit  $Ku = 0$  für alle  $u \in H$ , d.h.  $K = 0$ . ■

Wir konstruieren uns nun weitere, betragsmäßig monoton fallende Eigenwerte von  $K$  mit Hilfe der folgenden Konstruktion:

Sei  $\lambda_1$  der nach Satz 9.5 existierende betragsmäßig größte Eigenwert. Es gilt

$$\begin{aligned} Ku &= \lambda_1 u & \text{für } u \in V_1^+, \\ Ku &= -\lambda_1 u & \text{für } u \in V_1^-. \end{aligned}$$

wobei einer der „Eigenräume“  $V_1^+, V_1^-$  trivial sein kann. Die „Eigenräume“  $V_1^+, V_1^-$  sind nach Lemma 9.3 endlich dimensional, nach Lemma 9.2 ist  $V_1^+ \perp V_1^-$  und  $\dim V_1^+ \oplus V_1^- \geq 1$  nach Satz 9.5. Zur Konstruktion des betragsmäßig zweitgrößten Eigenwertes lösen wir die *Optimierungsaufgabe*:

$$\begin{aligned} \text{Gesucht ist } u \in H \text{ mit } \|u\| = 1 \text{ und } u \perp W_1, \quad W_1 := V_1^+ \oplus V_1^-, \text{ so dass} \\ |(Ku, u)| = \max. \end{aligned} \quad (9.9)$$

(Im Gegensatz zur Konstruktion des betragsmäßig größten Eigenwertes besteht in der Optimierungsaufgabe noch die *zusätzliche Nebenbedingung*  $u \perp W_1$ .)

**9.10 Lemma.** *Sei  $K : H \rightarrow H$  linear, selbstadjungiert und kompakt und  $W_1 \neq H$ . Dann gibt es eine Lösung  $u$  der Aufgabe (9.9).*

BEWEIS : Wegen Lemma 9.8 nehmen wir o.B.d.A. an, dass

$$\sigma := \sup \{ |(Ku, u)| \mid \|u\| = 1, u \perp W_1 \} > 0.$$

Andernfalls ist  $W_1^\perp$  Eigenraum zum Eigenwert Null. Sei  $(u_m)$  eine Maximalfolge, d.h.  $\|u_m\| = 1$ ,  $u_m \perp W_1$ ,  $(Ku_m, u_m) \rightarrow \sigma$ . Wegen der schwachen Kompaktheit beschränkter Mengen im Hilbertraum gibt es eine Teilfolge  $(u_m)_{m \in \Lambda}$ , so dass

$$u_m \rightharpoonup u \quad \text{schwach in } H, \quad (m \rightarrow \infty, m \in \Lambda).$$

Wegen der Kompaktheit von  $K$  gilt  $Ku_m \rightarrow Ku$  und  $(Ku_m, u_m) \rightarrow (Ku, u)$ . Ferner gilt  $u \perp W_1$  und

$$\|u\|^2 = (u, u_m) + o(1) \leq \|u\| \|u_m\| + o(1) = \|u\| + o(1),$$

woraus  $\|u\| \leq 1$  folgt. Es muss  $u \neq 0$  gelten, da  $\sigma > 0$ , und es muss  $\|u\| = 1$  gelten, da einerseits  $\sigma = (Ku, u)$ , andererseits wegen der Maximaleigenschaft von  $\sigma$

$$\left( K \frac{u}{\|u\|}, \frac{u}{\|u\|} \right) \leq \sigma,$$

somit  $\sigma \leq \sigma \|u\|^2$ , woraus  $\|u\| \geq 1$  folgt. Der obige schwache Limes  $u$  ist somit Lösung von (9.9). ■

**9.11 Lemma.** *Jede Lösung  $u \in H$  von (9.9),  $H \neq W_1$ , ist Eigenvektor zum Eigenwert  $\lambda_2 = (Ku, u)$ . Es gilt  $|\lambda_1| > |\lambda_2|$  und es gibt keine weiteren Eigenwerte  $\lambda$  von  $K$  mit  $|\lambda_1| > |\lambda| > |\lambda_2|$ .*

BEWEIS : Sei  $u$  Lösung von (9.9). Es gilt  $\|u\| = 1$  und  $u \perp W_1$ . Sei  $v \in H$  mit  $\|v\| = 1$ ,  $v \perp u$  und  $v \perp W_1$ . Wir definieren wieder die „Vergleichsfunktion“

$$w(t) = u \cos t + v \sin t.$$

Sei o.B.d.A.

$$\sigma = \sup \{ (Kz, z) \mid \|z\| = 1, z \perp W_1 \} > 0.$$

( $\sigma = 0$  ist nicht möglich wegen Lemma 9.8.) Wegen der Maximaleigenschaft von  $u$  gilt

$$\operatorname{Re}(Ku, u) \geq (Kw(t), w(t)),$$

und wie beim Beweis von Satz 9.5 schließen wir wieder

$$\operatorname{Re}(v, Ku) = 0 \quad \text{für alle } v \perp u, v \perp W_1. \quad (9.12)$$

(Zweimalige) Anwendung des Projektionssatzes ergibt

$$Ku = \lambda_2 u + w + v, \quad w \in W_1, v \perp u, v \perp W_1. \quad (9.13)$$

Hieraus und aus (9.12) erhalten wir durch skalare Multiplikation mit  $v$ , dass  $\|v\|^2 = 0$  und somit  $v = 0$  ist. Skalare Multiplikation von (9.13) mit  $w$  ergibt

$$0 = (u, \lambda_1 w) = (u, Kw) = (Ku, w) = (\lambda_2 u, w) + \|w\|^2 = \|w\|^2, \quad \text{da } u \perp w.$$

Somit ist  $w = 0$  und  $w \in W_1$

$$Ku = \lambda_2 u.$$

$u$  ist also Eigenvektor. Es ist klar, dass  $\sigma = |\lambda_2| < |\lambda_1|$  ist, da im Falle  $\sigma = |\lambda_1|$  gelten würde  $u \perp W_1, u \in W_1$  (nach Definition) und somit  $u = 0$ , was ein Widerspruch zu  $\|u\| = 1$  ist. Wir zeigen noch, dass es keinen Eigenwert mit Betrag zwischen  $|\lambda_1|$  und  $|\lambda_2|$  geben kann:

Sei  $|\lambda_1| > |\lambda| > |\lambda_2|$  und  $\lambda$  Eigenwert zum Eigenraum  $V$ . Da Eigenräume selbstadjungierter Operatoren zu verschiedenen Eigenwerten zueinander orthogonal sind, gilt  $V \perp W_1$  und für  $v \in V, \|v\| = 1$

$$|\lambda| = |(Kv, v)| \leq \sup \{ |(Kz, z)| \mid z \perp W_1, \|z\| = 1 \} = |\lambda_2|.$$

Dies ist ein Widerspruch. ■

Konstruktion der Eigenräume  $V_N^+$  und  $V_N^-$  zum Eigenwert  $\lambda_N > 0$  bzw.  $-\lambda_N$ :  
Die Eigenräume  $V_j^+$  und  $V_j^-$  zum Eigenwert  $\lambda_j > 0$  bzw.  $-\lambda_j$  seien schon konstruiert für  $j = 1, \dots, N-1$ . Es gelte

$$\begin{aligned} |\lambda_1| &> |\lambda_2| > \dots > |\lambda_{N-1}| > 0, \\ |\lambda_1| &= \sup \{ |(Ku, u)| \mid \|u\| = 1, u \in H \}, \end{aligned}$$

und es existiere kein Eigenwert mit  $|\lambda_{j-1}| > |\lambda| > |\lambda_j|, j = 1, \dots, N-1$ . Wir konstruieren  $V_N^+$  und  $V_N^-$  durch die folgende Optimierungsaufgabe:

Gesucht ist  $u \in H, \|u\| = 1, u \perp W_j = V_j^+ \oplus V_j^-, j = 1, \dots, N-1$ , so dass

$$|(Ku, u)| = \sigma_N := \sup \{ |(Kz, z)| \mid \|z\| = 1, z \perp W_j, j = 1, \dots, N-1 \}. \quad (9.14)$$

**9.15 Lemma.** Sei  $H \neq \bigoplus_{j=1}^{N-1} W_j$ . Unter den Voraussetzungen von Lemma 9.10 besitzt die Aufgabe (9.14) eine Lösung  $u$ .

Der Beweis geschieht ähnlich wie in Lemma 9.10 und wird daher hier nicht ausgeführt.

**9.16 Lemma.** Jede Lösung von (9.14) ist Eigenvektor zum Eigenwert  $\lambda_N$ . Es gilt  $|\lambda_{N-1}| > |\lambda_N|$  und es gibt keinen Eigenwert  $\lambda$  von  $K$  mit  $|\lambda_{N-1}| > |\lambda| > |\lambda_N|$ .

BEWEIS : Das Element  $Ku$  von (9.14) hat nach dem Projektionssatz die Darstellung

$$Ku = \pm \lambda_N u + \sum_{j=1}^{N-1} w_j + v, \quad w_j \in W_j \quad (9.17)$$

mit  $v \perp w_j, j = 1, \dots, N-1, v \perp u$ . Die Orthogonalität  $u \perp W_j, j = 1, \dots, N-1$  gilt nach Konstruktion ohnehin. Wir zeigen, dass die Elemente  $w_j$  und  $v$  in (9.17)

verschwinden, so dass die Eigenwertgleichung  $Ku = \pm\lambda_N u$  gilt. Dies folgt wieder durch skalare Multiplikation mit  $w_j$

$$0 = \pm\lambda_j(w_j, u) = \pm(Kw_j, u) = \pm(w_j, Ku) = \|w_j\|^2$$

bzw. mit  $v$

$$0 = \mathcal{R}e(v, Ku) = \|v\|^2,$$

wobei die Beziehung  $\mathcal{R}e(v, Ku) = 0$  aus der Beziehung

$$\frac{d}{dt}(Kw(t), w(t)) \Big|_{t=0} = 0, \quad w(t) := u \cos t + \frac{v}{\|v\|} \sin t$$

folgt. Ist schließlich  $\lambda$  ein Eigenwert mit  $|\lambda_{N-1}| > |\lambda| > |\lambda_N|$  zum Eigenraum  $V$ , so ist  $V \perp W_j$ ,  $j = 1, \dots, N-1$ , und ähnlich wie in Lemma 9.1 ist die Extremaleigenschaft von  $|\lambda_N|$  verletzt. ■

Durch die vorangegangene Konstruktion erhalten wir Eigenwerte  $\pm\lambda_1, \pm\lambda_2, \dots$  mit endlich-dimensionalen zugehörigen Eigenräumen. Das Verfahren bricht nach endlich vielen Schritten ab, wenn der Raum  $H$  endlich-dimensional ist oder wenn  $\lambda_N = 0$ . Wenn  $\dim H = \infty$  und  $\lambda_N = 0$ , liegt ein unendlich-dimensionaler Nullraum vor. Der Standardfall ist jedoch, dass das Verfahren nicht abbricht und wir abzählbar unendlich viele Eigenwerte  $\pm\lambda_j$ ,  $j \in \mathbb{N}$ , mit  $|\lambda_j| > |\lambda_{j-1}|$  erhalten. Ferner ist noch zusätzlich der Eigenwert Null möglich. Die Eigenwerte  $\lambda_j$  müssen wegen Lemma 9.4 gegen Null gehen.

**9.18 Satz.** *Sei  $K$  eine lineare, vollstetige, selbstadjungierte Abbildung eines Hilbertraumes  $H$  in sich und  $V_j^+, V_j^-$  die in Lemma 9.16 konstruierten Eigenräume von  $K$  zum Eigenwerte  $\pm\lambda_j$ ,  $j \in \mathbb{N}$ . Dann gilt*

$$H = N(K) \oplus H_0, \quad H_0 := \bigoplus_{j=1}^{\infty} (V_j^+ \oplus V_j^-).$$

Erläuterung:  $H_0$  besteht aus allen konvergenten Reihen  $\sum_{j=1}^{\infty} (\mu_j^+ v_j^+ + \mu_j^- v_j^-)$ ,  $v_j^{\pm} \in V_j^{\pm}$ ,  $\mu_j^{\pm} \in \mathbb{R}$  bzw.  $\mathbb{C}$ . Die Eigenfunktionen zu den von Null verschiedenen Eigenwerten bilden also ein vollständiges Orthonormalsystem in  $N(K)^{\perp}$ .

BEWEIS : Angenommen, dass  $N(K) \oplus H_0 \neq H$ . Dann betrachten wir die Optimierungsaufgabe

$$|(Ku, u)| = \max!, \quad \|u\| = 1, \quad u \perp N(K) \oplus H_0. \quad (9.19)$$

Mit den schon bekannten Methoden beweist man die Existenz einer Lösung  $u$ . Wir behaupten, dass  $u$  Eigenvektor von  $K$  ist: O.B.d.A. sei  $(Ku, u) > 0$ . Man setze

$$w(t) = \cos t u + \sin t w, \quad \|w\| = 1, \quad w \perp u, \quad w \perp N(K) \oplus H_0.$$



(Falls es ein solches  $w$  nicht gibt, gilt  $Ku = \lambda u + v$ ,  $v \perp N(K) \oplus H_0$  und man springe zu (9.22) mit  $w = 0$ .) Es gilt

$$\frac{d}{dt}(Kw(t), w(t)) \Big|_{t=0} = 0, \quad (9.20)$$

da  $w(t)$  zur Konkurrenz in (9.19) zugelassen ist. Aus (9.20) folgt

$$\operatorname{Re}(Ku, w) = 0 \quad \text{für } w \perp u, w \perp N(K) \oplus H_0. \quad (9.21)$$

Allgemein gilt wegen des Projektionssatzes

$$Ku = \lambda u + w + v, \quad w \perp u, v \perp u, v \in N(K) \oplus H_0, w \perp v. \quad (9.22)$$

Durch skalare Multiplikation mit  $w$  ergibt sich

$$\|w\|^2 = \operatorname{Re}(Ku, w) = 0, \quad \text{also } Ku = \lambda u + v,$$

und durch skalare Multiplikation mit  $v$

$$\|v\|^2 = \operatorname{Re}(Ku, v) = \operatorname{Re}(u, Kv).$$

Da  $K$  den Raum  $N(K) \oplus H_0$  in sich selbst abbildet, folgt  $(u, Kv) = 0$  und  $\|v\|^2 = 0$ . (Warum bildet  $K$  den Raum  $N(K) \oplus H_0$  in sich selbst ab? Ist  $z \in N(K) \oplus H_0$ , so gilt

$$z = z_0 + \sum_{j=1}^{\infty} (\zeta_j^+ v_j^+ + \zeta_j^- v_j^-), \quad v_j^\pm \in V_j^\pm, \quad \|v_j^\pm\| = 1$$

mit  $\sum_{j=1}^{\infty} (|\zeta_j^+|^2 + |\zeta_j^-|^2) < \infty$ . Da  $K$  insbesondere stetig, folgt  $Kz = \sum_{j=1}^{\infty} (\zeta_j^+ \lambda_j^+ v_j^+ + \zeta_j^- \lambda_j^- v_j^-)$ . Beachte, dass  $|\lambda_j^\pm| \leq \|K\|$ . Die letzte unendliche Summe konvergiert somit in  $H$ , und damit gilt  $Kz \in H_0$ .)

Aus (9.22) folgt damit die Eigenwertgleichung

$$Ku = \lambda u, \quad u \neq 0.$$

Da  $u \notin N(K)$ , gilt  $|\lambda| > 0$ , und  $|\lambda|$  muss die Ungleichung  $\lambda_j > |\lambda| > \lambda_{j-1}$  für einen Index  $j$  erfüllen, da  $|\lambda_j| \rightarrow 0$ . Dies ist aufgrund der Konstruktion des  $\lambda_j$  nicht möglich. ■

Insbesondere haben wir gezeigt, dass nur abzählbar viele Eigenwerte existieren können.

### Anwendungen auf Rand-Eigenwertprobleme für elliptische Differentialoperatoren der Gestalt

$$-\sum_{i,k=1}^n D_i(a_{ik}D_k u) + c u$$

Sei  $\Omega$  ein beschränktes Gebiet des  $\mathbb{R}^n$ ,  $a_{ik} \in L^\infty(\Omega)$ ,  $a_{ik} = a_{ki}$ ,  $c \in L^\infty(\Omega)$ . Es gelte die Elliptizitätsbedingung

$$\sum_{i,k=1}^n a_{ik}(x) \xi_i \xi_k \geq \alpha_0 |\xi|^2, \quad \xi \in \mathbb{R}^n \quad (9.23)$$

mit einer Konstanten  $\alpha_0 > 0$ .

Wir betrachten das Eigenwertproblem

$$\begin{aligned} - \sum_{i,k=1}^n D_i(a_{ik}D_k u) + c u &= \lambda u, & u &\neq 0 \\ u|_{\partial\Omega} &= 0 \end{aligned} \quad (9.24)$$

In der schwachen Formulierung lautet dies: Gesucht ist  $u \in H_0^{1,2}(\Omega)$  mit

$$(Au, \varphi)_{H^{1,2}} := \sum_{i,k=1}^n (a_{ik}D_k u, D_i \varphi)_{L^2} + (c u, \varphi)_{L^2} = \lambda(u, \varphi)_{L^2}$$

für alle  $\varphi \in H_0^{1,2}(\Omega)$ .

Wir könnten auch natürliche Randbedingungen behandeln, indem wir  $H^{1,2}(\Omega)$  als Grundraum wählen. Hierbei bräuchte man jedoch die Zusatzvoraussetzung, dass die Einbettung von  $H^{1,2}(\Omega)$  als Grundraum nach  $L^2(\Omega)$  kompakt ist. (Aus diesem Grunde benötigen wir auch die Voraussetzung, dass  $\Omega$  beschränkt ist.)

Wir führen noch die mit der Einbettung von  $H_0^{1,2}$  nach  $L^2$  zusammenhängende Abbildung  $J : H_0^{1,2} \rightarrow H_0^{1,2}$  ein:

$$(Ju, v)_{H^{1,2}} := (u, v)_{L^2}$$

Mit diesen Bezeichnungen führt die schwache Formulierung des Eigenwertproblems (9.24) auf die Gleichung

$$Au = \lambda Ju. \quad (9.25)$$

Hierbei ist  $A : H_0^{1,2} \rightarrow H_0^{1,2}$  eine lineare Abbildung, welche der Gårdingschen Ungleichung (Satz 8.21)

$$(Au, u)_{H^{1,2}} \geq \alpha \|u\|_{H^{1,2}}^2 - \lambda_0 \|u\|_{L^2}^2$$

mit Konstanten  $\alpha, \lambda_0 > 0$  genügt. Dies schreiben wir in der Form

$$((A + \lambda_0 J)u, u)_{H^{1,2}} \geq \alpha \|u\|_{H^{1,2}}^2. \quad (9.26)$$

Das Eigenwertproblem (9.25) formulieren wir äquivalent um:

$$(A + \lambda_0 J)u = (\lambda + \lambda_0)Ju,$$

Wir setzen

$$B = A + \lambda_0 J, \quad \mu = \lambda + \lambda_0.$$

Damit lautet das Eigenwertproblem

$$Bu = \mu Ju. \quad (9.27)$$

Sei  $C = B^{\frac{1}{2}}$  der eindeutig bestimmte lineare, beschränkte, selbstadjungierte positive Operator mit

$$C^2 = B.$$

Beachte, dass  $(u, Bu) \geq \alpha \|u\|^2$  für alle  $u \in H$  mit einem  $\alpha > 0$ , d.h. „ $B$  ist positiv“. Die Existenz einer solchen Abbildung  $B^{\frac{1}{2}} = C$  wurde auf dem Übungsblatt 9 gezeigt. Mit den Bezeichnungen

$$v = B^{\frac{1}{2}}u, \quad B^{\frac{1}{2}}v = \mu JB^{-\frac{1}{2}}v,$$

formen wir (9.27) um in

$$\frac{1}{\mu}v = B^{-\frac{1}{2}}JB^{-\frac{1}{2}}v. \quad (9.28)$$

Die Abbildung  $B^{-\frac{1}{2}}$  ist linear, beschränkt und selbstadjungiert, die Abbildung  $J$  ebenfalls; überdies ist  $J$  kompakt. Damit ist die Abbildung

$$K = B^{-\frac{1}{2}}JB^{-\frac{1}{2}}$$

linear, vollstetig und selbstadjungiert, und wir können die in Paragraph 2.9 entwickelte Theorie anwenden. Zur Vereinfachung ist es nützlich zu wissen, dass die Kerne  $N(K)$  und  $N(J)$  verschwinden.

### 9.29 Lemma.

$$N(K) = 0, \quad N(J) = 0.$$

BEWEIS : Es ist  $(Ju, v)_{H^{1,2}} = (u, v)_{L^2}$ . Gilt also  $u \in N(J)$ , so folgt

$$(u, v)_{L^2} = 0 \quad \text{für alle } v \in H_0^{1,2}(\Omega).$$

Daraus folgt sofort  $u = 0$ , da  $C_0^\infty(\Omega)$  dicht in  $H_0^{1,2}(\Omega)$  ist. Da  $B^{-\frac{1}{2}}$  invertierbar, folgt auch  $N(K) = 0$ . ■

Wir betrachten somit das Eigenwertproblem

$$Ku = \frac{1}{\mu}u \quad \text{in } H^{1,2},$$

$N(K) = 0$ ,  $K = K^*$ ,  $K$  kompakt. Aus Lemma 9.29 sowie Satz 9.5 und Satz 9.18 erhalten wir

**9.30 Satz.** *Das elliptische Rand-Eigenwertproblem (9.24) besitzt - unter den an  $a_{ik}$  und  $c$  genannten Voraussetzungen - abzählbar-unendlich viele Eigenwerte  $\nu_j$ . Die Eigenwerte sind reell und gehen gegen unendlich. Die Vielfachheit der Eigenräume  $V_j$  ist endlich, Eigenräume zu verschiedenen Eigenwerten sind orthogonal. Die Gesamtheit der Eigenvektoren ist vollständig in  $H_0^{1,2}(\Omega)$ .*

BEWEIS : Die Zahl  $\nu$  ist genau dann Eigenwert des Problems (9.24), wenn die Zahl  $\frac{1}{\nu + \lambda_0}$  Eigenwert von  $K = B^{-\frac{1}{2}}JB^{-\frac{1}{2}}$  ist. Da die Eigenwerte von  $K$  gegen Null gehen, d.h.  $\frac{1}{\nu_j + \lambda_0} \rightarrow 0$ , folgt  $|\nu_j| \rightarrow \infty$ , und da  $(Ku, u) \geq 0$ , gilt  $\nu_j \rightarrow \infty$  ( $j \rightarrow \infty$ ) für die Eigenwerte  $\nu_j$  von (9.24). ■

Anmerkung: Selbstverständlich lässt sich ein analoger Satz für  $H^{1,2}(\Omega)$  als Grundraum - dies entspricht natürlichen Randbedingungen - oder auch elliptische Operatoren höherer Ordnung formulieren.



# Kapitel 3

## Allgemeine lineare Funktionalanalysis

### 3.1 Grundbegriffe der Funktionalanalysis

#### Metrische Räume

**1.1 Definition.** Ein **metrischer Raum** ist eine Menge von Punkten, auf der eine **Abstandsfunktion**  $\rho : V \times V \rightarrow \mathbb{R}$  gegeben ist, die die folgenden Eigenschaften für alle  $x, y, z \in V$  besitzt:

(i) *positive Definitheit:*  $\rho(x, y) \geq 0$ ,  $\rho(x, x) = 0$ , und falls  $\rho(x, y) = 0$ , dann folgt  $x = y$

(ii) *Symmetrie:*  $\rho(x, y) = \rho(y, x)$

(iii) *Dreiecksungleichung:*  $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$

Man sagt auch „ $V$  wird durch  $\rho$  metrisiert“ und nennt  $\rho$  eine **Metrik**. Im metrischen Raum lässt sich ein Konvergenzbegriff einführen:

**1.2 Definition.** Eine **Folge**  $(u_m)_{m \in \mathbb{N}}$ ,  $u_m \in V$  **konvergiert** genau dann gegen ein Element  $u$ , wenn

$$\rho(u_m, u) \rightarrow 0 \quad (m \rightarrow \infty).$$

Man schreibt  $u_m \rightarrow u$  in  $V$  oder bezüglich  $\rho$  ( $m \rightarrow \infty$ ). Ein metrischer Raum  $V$  heisst **vollständig**, wenn jede Cauchyfolge  $(u_m)$  in  $V$  einen Limes bezüglich der Metrik hat, d.h. gibt es zu jedem  $\varepsilon > 0$  ein  $n_0 \in \mathbb{N}$  mit

$$\rho(u_m, u_k) < \varepsilon \quad \text{für } m, k \geq n_0,$$

so gibt es ein  $u \in V$  mit  $u_m \rightarrow u$  ( $m \rightarrow \infty$ ).

#### Beispiele metrischer Räume:

a)  $\mathbb{R}^n$  mit  $\rho(x, y) = |x - y|$  - kennt jeder!

b) Der Folgenraum  $l = \{(c_1, c_2, \dots) \mid c_j \in \mathbb{R} \text{ bzw. } \mathbb{C}\}$  mit der Metrik

$$\rho(a, b) = \sum_{j=1}^{\infty} \frac{1}{2^j} \frac{|a_j - b_j|}{1 + |a_j - b_j|}. \quad (1.3)$$

Wir zeigen, dass  $\rho$  in der Tat die Eigenschaften einer Metrik besitzt. Die positive Definitheit und die Symmetrie sind offensichtlich. Die Dreiecksungleichung ergibt sich folgendermaßen:

$$\begin{aligned} \rho(a, b) &= \sum_{j=1}^{\infty} \frac{1}{2^j} \left(1 - \frac{1}{1 + |a_j - b_j|}\right) \leq \sum_{j=1}^{\infty} \frac{1}{2^j} \left(1 - \frac{1}{1 + |a_j - c_j| + |c_j - b_j|}\right) \\ &= \sum_{j=1}^{\infty} \frac{1}{2^j} \left(\frac{|a_j - c_j| + |c_j - b_j|}{1 + |a_j - c_j| + |c_j - b_j|}\right) \leq \rho(a, c) + \rho(c, b). \end{aligned}$$

Man überlegt sich leicht:

**1.4 Lemma.**  $c^{(k)} \rightarrow c$ ,  $c^{(k)}, c \in l$ , konvergiert genau dann bezüglich der Metrik  $\rho$  in (1.3), wenn die Komponenten  $c_j^{(k)}$  von  $c^{(k)}$  gegen die Komponenten  $c_j$  von  $c$  konvergieren.

c) Der Raum der messbaren Funktionen

Sei  $\Omega$  eine messbare beschränkte Menge. Wir betrachten den Raum

$$\mathcal{M} = \{f : \Omega \rightarrow \mathbb{R} \text{ bzw. } \mathbb{C} \mid f \text{ messbar}\}.$$

Es sei darauf hingewiesen, dass nach Paragraph 1.2 nicht zwischen der Funktion  $f$  und seiner Äquivalenzklasse  $[f]$  bzgl. fast überall Gleichheit unterschieden wird. Als Metrik in  $\mathcal{M}$  definieren wir

$$\rho(f, g) = \int_{\Omega} \frac{|f - g|}{1 + |f - g|} dx.$$

Man prüft leicht nach, dass  $\rho$  eine Metrik in  $\mathcal{M}$  definiert. Die Dreiecksungleichung beweist man wie in Beispiel b).

**1.5 Satz.** Die Räume  $\mathbb{R}^n, l, \mathcal{M}$  sind bezüglich der angegebenen Metriken vollständig.

BEWEIS : Für den  $\mathbb{R}^n$  weiß dies jeder. Für den Raum  $l$  schließt man aus  $\rho(a^{(m)}, a^{(k)}) \rightarrow 0$  ( $m, k \rightarrow \infty$ ), dass die Komponenten  $a_j^{(m)}$  Cauchy-Folgen in  $\mathbb{R}$  bzw.  $\mathbb{C}$  bilden. Es gibt daher ein  $a \in l$ , so dass  $a^{(m)} \rightarrow a$  *komponentenweise*. Daraus folgt

$$\rho(a^{(m)}, a) \leq \sum_{j=1}^N \frac{1}{2^j} \frac{|a_j^{(m)} - a_j|}{1 + |a_j^{(m)} - a_j|} + \sum_{j=N+1}^{\infty} \frac{1}{2^j} = o(1) + \frac{1}{2^N} \rightarrow 0 \text{ für } N \rightarrow \infty,$$

was zu zeigen war.

Für den Raum  $\mathcal{M}$  betrachten wir eine Folge  $f_j$  messbarer Funktionen mit

$$\int_{\Omega} \frac{|f_j - f_k|}{1 + |f_j - f_k|} dx \rightarrow 0 \quad (j, k \rightarrow \infty),$$

d.h. die Folge  $\frac{|f_j - f_k|}{1 + |f_j - f_k|}$  konvergiert stark in  $L^1(\Omega)$ , und somit gilt für eine Teilfolge  $f_j - f_k \rightarrow 0$  fast überall. Für fast alle  $x \in \Omega$  gibt es also einen Grenzwert  $f_j(x) \rightarrow f(x)$ . Die Funktion  $f(x)$  ist nach Kapitel 1 Satz 2.3 wieder messbar. Für  $j, k \geq n_0$  gilt

$$\rho(f_j, f_k) = \int_{\Omega} \frac{|f_j - f_k|}{1 + |f_j - f_k|} dx < \frac{\varepsilon}{2}.$$

Wir lassen den Index  $k$  die konstruierte Teilfolge durchlaufen,  $k \rightarrow \infty$ , und gehen nach dem Satz über dominierte Konvergenz (Kapitel 1 Satz 3.12) zum Limes über. Dies ergibt

$$\rho(f_j, f) \leq \frac{\varepsilon}{2} < \varepsilon.$$

Die Vollständigkeit ist damit bewiesen. ■

Ein Beispiel eines *unvollständigen* metrischen Raumes ist der Raum  $C(\Omega)$  der reell- oder komplexwertigen Funktionen auf  $\Omega$  mit der obigen Metrik des Raumes der messbaren Funktionen. Offensichtlich führen Cauchy-Folgen aus diesem Raum heraus.

In einem metrischen Raum lassen sich in Analogie zum  $\mathbb{R}^n$  die Begriffe „beschränkte Menge“, „offene Menge“, „abgeschlossene Menge“ und „Häufungspunkt einer Folge oder Menge“ definieren. Als  $\varepsilon$ -Umgebung eines Punktes  $x$  eines metrischen Raumes  $V$  definiert man naturgemäß

$$U_{\varepsilon}(x) = \{y \in V \mid \rho(x, y) < \varepsilon\}.$$

Die Begriffe „folgenkompakt“ und „überdeckungskompakt“ sind in metrischen Räumen äquivalent, obwohl der Raum unendlich-dimensional sein kann.

Ist der Raum unendlich-dimensional, so sind jedoch beschränkte, abgeschlossene Mengen nicht mehr folgenkompakt.

## Normierte Räume und Banachräume

**1.6 Definition.** Ein *normierter Raum* ist ein Vektorraum  $V$  über  $\mathbb{R}$  bzw.  $\mathbb{C}$  mit einer **Abbildung**  $\|\cdot\| : V \rightarrow \mathbb{R}$ , so dass die folgenden Eigenschaften für alle  $u, v \in V$  erfüllt sind:

(i) *Positive Definitheit:*  $\|u\| \geq 0$ ,  $\|u\| = 0$  gilt genau dann, wenn  $u = 0$ .

(ii) *Positive Homogenität:*  $\|\lambda u\| = |\lambda| \|u\|$ ,  $\lambda \in \mathbb{R}$  bzw.  $\mathbb{C}$ .

(iii) *Dreiecksungleichung:*  $\|u + v\| \leq \|u\| + \|v\|$ .

Setzt man  $\rho(u, v) = \|u - v\|$ , so wird auf  $V$  eine Metrik definiert. Umgekehrt lässt sich leicht beweisen

**1.7 Lemma.** Sei  $V$  ein metrischer Raum mit einer translation- und skalierungsinvarianten Metrik, d.h.

$$\rho(u+w, v+w) = \rho(u, v), \quad \rho(\lambda u, \lambda v) = |\lambda| \rho(u, v).$$

Dann wird durch  $\|u\| := \rho(u, 0)$  eine Norm auf  $V$  definiert.

Da jede Norm auch eine Metrik erklärt, lassen sich die bei den metrischen Räumen definierten Begriffe wie Konvergenz, Vollständigkeit, Umgebungen, Offenheit und Abgeschlossenheit einer Teilmenge des Raumes verwenden. Insbesondere definiert man die Konvergenz

$$u_m \rightarrow u \text{ in } V, (m \rightarrow \infty) \Leftrightarrow \|u_m - u\| \rightarrow 0.$$

**1.8 Definition.** Ein **Banach-Raum** ist ein vollständiger normierter Raum.

Hilbert- und Banachräume sind die wichtigsten Raumtypen der Funktionalanalysis.

**Beispiele:**

a) Der Euklidische Raum  $\mathbb{R}^n$  bzw.  $\mathbb{C}^n$  mit  $\|a\| = \left(\sum_{j=1}^n |a_j|^2\right)^{\frac{1}{2}}$  als Norm.

b) Der Folgenraum  $l^p = \{a = (a_1, a_2, \dots) \mid a_j \in \mathbb{R} \text{ bzw. } \mathbb{C}, \|a\|_p < \infty\}$ ,  $1 \leq p < \infty$ , mit der Norm

$$\|a\|_p = \left(\sum_{j=1}^{\infty} |a_j|^p\right)^{\frac{1}{p}}.$$

Die Dreiecksungleichung - zunächst für endliche Summen - folgt zum Beispiel aus der Darstellung

$$\|a\|_p = \sup \left\{ \sum_{j=1}^N a_j \bar{b}_j \mid \sum_{j=1}^N |b_j|^{\frac{p}{p-1}} = 1 \right\}.$$

Dies folgert man seinerseits über die Theorie der Extremwerte mit Nebenbedingungen.

Im Fall  $p = \infty$  setzt man für  $a \in l^\infty$

$$\|a\|_\infty = \sup_j |a_j|.$$

Im Fall  $0 < p < 1$  definiert  $\|a\|_p$  keine Norm (wohl aber noch eine mit den algebraischen Operationen verträgliche Topologie).

c) Die Lebesgueschen Funktionenräume  $L^p$ ,  $1 \leq p \leq \infty$ .

Es sei  $\Omega$  eine messbare Punktmenge des  $\mathbb{R}^n$  und  $\mathcal{L}^p = \mathcal{L}^p(\Omega)$  die Menge aller messbaren Funktionen  $f : \Omega \rightarrow \mathbb{R}$  bzw.  $\mathbb{C}$  (oder  $\mathbb{R}^n$  bzw.  $\mathbb{C}^n$ ) mit (vgl. Abschnitt 1.3)

$$\|f\|_p := \left( \int_{\Omega} |f|^p dx \right)^{\frac{1}{p}} < \infty$$



für  $1 \leq p < \infty$  und

$$\|f\|_\infty := \text{ess sup } |f| < \infty.$$

In Abschnitt 1.3 haben wir gezeigt, dass die so definierte Größe  $\|\cdot\|_p$ ,  $1 \leq p \leq \infty$ , alle Eigenschaften einer Norm, mit Ausnahme der Definitheit erfüllt. Diese wird durch den Übergang zum Restklassenraum erreicht. Man betrachtet also Restklassen

$$[f] \equiv \{g \in \mathcal{L}^p \mid g = f \text{ fast überall}\}$$

und definiert

$$L^p = L^p(\Omega) \equiv \{[f] \mid f \in \mathcal{L}^p\}.$$

Wie bereits in Abschnitt 1.3 erläutert, wird nicht zwischen  $f$  und  $[f]$  unterschieden. Auf Übungsblatt 3, Aufgabe 2 wurde gezeigt, dass  $L^p$ ,  $1 \leq p \leq \infty$  ein Banachraum ist.

d) Der Raum der stetigen Funktionen

Es sei  $M \subset \mathbb{R}^n$  bzw.  $\mathbb{C}^n$  und

$$C(\overline{M}) = \{f : M \rightarrow \mathbb{R} \text{ bzw. } \mathbb{C} \mid f \text{ ist gleichmäßig stetig und beschränkt auf } M\}.$$

Mit der Norm

$$\|f\|_\infty = \sup \{|f(x)| \mid x \in M\}$$

ist  $C(\overline{M})$  ein Banachraum.

Es herrscht eine Verwechslungsmöglichkeit in der Bezeichnung  $\|f\|_\infty$ , wenn  $M$  das Mass Null hat - das wesentliche Supremum würde diese Menge „vergessen“. Ist  $M$  offen oder der Abschluß einer offenen Menge, so ist für stetige Funktionen  $\text{ess sup } |f| = \|f\|_\infty$ . (Ist  $M$  kompakt, wird das Supremum sogar angenommen.) Da  $C(\overline{M})$  ein Banachraum ist, ist in dem Fall, dass  $\text{ess sup}_M |f| = \sup_M |f|$  gilt, der Raum  $C(\overline{M})$  ein abgeschlossener Teilraum von  $L^\infty(M)$ .

Ist  $\Omega$  offen, so bezeichnet man üblicherweise  $C(\Omega)$  als den Raum der stetigen Funktionen über  $\Omega$ . Da für  $f \in C(\Omega)$  der Fall eintreten kann, dass  $f(x) \rightarrow \infty$  für  $x \in \Omega$ ,  $x \rightarrow x_0 \in \partial\Omega$ , ist  $C(\Omega)$  kein Banachraum bezüglich  $\|\cdot\|_\infty$ .

e) Der Raum der Hölderstetigen Funktionen  $C^\alpha$ ,  $0 < \alpha \leq 1$

Sei  $M \subset \mathbb{R}^n$ . Die **Hölderseminorm** einer Funktion  $f : M \rightarrow \mathbb{R}$  bzw.  $\mathbb{C}$  ist definiert durch

$$[f]_\alpha := \sup \left\{ \frac{|f(x) - f(y)|}{|x - y|^\alpha} \mid x, y \in M, x \neq y \right\},$$

der Raum der Hölderstetigen Funktionen auf  $M$  ist definiert durch

$$C^\alpha(\overline{M}) = \{f : M \rightarrow \mathbb{R} \mid [f]_\alpha < \infty\}.$$

Mit der Norm

$$\|f\|_\alpha = \|f\|_\infty + [f]_\alpha$$

ist  $C^\alpha$  ein Banachraum. Hölderstetige Funktionen zum Hölderexponenten  $\alpha = 1$  sind genau die (global) Lipschitzstetigen Funktionen. Man verwendet für diesen Raum nicht das Symbol  $C^1$ , sondern  $C^{0,1}$  (oder auch  $H^{1,\infty}$ ).

f) Die Räume  $C^m$  und  $C^{m,\alpha}$ ,  $m \in \mathbb{N}$ ,  $0 < \alpha \leq 1$ .

Sei  $\Omega$  eine offene Punktmenge des  $\mathbb{R}^n$ . Dann bezeichnet  $C^m(\overline{\Omega})$  bzw.  $C^{m,\alpha}(\overline{\Omega})$  die Menge aller Funktionen  $f : \Omega \rightarrow \mathbb{R}$  bzw.  $C$ , die in  $\Omega$   $m$ -mal differenzierbar sind und deren  $m$ -te Ableitung gleichmäßig stetig bzw. Hölderstetig ist. Bezüglich der Norm

$$\|f\|_{\infty,m} = \sum_{j=0}^m \|\nabla^j f\|_{\infty}$$

bzw.

$$\|f\|_{\infty,m+\alpha} = \|f\|_{\infty,m} + [\nabla^m f]_{\alpha}$$

sind  $C^m$  und  $C^{m,\alpha}$  Banachräume.

g) Die Sobolevräume  $H^{m,p}(\Omega)$ ,  $m \in \mathbb{N}$ ,  $1 \leq p \leq \infty$ .

Es sei  $\Omega$  eine offene Punktmenge. Mit  $H^{1,p}(\Omega)$  bezeichnen wir den linearen Raum aller Funktionen aus  $L^p(\Omega)$ , die erste verallgemeinerte Ableitungen in  $L^p(\Omega)$  haben, d.h. es existieren Funktionen  $u^{(i)} \in L^p(\Omega)$  mit

$$\int_{\Omega} u \frac{\partial \varphi}{\partial x_i} dx = - \int_{\Omega} u^{(i)} \varphi dx$$

für alle  $\varphi \in C_0^{\infty}(\Omega)$ . Mit  $H^{m,p}(\Omega)$  bezeichnen wir die Menge aller Funktionen aus  $L^p(\Omega)$ , die  $k$ -te verallgemeinerte Ableitungen,  $k = 1, \dots, m$ , aus  $L^p(\Omega)$  haben.  $H^{m,p}(\Omega)$  ist mit der Norm

$$\|u\|_{m,p} = \sum_{j=0}^m \|\nabla^j u\|_p$$

ein Banachraum. Der Banachraum  $H_0^{m,p}(\Omega)$  ist definiert als der Abschluß von  $C_0^{\infty}(\Omega)$  bzgl. der  $\|\cdot\|_{m,p}$ -Norm. Im Falle eines genügend glatten Randes  $\partial\Omega$  kann man zeigen, dass

$$H_0^{m,p}(\Omega) = \{u \in H^{m,p}(\Omega) \mid \nabla^k u|_{\partial\Omega} = 0, k = 1, \dots, m-1\},$$

d.h. alle verallgemeinerten Ableitungen bis zur Ordnung  $m-1$  verschwinden auf dem Rand  $\partial\Omega$ .

In Anwendungen werden oft sogenannte *Sobolev-Ungleichungen* benötigt. Dies sind Ungleichungen der Form

$$\|u\|_q \leq c \|\nabla u\|_p, \quad (1.9)$$

wobei die Konstante  $c$  und die Zahl  $q$  nicht von  $u$  abhängen soll. Zuerst überlegen wir uns für welche  $q$  dies möglich ist im Falle von Funktionen  $u \in C_0^{\infty}(\mathbb{R}^n)$  und  $1 \leq p < n$ . Sei also  $u \in C_0^{\infty}(\mathbb{R}^n)$ ,  $u \not\equiv 0$  und sei  $\lambda > 0$ . Wir setzen

$$u_{\lambda}(x) := u(\lambda x), \quad x \in \mathbb{R}^n. \quad (1.10)$$

Man rechnet leicht nach

$$\begin{aligned} \int_{\mathbb{R}^n} |u_{\lambda}|^q dx &= \int_{\mathbb{R}^n} |u(\lambda x)|^q dx = \frac{1}{\lambda^n} \int_{\mathbb{R}^n} |u(y)|^q dy, \\ \int_{\mathbb{R}^n} |\nabla u_{\lambda}|^p dx &= \lambda^p \int_{\mathbb{R}^n} |\nabla u(\lambda x)|^p dx = \frac{\lambda^p}{\lambda^n} \int_{\mathbb{R}^n} |\nabla u(y)|^p dy. \end{aligned}$$

Dies eingesetzt in (1.9) liefert

$$\frac{1}{\lambda^{\frac{n}{q}}} \|u\|_q \leq c \frac{\lambda}{\lambda^{\frac{n}{p}}} \|\nabla u\|_p,$$

d.h.

$$\|u\|_q \leq c \lambda^{1 - \frac{n}{p} + \frac{n}{q}} \|\nabla u\|_p. \quad (1.11)$$

Wenn  $1 - \frac{n}{p} + \frac{n}{q} \neq 0$  ist, kann man  $\lambda$  gegen 0 bzw.  $\infty$  konvergieren lassen und erhält einen Widerspruch. Somit kann (1.9) nur gelten, wenn

$$\frac{1}{q} = \frac{1}{p} - \frac{1}{n}$$

gilt. Dies kann man auch alternativ schreiben als

$$q = \frac{np}{n-p}.$$

**1.12 Satz.** Sei  $1 \leq p < n$ . Dann existiert eine Konstante  $c$ , die nur von  $p$  und  $n$  abhängt, so dass

$$\|u\|_q \leq c \|\nabla u\|_p, \quad q = \frac{np}{n-p}, \quad (1.13)$$

für alle  $u \in C_0^1(\mathbb{R}^n)$  gilt.

Wir benötigen hier wirklich  $u \in C_0^1(\mathbb{R}^n)$ , wie  $u \equiv 1$  zeigt. Bemerkenswert ist, dass die Konstante  $c$  nicht von der Größe des Trägers von  $u$  abhängt.

BEWEIS : (i) Der Fall  $p = 1$ .

Da  $u$  kompakten Träger hat gilt

$$u(x) = \int_{-\infty}^{x_i} D_{x_i} u(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_n) dy_i$$

und somit

$$|u(x)| \leq \int_{-\infty}^{\infty} |\nabla u(x_1, \dots, y_i, \dots, x_n)| dy_i, \quad i = 1, \dots, n.$$

Daraus folgt

$$|u(x)|^{\frac{n}{n-1}} \leq \prod_{i=1}^n \left( \int_{-\infty}^{\infty} |\nabla u(x_1, \dots, y_i, \dots, x_n)| dy_i \right)^{\frac{1}{n-1}}.$$

Diese Ungleichung wird bezüglich  $x_1$  über  $\mathbb{R}$  integriert, was

$$\begin{aligned} \int_{-\infty}^{\infty} |u|^{\frac{n}{n-1}} dx_1 &\leq \int_{-\infty}^{\infty} \prod_{i=1}^n \left( \int_{-\infty}^{\infty} |\nabla u| dy_i \right)^{\frac{1}{n-1}} dx_1 \\ &= \left( \int_{-\infty}^{\infty} |\nabla u| dy_1 \right)^{\frac{1}{n-1}} \int_{-\infty}^{\infty} \prod_{i=2}^n \left( \int_{-\infty}^{\infty} |\nabla u| dy_i \right)^{\frac{1}{n-1}} dx_1 \\ &\leq \left( \int_{-\infty}^{\infty} |\nabla u| dy_1 \right)^{\frac{1}{n-1}} \left( \prod_{i=2}^n \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\nabla u| dx_1 dy_i \right)^{\frac{1}{n-1}} \end{aligned}$$

liefert. In der letzten Ungleichung wurde die verallgemeinerte Hölder-Ungleichung

$$\int_{\Omega} \prod_{i=1}^k |f_i| dx \leq \prod_{i=1}^k \|f_i\|_{p_i}, \quad \sum_{i=1}^k \frac{1}{p_i} = 1,$$

angewendet. Die obige Ungleichung wird nun bzgl.  $x_2$  über  $\mathbb{R}$  integriert. Wir erhalten

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |u|^{\frac{n}{n-1}} dx_1 dx_2 \leq \left( \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\nabla u| dx_1 dy_2 \right)^{\frac{1}{n-1}} \int_{-\infty}^{\infty} \prod_{\substack{i=1 \\ i \neq 2}}^n I_i^{\frac{1}{n-1}} dx_2,$$

wobei

$$I_1 = \int_{-\infty}^{\infty} |\nabla u| dy_1, \quad I_i = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\nabla u| dx_1 dy_i, \quad i = 3, \dots, n.$$

Die verallgemeinerte Hölder-Ungleichung liefert

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |u|^{\frac{n}{n-1}} dx_1 dx_2 &\leq \left( \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\nabla u| dx_1 dy_2 \right)^{\frac{1}{n-1}} \left( \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\nabla u| dy_1 dx_2 \right)^{\frac{1}{n-1}} \times \\ &\quad \times \prod_{i=3}^n \left( \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\nabla u| dx_1 dx_2 dy_i \right)^{\frac{1}{n-1}}. \end{aligned}$$

Weitere Integrationen bzgl.  $x_3, \dots, x_n$  liefern dann

$$\begin{aligned} \int_{\mathbb{R}^n} |u|^{\frac{n}{n-1}} dx &\leq \prod_{i=1}^n \left( \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} |\nabla u| dx_1 \dots dy_i \dots dx_n \right)^{\frac{1}{n-1}} \\ &= \left( \int_{\mathbb{R}^n} |\nabla u| dx \right)^{\frac{n}{n-1}}, \end{aligned} \tag{1.14}$$

was (1.13) für  $p = 1$  ist.

(ii) Der Fall  $1 < p < n$ .

Wir benutzen (1.14) für  $v = |u|^\gamma$ , mit  $\gamma > 1$  welches später gewählt wird, und erhalten

$$\begin{aligned} \left( \int_{\mathbb{R}^n} |u|^{\frac{\gamma n}{n-1}} dx \right)^{\frac{n-1}{n}} &\leq \int_{\mathbb{R}^n} |\nabla |u|^\gamma| dx \\ &= \int_{\mathbb{R}^n} \gamma |u|^{\gamma-1} |\nabla u| dx \\ &\leq \gamma \left( \int_{\mathbb{R}^n} |u|^{(\gamma-1)\frac{p}{p-1}} dx \right)^{\frac{p-1}{p}} \left( \int_{\mathbb{R}^n} |\nabla u|^p dx \right)^{\frac{1}{p}}. \end{aligned} \quad (1.15)$$

Wir wählen nun  $\gamma$  derart, dass

$$\frac{\gamma n}{n-1} = (\gamma-1) \frac{p}{p-1},$$

d.h.  $\gamma = p \frac{(n-1)}{n-p} > 1$ . Für dieses  $\gamma$  haben wir

$$\frac{\gamma n}{n-1} = (\gamma-1) \frac{p}{p-1} = \frac{np}{n-p},$$

was zusammen mit (1.15)

$$\|u\|_{\frac{np}{n-p}}^\gamma \leq \gamma \|u\|_{\frac{np}{n-p}}^{\gamma-1} \|\nabla u\|_p$$

liefert. Hieraus folgt sofort die Behauptung. ■

**1.16 Satz.** Sei  $\Omega \subset \mathbb{R}^n$ , ein beschränktes Gebiet mit  $\partial\Omega \in C^{0,1}$ . Dann existiert eine Konstante  $K = K(n, \partial\Omega)$  so, dass für alle  $u \in H^{1,p}(\Omega)$  gilt:

$$\|u\|_q \leq K \|u\|_{1,p}, \quad q = \frac{np}{n-p} \quad (1.17)$$

sofern  $1 \leq p < n$ . Ist  $u \in H_0^{1,p}(\Omega)$ , so gilt

$$\|u\|_q \leq K \|\nabla u\|_p, \quad q = \frac{np}{n-p}. \quad (1.18)$$

Da  $q > p$  ist, erhält man aufgrund der investierten Differenzierbarkeit eine höhere Integrabilität der Funktion.

BEWEIS : (i) Analog wie im Falle des Hilbertraumes  $H^{1,2}(\Omega)$  (vgl. Abschnitt 2.8) existiert ein Fortsetzungsoperator  $C : H^{1,p}(\Omega) \rightarrow H^{1,p}(\mathbb{R}^n)$  mit

$$\begin{aligned} (Cu)|_\Omega &= u, \quad \forall u \in H^{1,p}(\Omega), \\ Cu &\text{ hat kompakten Träger,} \\ \|Cu\|_{H^{1,p}(\mathbb{R}^n)} &\leq c \|u\|_{H^{1,p}(\Omega)}. \end{aligned} \quad (1.19)$$

Da  $Cu$  einen kompakten Träger hat liefert Kapitel 1, Satz 4.5, die Existenz von Funktionen  $u_m \in C_0^\infty(\mathbb{R}^n)$  mit

$$u_m \rightarrow Cu \quad \text{in } H^{1,p}(\mathbb{R}^n). \quad (1.20)$$

Satz 1.12 liefert

$$\|u_m - u_k\|_{\frac{pn}{n-p}} \leq c \|\nabla(u_m - u_k)\|_p \rightarrow 0,$$

woraus

$$u_m \rightarrow Cu \quad \text{in } L^{\frac{pn}{n-p}}(\mathbb{R}^n) \quad (1.21)$$

folgt. Satz 1.12 liefert auch

$$\|u_m\|_{L^{\frac{np}{n-p}}(\mathbb{R}^n)} \leq c \|\nabla u_m\|_{L^p(\mathbb{R}^n)}$$

was zusammen mit (1.20) und (1.21)

$$\|Cu\|_{L^{\frac{np}{n-p}}(\mathbb{R}^n)} \leq c \|\nabla(Cu)\|_{L^p(\mathbb{R}^n)}$$

ergibt. Dies und (1.19) liefern die Behauptung (1.17).

(ii) Sei  $u \in H_0^{1,p}(\Omega)$ . Dann existieren Funktionen  $u_m \in C_0^\infty(\Omega)$  mit  $u_m \rightarrow u$  in  $H^{1,p}(\Omega)$ . Wir setzen die Funktionen  $u_m$  durch Null auf ganz  $\mathbb{R}^n$  fort und erhalten aus Satz 1.12

$$\|u_m\|_{L^{\frac{np}{n-p}}(\mathbb{R}^n)} \leq c \|\nabla u_m\|_{L^p(\mathbb{R}^n)}$$

sowie  $u_m \rightarrow u$  in  $L^{\frac{np}{n-p}}(\mathbb{R}^n)$ . Der Grenzübergang  $m \rightarrow \infty$  liefert (1.18).  $\blacksquare$

Im Grenzfall  $p = n$  gelten zahlreiche Besonderheiten. Für den Fall  $p > n$  gilt

**1.22 Satz.** *Sei  $p > n$ . Unter den Voraussetzungen an  $\Omega$  wie in Satz 1.16 gilt*

$$H^{1,p}(\Omega) \subset C^\alpha(\bar{\Omega})$$

mit  $\alpha = 1 - \frac{n}{p}$ . Für die Normen gilt entsprechend

$$\|u\|_{C^\alpha} \leq K \|u\|_{H^{1,p}}, \quad \forall u \in H^{1,p}(\Omega). \quad (1.23)$$

BEWEIS : Wir zeigen zunächst das (1.23) für  $\Omega = \mathbb{R}^n$  und  $u \in C^1(\mathbb{R}^n)$  gilt.

(i) Sei  $B_r(x) \subseteq \mathbb{R}^n$  eine Kugel. Wir beweisen zuerst, dass

$$\int_{B_r(x)} |u(y) - u(x)| dy \leq c \int_{B_r(x)} \frac{|\nabla u(y)|}{|x-y|^{n-1}} dy, \quad (1.24)$$

mit einer Konstanten  $c$ , die nur von  $n$  abhängt gilt. Sei  $\omega \in \partial B_1(0)$  und  $0 < s < r$ . Es gilt

$$\begin{aligned} |u(x + s\omega) - u(x)| &= \left| \int_0^s \frac{d}{dt} u(x + t\omega) dt \right| = \left| \int_0^s \nabla u(x + t\omega) \cdot \omega dt \right| \\ &\leq \int_0^s |\nabla u(x + t\omega)| dt \end{aligned}$$

und somit

$$\int_{\partial B_1(0)} |u(x + s\omega) - u(x)| \, dS \leq \int_0^s \int_{\partial B_1(0)} |\nabla u(x + t\omega)| \frac{t^{n-1}}{t^{n-1}} \, dS \, dt.$$

Wir setzen  $y = x + t\omega$ , und somit  $t = |x - y|$ , und gehen von Polarkoordinaten zu kartesischen Koordinaten über. Wir erhalten

$$\int_{\partial B_1(0)} |u(x + s\omega) - u(x)| \, dS \leq \int_{B_s(x)} \frac{|\nabla u(y)|}{|x - y|^{n-1}} \, dy \leq \int_{B_r(x)} \frac{|\nabla u(y)|}{|x - y|^{n-1}} \, dy.$$

Diese Ungleichung wird mit  $s^{n-1}$  multipliziert und bzgl.  $s$  von 0 bis  $r$  integriert, welches

$$\int_{B_r(x)} |u(y) - u(x)| \, dy \leq \frac{r^n}{n} \int_{B_r(x)} \frac{|\nabla u(y)|}{|x - y|^{n-1}} \, dy$$

liefert. Dies ist aber (1.24).

(ii) Sei nun  $x \in \mathbb{R}^n$  fest aber beliebig. Dann gilt aufgrund von (1.24)

$$\begin{aligned} |u(x)| &\leq \int_{B_1(x)} |u(x) - u(y)| \, dy + \int_{B_1(x)} |u(y)| \, dy \\ &\leq c \int_{B_1(x)} \frac{|\nabla u(y)|}{|x - y|^{n-1}} \, dy + c \|u\|_{L^p(B_1(x))} \\ &\leq c \|\nabla u\|_{L^p(B_1(x))} \left( \int_{B_1(x)} \frac{dy}{|x - y|^{(n-1)\frac{p}{p-1}}} \right)^{\frac{p-1}{p}} + c \|u\|_{L^p(B_1(x))} \\ &\leq c \|u\|_{H^{1,p}(\mathbb{R}^n)}, \end{aligned} \tag{1.25}$$

wobei in der letzten Zeile benutzt wurde, dass  $(n-1)\frac{p}{p-1} < n$  ist, da  $p > n$ , und somit

$$\int_{B_1(x)} \frac{dy}{|x - y|^{(n-1)\frac{p}{p-1}}} < \infty.$$

Da  $x$  beliebig war, liefert (1.25) sofort

$$\|u\|_\infty \leq c \|u\|_{1,p}. \tag{1.26}$$

(iii) Seien  $x, y \in \mathbb{R}^n$ ,  $r = |x - y|$  und  $W = B_r(x) \cap B_r(y)$ . Wir haben

$$|u(x) - u(y)| \leq \int_W |u(x) - u(z)| \, dz + \int_W |u(y) - u(z)| \, dz. \tag{1.27}$$

Mit Hilfe von (1.24) erhalten wir

$$\begin{aligned}
 \int_W |u(x) - u(z)| \, dz &\leq c \int_{B_r(x)} |u(x) - u(z)| \, dz \\
 &\leq c \|\nabla u\|_{L^p(B_r(x))} \left( \int_{B_r(x)} \frac{dz}{|x-z|^{(n-1)\frac{p}{p-1}}} \right)^{\frac{p-1}{p}} \\
 &\leq c \left( r^{n-(n-1)\frac{p}{p-1}} \right)^{\frac{p-1}{p}} \|\nabla u\|_{L^p(\mathbb{R}^n)} \\
 &= cr^{1-\frac{n}{p}} \|\nabla u\|_p
 \end{aligned}$$

und eine analoge Abschätzung für den zweiten Term in (1.27). Somit gilt

$$|u(x) - u(y)| \leq c |x - y|^{1-\frac{n}{p}} \|\nabla u\|_p,$$

was sofort

$$[u]_{1-\frac{n}{p}} \leq c \|\nabla u\|_p$$

liefert. Dies zusammen mit (1.26) ist (1.23) im Falle  $\Omega = \mathbb{R}^n$  und  $u \in C^1(\mathbb{R}^n)$ .

(iv) Sei nun  $\Omega$  ein beschränktes Gebiet mit  $\partial\Omega \in C^{0,1}$  und  $C : H^{1,p}(\Omega) \rightarrow H^{1,p}(\mathbb{R}^n)$  der Fortsetzungsoperator mit den Eigenschaften (1.19). Da  $Cu$  einen kompakten Träger hat, erhalten wir die Existenz einer Folge  $u_m \in C_0^\infty(\mathbb{R}^n)$  mit

$$u_m \rightarrow Cu \quad \text{in} \quad H^{1,p}(\mathbb{R}^n).$$

Die Ungleichung (1.23) angewendet auf  $u_m - u_k$  liefert, dass  $(u_m)$  eine Cauchyfolge in  $C^{1-\frac{n}{p}}$  ist, d.h.

$$u_m \rightarrow Cu \quad \text{in} \quad C^{1-\frac{n}{p}}(\mathbb{R}^n).$$

Die Ungleichung (1.23) angewendet auf  $u_m$  liefert

$$\|u_m\|_{C^{1-\frac{n}{p}}(\mathbb{R}^n)} \leq c \|u_m\|_{H^{1,p}(\mathbb{R}^n)},$$

welches zusammen mit den obigen Konvergenzen von  $u_m$

$$\|Cu\|_{C^{1-\frac{n}{p}}(\mathbb{R}^n)} \leq c \|Cu\|_{H^{1,p}(\mathbb{R}^n)}$$

liefert. Dies und die Eigenschaften (1.19) ergeben sofort (1.23). ■

h) Die schwachen  $L^p$ -Räume („Marcinkiewicz-Räume“)

Zur Motivation der folgenden Definition überlegen wir uns das

**1.28 Lemma.** *Sei  $u \in L^p(\Omega)$ . Dann gilt für alle  $K > 0$*

$$\mu(\{x \in \Omega \mid |u(x)| \geq K\}) \leq \frac{\|u\|_p^p}{K^p}. \quad (1.29)$$



BEWEIS :

$$\|u\|_p^p \geq \int_{\{|u| \geq K\}} |u|^p dx \geq K^p \mu(\{|u| \geq K\}).$$

■

Die Eigenschaft (1.29) gibt Anlass zur folgenden Definition:

**1.30 Definition.** Eine messbare Funktion  $u$  liegt im Raum  $L_\omega^p(\Omega)$  (*schwach- $L^p$* ), wenn eine Konstante  $c_0 > 0$  existiert, so dass für alle  $K > 0$

$$\mu(\{x \in \Omega \mid |u(x)| \geq K\}) \leq \frac{c_0^p}{K^p}. \quad (1.31)$$

Das Infimum der Konstanten in (1.31) ist eine *Quasi-Norm* in  $L_\omega^p$ , d.h. die Dreiecksungleichung wird ersetzt durch

$$\|u + v\|_{L_\omega^p} \leq c (\|u\|_{L_\omega^p} + \|v\|_{L_\omega^p})$$

mit einer Konstanten  $c$ . Man kann allerdings eine Norm auf  $L_\omega^p(\Omega)$  finden, welche die gleiche Topologie erzeugt. Man kann zeigen, dass für beschränktes  $\Omega$  für  $\varepsilon > 0$ ,  $p - \varepsilon \geq 1$

$$L^p(\Omega) \subsetneq L_\omega^p(\Omega) \subsetneq L^{p-\varepsilon}(\Omega)$$

gilt. Die Beweise kann man in Kufner, John, Fučík: Function Spaces, Academia, finden.

Die schwachen  $L^p$ -Räume werden in der sogenannten Interpolationstheorie und in der Theorie der singulären Integrale benötigt. Ein Beispiel ist der Satz

**1.32 Satz.** Sei  $\Omega$  eine beschränkte, offene Punktmenge mit glattem Rand und  $u \in H_0^{1,2}(\Omega)$  eine schwache Lösung von

$$-\Delta u = f \in L^1(\Omega).$$

Dann gilt  $\nabla^2 u \in L_\omega^1$ .

Leider gilt nicht  $\nabla^2 u \in L^1$ , das Analogon für  $1 < p < \infty$  ist jedoch richtig, d.h. gilt in obigem Satz zusätzlich  $f \in L^p$ , dann folgt  $\nabla^2 u \in L^p$ , falls  $1 < p < \infty$ . Einen Beweis dieser Aussagen kann man in Stein, Singular integrals and differentiability properties of functions, Princeton, finden.

i) Die Morrey-Räume

Für die vertiefende Theorie der partiellen Differentialgleichungen sind die Morrey-Bedingung und die dazugehörigen Räume sehr wichtig.

**1.33 Definition.** Eine Funktion  $u \in L^p(\Omega)$ ,  $\Omega \subset \mathbb{R}^n$  messbar, genügt einer **Morrey-Bedingung**, wenn eine Konstante  $K$  sowie ein  $\lambda \in [0, n)$  existiert, so dass für alle Kugeln  $B_R \subset \mathbb{R}^n$  mit Radius  $R \in (0, R_0)$ ,  $R_0$  fest, z.B.  $R_0 = 1$ , die Ungleichung

$$\int_{B_R \cap \Omega} |u|^p dx \leq K R^{n-\lambda}. \quad (1.34)$$

gilt.

Der **Morrey-Raum**  $L^{p,\lambda}(\Omega)$  ist derjenige Teilraum von  $L^p(\Omega)$ , dessen Elemente die Bedingung (1.34) erfüllen. Sehr ähnlich zu  $L^{p,\lambda}(\Omega)$  ist der gewichtete  $L^p$ -Raum, der aus allen  $L^p$ -Funktionen mit der Eigenschaft

$$\int_{\Omega} |u|^p \frac{1}{|x - x_0|^{n-\lambda}} dx \leq K \quad (1.35)$$

besteht.

Die Norm in  $L^{p,\lambda}$  wird definiert durch das Infimum der zulässigen Konstanten  $K$  in (1.34) oder äquivalent

$$\|u\|_{L^{p,\lambda}} = \sup \left\{ R^{\lambda-n} \int_{B_R(x_0) \cap \Omega} |u|^p dx \mid 0 < R < R_0, x_0 \in \mathbb{R}^n \right\}.$$

Die Morrey-Bedingung beinhaltet, dass die Singularitäten von  $u$  eingeschränkt werden. Wenn z.B.  $u \in L^{1,\lambda}$ , so kann  $u$  z.B. nicht gleich der Funktion  $\frac{1}{|x-x_0|^{\lambda+\varepsilon}}$ ,  $\varepsilon > 0$ , sein. Besonders wichtig sind die folgenden beiden Einbettungssätze:

**1.36 Satz.** Sei  $\Omega \subset \mathbb{R}^n$  ein beschränktes Gebiet mit  $\partial\Omega \in C^{0,1}$ . Die  $H^{1,p}(\Omega)$ -Funktion  $u$  genüge der Morrey-Bedingung

$$\int_{B_R \cap \Omega} |\nabla u|^p dx \leq KR^{n-p+\alpha p} \quad (1.37)$$

mit einer Konstanten  $K$  für alle Kugeln  $B_R \subset \mathbb{R}^n$ ,  $0 < R \leq R_0$  sowie einer Konstanten  $\alpha \in (0, 1)$ . Dann ist  $u \in C^\alpha(\overline{\Omega})$ .

Anders ausgedrückt:  $H^{1,p;p-\alpha p}(\Omega) \subset C^\alpha(\overline{\Omega})$ , wobei  $H^{1,p;\lambda}$  der Teilraum von  $H^{1,p}$  bezeichnet, dessen Elemente  $u$  die Inklusion  $\nabla u \in L^{p,\lambda}$  erfüllen. Satz 1.36 ist ein äußerst wichtiger Stetigkeitstest.

**1.38 Satz (Marcinkiewicz-Stampacchia).** Sei  $\Omega \subset \mathbb{R}^n$  ein beschränktes Gebiet mit  $\partial\Omega \in C^{0,1}$  dann gilt  $H^{1,p;\lambda} \subset L^q(\Omega)$  mit

$$\frac{1}{q} = \frac{1}{p} - \frac{1}{\lambda},$$

sofern  $\lambda > 0$ . Die entsprechenden Normen werden entsprechend abgeschätzt:

$$\|u\|_q \leq K \|u\|_{1,p;\lambda}.$$

Erläuterung: Für  $\lambda = n$  ist die Aussage dieses Satzes der übliche Sobolevsche Einbettungssatz, für  $n > \lambda > p$  erhält man eine bessere  $L^q$ -Inklusion, als sie die Sobolevschen Ungleichungen liefern würden.

j) Der Raum  $BV[a, b]$  von Funktionen beschränkter Variation

**1.39 Definition.** Eine Funktion  $f : [a, b] \rightarrow \mathbb{R}$  hat **beschränkte Variation**, wenn für alle  $N \in \mathbb{N}$  und alle Paare  $t_j, t'_j$ ,  $j = 1, \dots, N$  mit  $t_j < t'_j < t_{j+1}$  die sogenannte **Totalvariation**

$$V(f) = \sup \left\{ \sum_{j=1}^N |f(t_j) - f(t'_j)| \mid a \leq t_j < t'_j < t_{j+1} \leq b, j = 1, \dots, N \in \mathbb{N} \right\} \quad (1.40)$$

endlich ist.

Die Menge der Funktionen mit beschränkter Variation bildet einen Banachraum  $BV[a, b]$  bezüglich der Norm

$$\|f\|_{BV} = \|f\|_{L^1} + V(f).$$

Alternativ kann man die Totalvariation auch durch

$$V(f) = \sup_{\substack{\varphi \in C_0^\infty([a,b]) \\ \|\varphi\|_{L^\infty([a,b])} \leq 1}} \int_a^b f \varphi' dx$$

definieren. Diese Definition lässt sich leicht auf den Fall  $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$  durch

$$V(f) = \sup_{\substack{\varphi \in C_0^\infty(\Omega)^n \\ \|\varphi\|_{L^\infty(\Omega)} \leq 1}} \int_{\Omega} f \operatorname{div} \varphi dx$$

verallgemeinern. Der Raum  $BV$  ist etwas schwächer als der Raum  $H^{1,1}$ . In der Tat gilt:

$$H^{1,1}(\Omega) \subset BV(\Omega) \subset L^{\frac{n}{n-1}}(\Omega).$$

Die Beweise können in Giusti: Minimal surfaces and functions of bounded variation, Birkhäuser, gefunden werden. Der Raum  $BV$  spielt eine große Rolle in der Theorie von Erhaltungssätzen, der Variationsrechnung und der Geometrie.

Sehr viel allgemeiner als Banachräume oder normierte Räume sind topologische lineare Räume, die jedoch im Großen und Ganzen zu allgemein sind, um brauchbare Anwendungen zu liefern.

**1.41 Definition.** Ein **topologischer linearer Raum** ist ein Vektorraum  $V$  über  $\mathbb{R}$  bzw.  $\mathbb{C}$ , auf dem eine Topologie erklärt ist, die mit der Addition in  $V$  und der Multiplikation mit Skalaren verträglich ist.

**Erläuterung:** Eine **Topologie** auf  $V$  ist durch ein System  $\tau$  von Teilmengen von  $V$  gegeben, welches den folgenden Bedingungen genügt:

(O1)  $V \in \tau, \emptyset \in \tau$ ,

(O2) die Vereinigungen einer beliebigen Familie von Mengen aus  $\tau$  ist wieder eine Menge aus  $\tau$ ,

(O3) der Durchschnitt einer Familie von endlich vielen Mengen aus  $\tau$  ist wieder eine Menge aus  $\tau$ .

Die Mengen des Systems  $\tau$  heißen **offenen Mengen**. Eine Menge  $U \subset V$  heißt **Umgebung eines Punktes**  $x \in V$ , wenn es eine offene Menge  $O \in \tau$  gibt mit  $x \in O \subset U$ .

Äquivalent ist eine Topologie gegeben, wenn es zu jedem Punkt  $x \in V$  ein System  $\mathfrak{U}_x$  von Umgebungen  $U \subset V$  gibt, so dass

(U1)  $x \in U$  für  $U \in \mathfrak{U}_x$ ,

(U2)  $U \cap U' \in \mathfrak{U}_x$ , falls  $U, U' \in \mathfrak{U}_x$ ,

(U3)  $W \in \mathfrak{U}_x$ , falls ein  $U \in \mathfrak{U}_x$  mit  $U \subset W$  existiert.

(U4) Ist  $U$  eine Umgebung von  $x$  dann existiert eine Umgebung  $W$  von  $x$ , so dass  $U$  eine Umgebung aller Punkte  $y \in W$  ist.

Eine Menge  $O \subset V$  heißt **offen**, wenn es für alle Punkte  $x \in O$  eine Umgebung  $U$  von  $x$  gibt mit  $x \in U \subset O$ . Lässt man den dritten Punkt weg, so spricht man von einer *Umgebungsbasis*. Um eine Topologie auf  $V$  zu definieren, genügt es, eine Basis von Nullumgebungen anzugeben, da man die Umgebungsbasis eines allgemeinen Punktes aufgrund der linearen Struktur von  $V$  durch Translation erhält.

Durch die Topologie wird ein Konvergenzbegriff und Stetigkeitsbegriff erklärt. Eine Folge  $(x_n)$ ,  $x_n \in V$  heißt **konvergent** gegen einen Punkt  $x \in V$ , wenn es zu jeder Umgebung  $U$  des Punktes  $x$  einen Index  $n_0$  gibt, so dass  $x_n \in U$  für alle  $n \geq n_0$ . Man schreibt  $x_n \rightarrow x$ , ( $n \rightarrow \infty$ ). Eine Abbildung  $f : V \rightarrow W$  eines topologischen linearen Raumes  $V$  in einen topologischen linearen Raum  $W$  heißt **stetig im Punkt**  $x \in V$ , wenn es zu jeder Umgebung  $V \subset W$  von  $f(x)$  eine Umgebung  $U \subset V$  von  $x$  gibt für die gilt  $f(U) \subset V$ . Die Abbildung  $f : V \rightarrow W$  heißt **stetig**, wenn sie stetig in allen Punkten  $x \in V$  ist.

Die Verträglichkeit der Addition in  $V$  und der Multiplikation mit Skalaren mit der Topologie bedeutet, dass diese stetige Funktionen bezüglich der Topologie sind.

Eine größere Bedeutung für Anwendungen in der Analysis haben die lokalkonvexen Räume.

**1.42 Definition.** Ein **lokalkonvexer** (topologischer linearer) Raum ist ein topologischer linearer Raum mit einer aus **konvexen** Mengen bestehenden Nullumgebungsbasis.

Eine wichtige Klasse von lokalkonvexen Räumen sind Räume mit *Multi-Norm*.

**1.43 Definition.** Sei  $V$  ein linearer Raum. Eine **Multi-Norm** ist eine Familie von **Semi-Normen**  $\|\cdot\|_i$ , auf  $V$ ,  $i \in I$ , die in ihrer Gesamtheit definit sind, d.h. aus  $\|u\|_i = 0$  für alle  $i \in I$  folgt  $u = 0$ .

Eine **Semi-Norm** erfüllt die Dreiecksungleichung und die positive Homogenität, braucht aber nicht definit zu sein. Die Indexmenge  $I$  braucht nicht abzählbar zu sein; ist sie abzählbar, nennt man den Raum **abzählbar normiert** und kann ihn durch die Metrik

$$\rho(x, y) = \sum_{j=1}^{\infty} \frac{1}{2^j} \frac{\|x - y\|_j}{1 + \|x - y\|_j}$$

metrisieren.

**Beispiele von lokalkonvexen Räumen:**

a) Die Räume  $C(\Omega)$ ,  $C^m(\Omega)$ .

Sei  $\Omega \subset \mathbb{R}^n$  offen und  $\Omega_j \subset \Omega$  offen und beschränkt,  $\bigcup_{j=1}^{\infty} \Omega_j = \Omega$ . Der Raum  $C(\Omega)$  bzw.  $C^m(\Omega)$  aller auf  $\Omega$  stetigen bzw.  $m$ -mal stetig differenzierbaren Funktionen wird mit Hilfe der Multi-Norm

$$\|u\|_j = \sup_{\overline{\Omega_j}} |u|$$

bzw.

$$\|u\|_j = \sup_{\overline{\Omega_j}} \left( |u| + \sum_{|\alpha|=1}^m |\nabla^\alpha u| \right)$$

ein lokalkonvexer Raum.

b) Die Räume  $L_{\text{loc}}^p(\mathbb{R}^n)$ ,  $H_{\text{loc}}^{m,p}(\mathbb{R}^n)$ .

Es handelt sich hierbei um die linearen Räume von Funktionen, deren Restriktion auf beschränkte Mengen des  $\mathbb{R}^n$  in  $L^p$  bzw.  $H^{m,p}$  liegt. Eine Multi-Norm ist gegeben durch

$$\|u\|_j = \|u\|_{L^p(\Omega_j)}$$

mit  $\Omega_j \subset \mathbb{R}^n$ , beschränkt und messbar,  $\bigcup_{j=1}^{\infty} \Omega_j = \mathbb{R}^n$ . Der Raum  $H_{\text{loc}}^{m,p}$  wird analog definiert.

Entsprechend bildet man  $L_{\text{loc}}^p(\Omega)$ -Räume für  $\Omega \neq \mathbb{R}^n$ .

c) Der Raum  $\mathcal{D} = C_0^\infty(\Omega)$

Es sei  $\Omega \subset \mathbb{R}^n$  offen.  $C_0^\infty(\Omega)$  besteht aus allen unendlich oft differenzierbaren Funktionen  $f$ , deren Träger  $\text{supp } f = \overline{\{x \in \Omega \mid f(x) \neq 0\}}$  eine kompakte Menge in  $\Omega$  ist.  $\bigcup_{j=1}^{\infty} \Omega_j = \Omega$ . Eine Multi-Norm in  $\mathcal{D}$  ist gegeben durch

$$\|f\|_{j,m} := \sup_{\overline{\Omega_j}} |\nabla^\alpha f|, \quad |\alpha| = 0, 1, \dots, m.$$

Der Nachteil dieser Multi-Norm ist, dass  $\mathcal{D}$  mit dieser Multi-Norm versehen *nicht* vollständig ist. Das Problem besteht darin, dass der Träger des Grenzwerts einer Cauchyfolge keine kompakte Teilmenge von  $\Omega$  sein muss. Allerdings kann man auf  $\mathcal{D}$  eine lokal konvexe Topologie  $\tau$  definieren, die nicht metrisierbar ist, in der Cauchy Folgen konvergieren. Im weiteren betrachten wir  $\mathcal{D}$  mit dieser Topologie  $\tau$  versehen (siehe Rudin: Functional Analysis, Mc Graw).

d) Der Raum  $\mathcal{S}$

Der Raum  $\mathcal{S}$  besteht aus allen unendlich oft differenzierbaren Funktionen, die im Unendlichen stärker als jedes Polynom fallen. Eine Multi-Norm ist gegeben durch

$$\|f\|_{j,m} = \sup_{\mathbb{R}^n, |\alpha|=m} |\nabla^\alpha f| (|x| + 1)^j.$$

Auch in lokalkonvexen linearen Räumen lässt sich definieren, wann eine Menge beschränkt ist.

**1.44 Definition.** Eine Menge  $M$  in einem lokalkonvexen linearen Raum  $E$  heißt **beschränkt**, wenn es zu jeder Umgebung  $U$  der Null eine Zahl  $t > 0$  gibt, so dass  $M \subset tU$  gilt.

Die Frage, ob die Beschränktheit einer Menge die relative Kompaktheit nach sich zieht, ist sehr wichtig und sehr gut studiert, so dass man heutzutage eine kristallklare, abstrakte Erklärung dieses Phänomens hat.

Für die Räume  $\mathcal{S}$  und  $\mathcal{D}$  gelten die Inklusion „Beschränktheit  $\Rightarrow$  relative Kompaktheit“, obwohl sie unendlich-dimensional sind.

## 3.2 Beschränkte lineare Funktionale und beschränkte lineare Operatoren

**2.1 Definition.** Ein **lineares Funktional**  $f$  auf einem normierten Raum  $V$  ist eine lineare Abbildung von  $V$  in die reellen bzw. komplexen Zahlen. Das lineare Funktional heißt **beschränkt**, wenn eine Konstante  $K$  existiert mit

$$|f(u)| \leq K \|u\| \quad \text{für alle } u \in V.$$

Die Norm eines beschränkten, linearen Funktionals  $f$  wird durch

$$\|f\| = \sup_{\|u\|=1} |f(u)| \tag{2.2}$$

definiert.

**2.3 Lemma.** Ein lineares Funktional ist genau dann beschränkt, wenn es stetig ist.

BEWEIS : Der Beweis verläuft wie im Hilbertraum. (Vergleiche Kapitel 2 Satz 3.3) ■

Die Menge der beschränkten linearen Funktionale bildet bezüglich der natürlichen Addition und Multiplikation mit Skalaren einen Vektorraum, welcher als **Dualraum**  $V^*$  von  $V$  bezeichnet wird. Hierbei ist die Bezeichnung

$$\langle f, u \rangle = f(u) \quad \text{für } f \in V^*, u \in V$$

üblich.

**2.4 Lemma.**  $V^*$  ist bezüglich der in (2.2) definierten Norm ein Banachraum.

BEWEIS : Der Beweis geschieht ähnlich wie im Hilbertraum. (Übungsaufgabe) ■

Mit Hilfe des Dualraums lässt sich die *schwache Konvergenz* in einem normierten Raum definieren.

**2.5 Definition.** Eine Folge  $(u_m)_{m \in \mathbb{N}}$  in einem normierten Raum  $V$  **konvergiert schwach** gegen ein Element  $u \in V$ , in Zeichen:

$$u_m \rightharpoonup u \quad \text{schwach in } V \text{ (} m \rightarrow \infty \text{),}$$

wenn für jedes  $f \in V^*$

$$\langle f, u_m \rangle \rightarrow \langle f, u \rangle.$$

Den Dualraum  $V^{**}$  zu  $V^*$  nennt man den **bidualen Raum** zu  $V$ . In  $V^*$  gibt es zwei Arten von Konvergenz, nämlich die schwache Konvergenz wie oben definiert, sowie die **\*-schwache Konvergenz**:

$$f_m \xrightarrow{*} f \quad \text{im Sinne der *-schwachen Konvergenz}$$

bedeutet, dass für alle  $u \in V$

$$\langle f_m, u \rangle \rightarrow \langle f, u \rangle \quad (m \rightarrow \infty).$$

### 2.6 Satz (\*-schwache relative Kompaktheit beschränkter Mengen).

Es sei  $(f_m)$ ,  $f_m \in V^*$  eine beschränkte Folge im Dualraum eines normierten, separablen Raumes. Dann gibt es eine \*-schwach konvergente Teilfolge.

Anmerkung: Separabilität wird wie im Hilbertraum definiert.

BEWEIS : Sei  $(u_j)$  eine Folge linear unabhängiger Elemente aus  $V$ , deren lineare Hülle dicht in  $V$  ist. Nach dem Doppelfolgensatz gibt es eine Teilfolge  $\Lambda \subset \mathbb{N}$ , so dass

$$\langle f_m, u_j \rangle \quad \text{konvergiert für alle } j, m \in \Lambda, m \rightarrow \infty.$$

Wir setzen

$$\langle f, u_j \rangle = \lim_{m \in \Lambda, m \rightarrow \infty} \langle f_m, u_j \rangle.$$

Da  $\text{span} \langle u_j \rangle$  dicht in  $V$  ist und die  $f_m$  beschränkt, lässt sich die Definition von  $\langle f, u \rangle$  und die Konvergenz  $\langle f_m, u \rangle \rightarrow \langle f, u \rangle$  auf alle  $u \in V$  durch Abschliessung ausdehnen. Die Beschränktheit von  $f$  sieht man aus der Gleichung bzw. Ungleichung

$$|\langle f, u \rangle| = o(1) + |\langle f_m, u \rangle| \leq o(1) + \|f_m\| \|u\| \leq o(1) + K \|u\|.$$

Der \*-schwache Limes der  $f_m$  und die Teilfolge sind damit konstruiert. ■

### Der Dualraum zu $L^p(\Omega)$

**2.7 Satz.** Sei  $\Omega$  eine messbare Teilmenge des  $\mathbb{R}^n$  und  $p \in [1, \infty)$ . Dann existiert zu jedem  $f \in (L^p(\Omega))^*$  eine eindeutig bestimmte Funktion  $g \in L^q(\Omega)$ ,  $q = \frac{p}{p-1}$  ( $q = \infty$ , falls  $p = 1$ ) mit

$$\langle f, u \rangle = \int_{\Omega} gu \, dx. \quad (2.8)$$

Umgekehrt beschreibt jedes  $g \in L^q(\Omega)$  über Gleichung (2.8) ein Element  $f \in (L^q(\Omega))^*$ . Aufgrund dieses Satzes sagt man, dass  $L^q(\Omega)$  der Dualraum zu  $L^p(\Omega)$  ist,  $q = \frac{p}{p-1}$ ,  $1 \leq p < \infty$ .

BEWEIS : Wir führen den Beweis nur in dem einfacheren Fall  $1 < p \leq 2$  bei beschränktem  $\Omega$  aus. Der Fall  $p > 2$  findet sich etwa in dem Buch von KANTOROWITSCH/AKILOV, „Funktionalanalysis in normierten Räumen“.

Sei also  $p \in (1, 2]$ ,  $\Omega$  beschränkt und  $f \in (L^p(\Omega))^*$ . Da dann  $L^2(\Omega) \subset L^p(\Omega)$ , ist  $f$  insbesondere ein beschränktes lineares Funktional auf  $L^2(\Omega)$  und wird nach dem Rieszschen Darstellungssatz durch ein Element  $g \in L^2(\Omega)$  dargestellt:

$$\langle f, u \rangle = \int_{\Omega} g u \, dx, \quad u \in L^2(\Omega). \quad (2.9)$$

Wir zeigen, dass sogar  $g \in L^{\frac{p}{p-1}}(\Omega)$  gilt. Durch Abschluss in  $L^p$ , da  $L^2$  dort dicht ist, folgt dann aus (2.9) die behauptete Darstellung. Hierzu definieren wir:

$$g_L = \begin{cases} g & \text{falls } |g| \leq L \\ 0 & \text{falls } |g| > L \end{cases}$$

und wählen in (2.9)  $u = |g_L|^s \operatorname{sign} g$ . Da  $g_L$  beschränkt, ist  $u \in L^2$  und damit zulässig in (2.9). Es gilt

$$\begin{aligned} \int_{\Omega} |g_L|^{s+1} \, dx &= \int_{\Omega} g u \, dx = \langle f, u \rangle \leq K \|u\|_p \\ &= K \left( \int_{\Omega} |g_L|^{sp} \, dx \right)^{\frac{1}{p}}. \end{aligned}$$

Wir wählen  $s$  so, dass  $s + 1 = sp$ , d.h.  $s = \frac{1}{p-1}$ . Wegen  $s + 1 = \frac{p}{p-1}$  folgt dann

$$\int_{\Omega} |g_L|^{\frac{p}{p-1}} \, dx \leq K \left( \int_{\Omega} |g_L|^{\frac{p}{p-1}} \, dx \right)^{\frac{1}{p}},$$

und wegen  $1 - \frac{1}{p} = \frac{p-1}{p}$

$$\int_{\Omega} |g_L|^{\frac{p}{p-1}} \, dx \leq K^{\frac{p}{p-1}}.$$

Mit dem Grenzübergang  $L \rightarrow \infty$  folgt aus dem Satz über monotone Konvergenz, dass  $g \in L^{\frac{p}{p-1}}$ . ■

## Der duale Raum zu $C[a, b]$

Zur Vorbereitung benötigen wir den Begriff des Riemann-Stieltjes-Integrals.

Es sei  $g \in BV[a, b]$ , also eine Funktion mit beschränkter Variation. Zu jedem  $f \in C[a, b]$  definieren wir die *Riemann-Stieltjes-Summe*

$$\sum_{k=1}^N f(\xi_k) (g(x_{k+1}) - g(x_k))$$



bezüglich einer Zerlegung  $a = x_0 < x_1 < \dots < x_N = b$  von  $[a, b]$  und einer Belegung  $\xi_k \in [x_k, x_{k+1}]$ . Das **Riemann-Stieltjes-Integral**  $\int_a^b f dg$  wird dann ähnlich wie das klassische Riemann-Integral als Grenzwert der Riemann-Stieltjes-Summen bei gegen Null gehender *Feinheit*  $\max_{k=1, \dots, N} |x_k - x_{k+1}|$  definiert. Es gilt die Abschätzung

$$\left| \int_a^b f dg \right| \leq \|f\|_\infty V(g),$$

so dass  $\int_a^b f dg$  ein beschränktes lineares Funktional auf  $C[a, b]$  definiert.  $V(g)$  ist hierbei die Totalvariation von  $g$ , siehe Definition 1.37. Umgekehrt lässt sich zeigen, dass jedes  $\varphi \in C^*[a, b]$  als Riemann-Stieltjes-Integral dargestellt werden kann (siehe Kantorovitch/Akilow Kap. VI.3). Die Zuordnung  $\varphi \in C^* \leftrightarrow g \in BV$  ist nicht eindeutig, man muss sich auf „normalisierte“ Elemente aus  $BV$  einigen, d.h. nur rechtsstetige oder nur linksstetige  $BV$ -Funktionen  $g$  zulassen.

**Beispiel** Wir betrachten den Raum  $C[-1, 1]$  und das *Delta-Funktional*  $\delta$  mit

$$\langle \delta, f \rangle = f(0).$$

Offensichtlich ist  $\delta \in C^*[-1, 1]$ . Die Darstellung als Stieltjes-Integral lautet:

$$\langle \delta, f \rangle = \int_{-1}^1 f dg$$

mit

$$g(t) = \begin{cases} 1, & t \geq 0, \\ 0, & t < 0. \end{cases}$$

### Der Dualraum zu $L^\infty(\Omega)$ .

Sei  $\Omega$  messbar und beschränkt,  $\Omega \subset \mathbb{R}^n$ , sowie  $\varphi$  eine **additive, reelle Mengenfunktion**, welche auf den Lebesgue-messbaren Teilmengen von  $\Omega$  definiert ist. Additiv bedeutet:

$$\varphi(e_1 \cup e_2) = \varphi(e_1) + \varphi(e_2), \quad e_1 \cap e_2 = \emptyset.$$

Man kann sich z.B.  $\varphi(e) = \int_e w dx$  vorstellen,  $w \in L^1$ .

**2.10 Definition.** Eine additive, reelle Mengenfunktion  $\varphi$  ist von **beschränkter Variation**, wenn

$$\sup \left\{ \left| \varphi \left( \bigcup_{k=1}^N e_k \right) \right| \mid e_k \subset \Omega, e_k \text{ messbar}, e_j \cap e_k = \emptyset, j \neq k \right\} < \infty.$$

Für jede additive Mengenfunktion  $\varphi$  von beschränkter Variation und  $f \in L^\infty(\Omega)$  lässt sich das Lebesgue-Stieltjes-Integral über die entsprechenden *Lebesgue-Stieltjes-Summen* definieren:

$$\sum_{j=1}^N l_j \varphi(\{l_{j-1} < f \leq l_j\}), \quad l_0 < -\|f\|_\infty, \quad l_N > \|f\|_\infty,$$

wobei  $l_0 < l_1 < \dots < l_{N-1} < l_N$  eine Zerlegung des Bildbereiches von  $f$  ist. Die entsprechenden Limites werden als Radon-Integral oder **Lebesgue-Stieltjes-Integral**  $\int_\Omega f d\varphi$  bezeichnet. Man kann zeigen, dass sich jedes beschränkte lineare Funktional auf  $L^\infty(\Omega)$  in der Form

$$\int_\Omega f d\varphi$$

mit einer additiven, auf den Lebesgue-messbaren Teilmengen von  $\Omega$  definierten Mengenfunktion  $\varphi$  darstellen lässt. Hierbei hat  $\varphi(e)$  *nicht* notwendig die Gestalt  $\int_e w dx$ .

Details können in Kantorowitsch-Akilov, VI.4 gefunden werden.

### 3.3 Das Prinzip der gleichmäßigen Beschränktheit

Dieser Abschnitt behandelt einen grundlegenden Satz über Folgen beschränkter, linearer Operatoren in Banachräumen.

**3.1 Satz.** Sei  $(A_n)$  eine Folge linearer, beschränkter Operatoren, die einen Banachraum  $B$  in einen normierten Raum  $V$  abbilden. Die Folge  $(A_n)$  sei **punktweise beschränkt**, d.h. für alle  $x \in B$  gelte

$$\sup_n \|A_n x\| < \infty.$$

Dann sind die Normen der Operatoren gleichmäßig beschränkt, d.h.

$$\sup_n \|A_n\| < \infty.$$

• Für *nichtlineare* Abbildungen ist die entsprechende Aussage bereits im  $\mathbb{R}$  falsch:

$$f_n(x) = \begin{cases} |x|^2 & \text{für } |x| \leq n \\ n^2 & \text{für } |x| \geq n \end{cases}$$

Eine alternative Formulierung ist der Satz von der Festlegung einer Singularität:

**3.2 Lemma.** Gilt  $\sup_n \|A_n\| = \infty$ , so gibt es ein  $x$  mit  $\sup_n \|A_n x\| = \infty$ .

BEWEIS (Satz 3.1): (i) Wir zeigen zunächst, dass für eine lineare Abbildung  $A$  aus der Ungleichung

$$\|Ax\| \leq C \quad \text{für } x \in B_\delta(x_0) = \{y \in B \mid \|y - x_0\| < \delta\}$$

die Ungleichung

$$\|A\| \leq \frac{2C}{\delta} \quad (3.3)$$

folgt, d.h. ist  $A$  punktweise auf einer kleinen Kugel mit Mittelpunkt  $x_0$  beschränkt - egal, wo  $x_0$  liegt - so erfüllt die Operatornorm eine konkrete Abschätzung. Dies gilt entsprechend dann für Folgen von Operatoren  $(A_n)$ .

Nachweis von (3.3): Für jedes  $y$  mit  $\|y\| < 1$  gilt  $x = x_0 + \delta y \in B_\delta(x_0)$ . Nach Voraussetzung gilt

$$\|A(x_0 + \delta y)\| \leq C,$$

daher

$$\delta\|Ay\| \leq C + \|Ax_0\| \leq 2C.$$

Übergang zum Supremum bezüglich  $\|y\| = 1$  liefert (3.3).

(ii) Wir führen einen Widerspruchsbeweis. Angenommen, dass  $\|A_n\| \rightarrow \infty$  ( $n \rightarrow \infty$ ). Sei  $p(x) = \sup_n \|A_n(x)\|$ . Die Abbildung  $p$  muss auf jeder Kugel  $B_\delta(x_0)$  unbeschränkt

sein, andernfalls schließt man aus (3.3), dass doch nicht  $\|A_n\| \rightarrow \infty$ .

Da die Abbildung  $p$  auf jeder noch so kleinen Kugel unbeschränkt ist, liegt die Menge  $E_k = \{x \in B \mid p(x) > k\}$  dicht in  $B$ . Ausserdem ist diese Menge offen, denn ist  $p(x_0) > k$ , so folgt  $\|A_{n_0}x_0\| > k$  für ein  $n_0 \in \mathbb{N}$ , und daher  $\|A_{n_0}x\| > k$  für  $x \in B_\varepsilon(x_0)$ . Dies wieder impliziert  $p(x) > k$  für  $x \in B_\varepsilon(x_0)$ .

In vollständigen Räumen gilt jedoch, dass der Schnitt einer Folge von offenen Mengen, die dicht in  $B$  liegen, nicht leer ist.

Die eben gemachte Aussage lässt sich mit Hilfe des Baireschen Kategoriebegriffs durchleuchten - siehe etwa KANTOROWITSCH/AKILOV; der Einfachheit halber verzichten wir darauf und beweisen sie direkt in dem folgendem Lemma.

Da  $E_{k+1} \subset E_k$  und  $E_k$  offen und dicht sind, ergibt sich aus Lemma 3.4 die Existenz eines Elementes  $x_0 \in \bigcap_{k=1}^{\infty} E_k$ . Für dieses gilt  $\sup_n \|A_n x_0\| \geq k$  für alle  $k$ , also  $\sup_n \|A_n x_0\| = \infty$ , was der Voraussetzung des Satzes widerspricht. ■

Zum Beweis von Satz 3.3 haben wir das folgende Lemma, welches eine duale Formulierung des sogenannten *Baireschen Kategoriensatzes* ist, benötigt.

**3.4 Lemma.** *Es sei  $\mathcal{O}_i$  eine Folge von offenen Mengen in einem Banachraum  $B$ . Jede der Mengen  $\mathcal{O}_i$  sei dicht in  $B$ . Dann ist*

$$\bigcap_{i=1}^{\infty} \mathcal{O}_i \neq \emptyset.$$

• Lässt man die Forderung der Dichtigkeit von  $\mathcal{O}_i$  fallen, ist die Aussage bereits im  $\mathbb{R}$  falsch. Beispiel:  $\mathcal{O}_i = (0, \frac{1}{i})$ . Es gilt  $\mathcal{O}_i \supseteq \mathcal{O}_{i+1}$  und  $\frac{1}{j} \notin \mathcal{O}_i$ ,  $i > j$ , sowie  $0 \notin \mathcal{O}_i$ . Der Schnitt der  $\mathcal{O}_i$  muss daher leer sein.

BEWEIS : Wir konstruieren eine Folge von ineinander geschachtelten Kugeln  $B_{\varepsilon_i}(x_i)$ ,  $\varepsilon_i \rightarrow 0$ , die in allen  $\mathcal{O}_j$  liegen, so dass  $x_i \rightarrow x^*$ . Von diesem Punkt  $x^*$  zeigen wir, dass  $x^* \in \mathcal{O}_i$  für alle  $i \in \mathbb{N}$ .

(i) Da  $\mathcal{O}_1$  offen ist, gibt es eine Kugel  $B_{\varepsilon_1}(x_1) \subset \mathcal{O}_1$ . Da  $\mathcal{O}_2$  dicht in  $B$  ist, ist

$\mathcal{O}_2$  auch dicht in  $B_{\varepsilon_1/2}(x_1)$ . Es gibt daher ein  $x_2 \in B_{\varepsilon_1/2}(x_1) \cap \mathcal{O}_2$ , und da  $\mathcal{O}_2$  offen, gibt es ein  $\varepsilon_2$  mit  $B_{\varepsilon_2}(x_2) \subset B_{\varepsilon_1/2}(x_1) \cap \mathcal{O}_2$  und somit  $B_{\varepsilon_2}(x_2) \subset \mathcal{O}_1 \cap \mathcal{O}_2$ . Es gilt  $\varepsilon_2 \leq \frac{\varepsilon_1}{2}$ . Diese Konstruktion wird induktiv fortgesetzt. Ist  $B_{\varepsilon_n}(x_n) \subset B_{\varepsilon_{n-1}/2}(x_{n-1}) \cap \mathcal{O}_n$  schon konstruiert, so finden wir, da  $\mathcal{O}_{n+1}$  dicht in  $B_{\varepsilon_n/2}(x_n)$  ist, eine Kugel  $B_{\varepsilon_{n+1}}(x_{n+1}) \subset B_{\varepsilon_n/2}(x_n) \cap \mathcal{O}_{n+1}$ . Es gilt  $\varepsilon_{n+1} \leq \frac{\varepsilon_n}{2}$  sowie  $B_{\varepsilon_n}(x_n) \subset \bigcap_{j=1}^n \mathcal{O}_j$ ,

da  $B_{\varepsilon_n}(x_n) \subset \mathcal{O}_n \cap B_{\varepsilon_{n-1}/2}(x_{n-1})$  und  $B_{\varepsilon_{n-1}}(x_{n-1}) \subset \bigcap_{j=1}^{n-1} \mathcal{O}_j$ .

(ii) Die Mittelpunkte  $x_n$  konvergieren. In der Tat gilt zunächst  $\varepsilon_j \leq \varepsilon_1 2^{-j+1}$ , ausserdem

$$x_{j+k} \in B_{\varepsilon_j/2}(x_j), \quad k \in \mathbb{N}, \quad \text{d.h. } \|x_{j+k} - x_j\| \leq \varepsilon_j/2 \rightarrow 0 \quad (j \rightarrow \infty), k \in \mathbb{N}.$$

Die Folge  $(x_j)$  konvergiert daher gegen ein  $x^*$ . Wir behaupten, dass  $x^* \in \mathcal{O}_j$  für alle  $j \in \mathbb{N}$ . In der Inklusion  $x_{j+k} \in B_{\varepsilon_j/2}(x_j)$  können wir zur Grenze  $k \rightarrow \infty$  übergehen, es ergibt sich  $x^* \in B_{\varepsilon_j/2}(x_j) \subset \mathcal{O}_j$ . ■

Eine äquivalente Folgerung aus Lemma 3.4 ergibt sich durch Komplementbildung: Die Komplemente der Mengen  $\mathcal{O}_j$  in Lemma 3.4 sind abgeschlossene, *nirgends dichte* Mengen  $A_j$ , d.h.  $A_j$  enthält keine Kugel. Da  $\bigcap_{j=1}^{\infty} \mathcal{O}_j \neq \emptyset$ , kann die Vereinigung der  $A_j$  nicht den ganzen Raum ergeben. Somit haben wir auch den folgenden Satz bewiesen.

**3.5 Satz (Bairesche Kategoriensatz).** *Die Vereinigung von abzählbar vielen abgeschlossenen, nirgends dichten Mengen eines Banachraumes kann nicht den ganzen Raum ergeben.*

- Satz und Beweis gelten analog in vollständigen metrischen Räumen.

Unter Benutzung schöner Vokabeln lässt sich Satz 3.5 folgendermaßen formulieren:

**3.6 Satz.** *Ein Banachraum (bzw. vollständiger metrischer Raum) ist von zweiter Kategorie.*

Hierbei heißt eine Menge von „zweiter Kategorie“, wenn sie nicht von „erster Kategorie“ ist, d.h. die Vereinigung abzählbar vieler nirgends dichter Mengen ist.

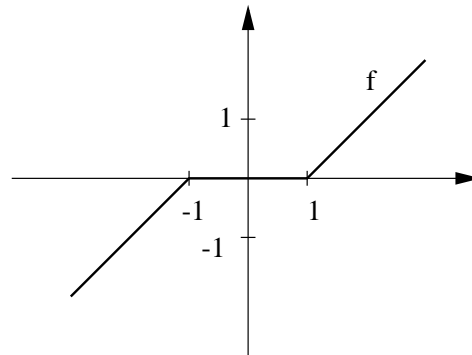
Ebenfalls eine Folge des Baireschen Kategoriensatzes ist der Satz von der offenen Abbildung.

**3.7 Definition.** *Eine Abbildung eines topologischen Raumes in einen anderen heißt **offen**, wenn sie offene Mengen in offene Mengen überführt.*

**3.8 Satz (von der offenen Abbildung).** *Jede stetige, lineare, surjektive Abbildung  $A : E \rightarrow F$  eines Banachraumes  $E$  in einen Banachraum  $F$  ist offen.*

- Betrachten wir den Fall  $E = F = \mathbb{R}^n$  mit der üblichen Topologie. Stetige, surjektive nichtlineare Abbildungen sind selbst hier nicht offen. Beispiel für  $n = 1$ :

$$f(x) = \begin{cases} 0, & \text{für } |x| \leq 1, \\ x - 1, & \text{für } x > 1, \\ x + 1, & \text{für } x < -1. \end{cases}$$



Offensichtlich ist  $f(U_\varepsilon(1))$  nicht notwendig offen.

Der Fall  $E = F = \mathbb{R}^n$  mit einer linearen, surjektiven Abbildung  $A$  ist elementar. Wenn  $A$  surjektiv ist, gilt  $\det A > 0$  und  $A^{-1}$  existiert auf  $\mathbb{R}^n$ . Da  $A^{-1}$  stetig, sind Urbilder offener Mengen offen, d.h.  $A$  ist offen.

Im  $\mathbb{R}^n$  gibt es den berühmten Satz von Brouwer über die Gebietsinvarianz. Dieser besagt, dass jede stetige, *injektive* Abbildung  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$  offen ist.

In der Theorie der komplexen Funktionen gibt es ebenfalls einen solchen Satz.

**3.9 Satz (über die Gebietsinvarianz).** *Sei  $f : \mathbb{C} \rightarrow \mathbb{C}$  holomorph und nicht konstant. Dann ist  $f$  offen.*

BEWEIS (Satz 3.8): (i) Wir zeigen, dass eine Konstante  $c > 0$  existiert mit

$$B_{2c} \subset \overline{AB_1} \quad (3.10)$$

Hierbei ist  $B_r(z) = \{x \in E \text{ bzw. } F \mid \|x - z\| < r\}$ ,  $B_r = B_r(0)$  und  $AB_1$  das Bild von  $B_1$  unter der Abbildung  $A$ , der Querstrich bedeutet den Abschluss der Menge.

Um (3.10) zu beweisen, setzen wir  $X_n = \overline{nAB_1}$ . Da  $A$  surjektiv ist, gilt  $\bigcup_{n=1}^{\infty} X_n = F$ , und nach dem Baireschen Kategoriensatz (Satz 3.5) muss eines der  $X_n$  eine Kugel enthalten, d.h.  $\text{int } X_{n_0} \neq \emptyset$ . Daraus folgt, dass auch  $\text{int } X_1 = \text{int } \overline{AB_1} \neq \emptyset$ , d.h.  $\exists c > 0$  und  $y_0 \in F$  mit

$$B_{4c}(y_0) \subset \overline{AB_1}.$$

Insbesondere ist  $y_0 \in \overline{AB_1}$  und aus Symmetriegründen  $-y_0 \in \overline{AB_1}$ . Daraus folgt

$$B_{4c}(0) = -y_0 + B_{4c}(y_0) \subset \overline{AB_1} + \overline{AB_1} = 2\overline{AB_1}.$$

Daraus ergibt sich (3.10).

(ii) Wir zeigen, dass mit obigem  $c$

$$B_c \subset AB_1. \quad (3.11)$$

Um (3.11) zu beweisen, müssen wir zu  $y \in B_c$  ein  $x \in B_1$  mit  $Ax = y$  finden. Wegen (3.10) existiert zu jedem  $\varepsilon > 0$  ein  $z \in E$  mit  $\|z\| < \frac{1}{2}$  und  $\|y - Az\| < \varepsilon$ , denn

$$y \in B_c \subset \overline{AB_{1/2}}.$$

Wählt man  $\varepsilon = \frac{c}{2}$ , hat man ein  $z_1 \in E$  mit

$$\|z_1\| < \frac{1}{2} \quad \text{und} \quad \|y - Az_1\| < \frac{c}{2}.$$

Wir wiederholen das Spielchen mit  $y - Az_1$  anstelle von  $y$  und wählen nun  $\varepsilon = \frac{c}{4}$ . Da  $y - Az_1 \in B_{c/2} \subset \overline{AB_{1/4}}$ , finden wir ein  $z_2$  mit

$$\|z_2\| < \frac{1}{4} \quad \text{und} \quad \|(y - Az_1) - Az_2\| < \frac{c}{4}.$$

Dieses Argument wird wiederholt, und wir erhalten eine Folge  $(z_n)$  mit

$$\|z_n\| < \frac{1}{2^n} \quad \text{und} \quad \|y - A(z_1 + z_2 + \dots + z_n)\| < \frac{c}{2^n}.$$

Die Elemente  $x_n = z_1 + z_2 + \dots + z_n$  bilden daher eine Cauchyfolge mit Limes  $x$ . Es gilt  $\|x_n\| \leq \|z_1\| + \sum_{j=2}^{\infty} \frac{1}{2^j}$  und somit

$$\|x\| \leq \|z_1\| + \sum_{j=2}^{\infty} \frac{1}{2^j} < \frac{1}{2} + \sum_{j=2}^{\infty} \frac{1}{2^j} = 1.$$

Ferner gilt  $\|y - Ax_n\| \rightarrow 0$ , und schliesslich  $y - Ax = 0$ . Das gewünschte  $x$  mit  $\|x\| < 1$  ist damit konstruiert.

(iii) Sei  $U \subseteq E$  offen und sei  $y \in AU$ . Dann gibt es ein  $x_0 \in U$  mit  $y = Ax_0$ . Da  $U$  offen ist gibt es ein  $\varepsilon > 0$  mit  $B_\varepsilon(x_0) \subseteq U$  und also gilt

$$AB_\varepsilon(x_0) \subseteq AU.$$

Aber  $AB_\varepsilon(x_0) = Ax_0 + \varepsilon AB_1$ , was mit Hilfe von (3.11) liefert  $AB_\varepsilon(x_0) \supseteq Ax_0 + \varepsilon B_c = B_{\varepsilon c}(Ax_0) = B_{\varepsilon c}(y)$ , d.h.  $B_{\varepsilon c}(y) \subseteq AU$ , d.h.  $AU$  ist offen. ■

Aus dem Satz über die offene Abbildung folgt der Satz von der stetigen Inversen, da die Stetigkeit einer Abbildung äquivalent dazu ist, dass Urbilder offener Mengen offen sind.

**3.12 Satz (von der stetigen Inversen).** *Es seien  $E$  und  $F$  Banachräume und  $A : E \rightarrow F$  linear, stetig und bijektiv. Dann ist  $A^{-1}$  stetig.*

BEWEIS : Wir bezeichnen  $B := A^{-1} : F \rightarrow E$ . Sei  $U \subseteq E$  offen. Das Urbild  $B^{-1}U$  ist aber die Menge  $AU$ , da  $B^{-1} = (A^{-1})^{-1} = A$  gilt. Aber nach Satz 3.8 ist die Abbildung  $A$  offen, d.h.  $AU$  ist offen. Somit ist auch  $B^{-1}U$  offen, d.h.  $B = A^{-1}$  ist stetig. ■

Eine weitere elegante Folge ist

**3.13 Satz (vom abgeschlossenen Graphen).** *Seien  $E$  und  $F$  Banachräume,  $A : E \rightarrow F$  linear. Der Graph  $G(A) = \{(x, Ax) \in E \times F\}$  sei abgeschlossen bezüglich der Graphennorm*

$$\|(x, Ax)\| = \|x\| + \|Ax\|.$$

*Dann ist  $A$  stetig.*

BEWEIS : Wir wählen als Raum  $E$  in Satz 3.12 den Raum  $G(A)$  mit der Graphen-norm, als Bildraum den Raum  $E$ ; als Abbildung  $A_0 : G(A) \rightarrow E$  definieren wir

$$A_0(x, Ax) = x.$$

Offensichtlich ist  $A_0$  bijektiv. Ausserdem ist  $A_0$  linear und beschränkt, da

$$\|A_0(x, Ax)\| = \|x\| \leq \|x\| + \|Ax\| = \|x\|_{G(A)}.$$

Nach Satz 3.12 ist damit  $A_0^{-1}$  stetig, d.h. für  $x_j \rightarrow x$  folgt  $(x_j, Ax_j) \rightarrow (x, Ax)$  und somit  $Ax_j \rightarrow Ax$ . ■

Eine weitere Folgerung ist die Äquivalenz von Normen.

**3.14 Satz.** *Der Banachraum  $B$  sei bezüglich der Normen  $\|\cdot\|_0$  und  $\|\cdot\|_1$  vollständig. Es gelte  $\|u\|_1 \leq \|u\|_0$  für alle  $u \in B$ . Dann gibt es eine Konstante  $K$ , so dass*

$$\|u\|_0 \leq K\|u\|_1.$$

BEWEIS : Wir wenden Satz 3.12 an mit  $A = I$  als identischer Abbildung. Der Raum  $E$  bzw.  $F$  sei der mit  $\|\cdot\|_0$  bzw.  $\|\cdot\|_1$  versehene Raum  $B$ . Als Abbildung von  $E$  nach  $F$  ist die identische Abbildung  $I$  stetig, da  $\|x\|_1 = \|Ix\|_1 \leq \|x\|_0$ . Nach Satz 3.12 ist daher die Inverse stetig; daraus folgt, dass ein  $K > 0$  existiert mit

$$\|x\|_0 = \|I^{-1}x\|_0 \leq K\|x\|_1.$$

Eine weitere überraschende Folgerung aus unseren Sätzen ist der *Satz von Hellinger-Toeplitz*:

**3.15 Satz.** *Sei  $H$  ein Hilbertraum und  $A : H \rightarrow H$  linear und selbstadjungiert. Dann ist  $A$  beschränkt.*

BEWEIS : Wir versehen  $H$  mit der Graphen-Norm  $\|x\|_A = \|x\| + \|Ax\|$ .  $H$  ist bezüglich  $\|x\|_A$  vollständig! Gilt nämlich  $\|x_j - x_k\|_A \rightarrow 0$ , so folgt  $\|x_j - x_k\| \rightarrow 0$  und  $\|Ax_j - Ax_k\| < \varepsilon$  für  $j, k \geq n_0(\varepsilon)$ . Hieraus schließen wir  $|(Ax_j - Ax_k, \varphi)| < \varepsilon$  für alle  $\varphi$  mit  $\|\varphi\| \leq 1$ ,  $j, k \geq n_0(\varepsilon)$ . Aufgrund der Vollständigkeit von  $(H, \|\cdot\|)$  folgt die Existenz von  $x \in H$  mit  $x_k \rightarrow x$ , und es gilt  $(Ax_k, \varphi) = (x_k, A\varphi) \rightarrow (x, A\varphi) = (Ax, \varphi)$  für  $k \rightarrow \infty$ . Weiter folgt

$$|(Ax_j - Ax, \varphi)| < \varepsilon \quad \|\varphi\| \leq 1,$$

und somit

$$\|Ax_j - Ax\| \leq \varepsilon \quad j \geq n_0.$$

Da  $\|x\| \leq \|x\|_A$  und  $H$  bezüglich beider Normen vollständig ist, folgt aus Satz 3.14 die Ungleichung

$$\|x\|_A \leq K\|x\|$$

und

$$\|Ax\| \leq (K + 1)\|x\|.$$

■

Eine Anwendung des Satzes von der stetigen Inversen ist der Nachweis der stetigen Abhängigkeit von Lösungen gewöhnlicher Differentialgleichungen von den Daten. Zur Illustration betrachten wir

$$Au := u' + cu = f \quad \text{in } [0, T], \quad u(0) = 0 \in \mathbb{R}^n \quad (3.16)$$

mit gegebenen Funktionen  $f \in L^p(0, T)$ ,  $c \in L^\infty(0, T)$ . Die vektorwertige Funktion  $u \in H^{1,p}(0, T; \mathbb{R}^n)$  ist gesucht. Der in (3.16) definierte Operator  $A$  bildet den Raum  $\tilde{H}_0^{1,p} = \{v \in H^{1,p}; v(0) = 0\}$  nach  $L^p(0, T)$  ab.  $A$  ist stetig als Abbildung von  $\tilde{H}_0^{1,p}$  nach  $L^p$ , außerdem ist  $A$  surjektiv und injektiv, da sich (3.16) nach den Sätzen über globale Lösbarkeit von Anfangswertproblemen gewöhnlicher Differentialgleichungen eindeutig lösen lässt. Nach dem Satz über die stetige Inverse folgt damit, dass die Lösung von (3.16) stetig von der rechten Seite  $f$  abhängt, d.h. konvergieren die Funktionen  $f_n$  gegen  $f$  in  $L^p$ , so konvergieren die Lösungen  $u_n$  gegen  $u$  in  $H^{1,p}$ . Das gleiche Argument lässt sich verwenden, wenn man  $f \in C[a, b]$  und  $u \in C^1[a, b]$  betrachtet und  $A : C^1[a, b] \cap H_0^{1,p} \rightarrow C[a, b]$  auffasst.

### Tonnelierte Räume (Englisch: barreled spaces)

Ein sehr allgemeiner Raumtyp in der Klasse der lokalkonvexen Räume, in der ein Analogon des Prinzips der gleichmässigen Beschränktheit gilt, sind die *tonnelierten* Räume. Man kann zeigen, dass vollständige, lokalkonvexe, metrische Räume tonneliert sind.

**3.17 Definition.** Ein lokalkonvexer Raum  $V$  heißt **tonneliert**, wenn jede abgeschlossene, konvexe, absorbierende Menge eine Nullumgebung ist.

Eine Menge  $M$  heißt **absorbierend**, wenn zu  $u \in V$  ein  $\lambda_0 \in \mathbb{R}$  existiert, so dass für alle  $\lambda \geq \lambda_0$  gilt:  $\lambda^{-1}u \in M$ .

- In lokalkonvexen Räumen kann man immer eine Umgebungsbasis der Null finden, die aus *absolut konvexen* Mengen  $V$  besteht, d.h. es gilt  $V = \{\lambda x \mid |\lambda| \leq 1, x \in V\}$  und  $V$  ist konvex.

Wir betrachten lokalkonvexe Räume  $E$  und  $F$  und eine Menge  $U := \{A_i \mid i \in J\}$  von linearen, stetigen Operatoren  $A_i : E \rightarrow F$ .

**3.18 Definition.** Die Menge  $U$  heißt **punktweise beschränkt**, wenn für jedes  $x$  die Menge  $U(x) = \{A_i(x) \mid i \in J\}$  beschränkt ist.

**3.19 Definition.** Die Menge  $U = \{A_i : E \rightarrow F \mid i \in J\}$  heißt **gleichgradig stetig**, wenn man zu jeder Nullumgebung  $W$  im Raum  $F$  eine Nullumgebung  $V$  im Raum  $E$  finden kann, so dass

$$A_i(V) \subset W \quad \text{für alle } i \in J$$

gilt.

Der Begriff „gleichgradig stetig“ ersetzt die gleichmäßige Beschränktheit der Operatornormen.



**3.20 Satz.** *Sei  $E$  tonneliert. Dann ist jede punktweise beschränkte Menge von stetigen, linearen Operatoren gleichgradig stetig.*

BEWEIS : Sei  $W$  eine absolut konvexe, abgeschlossene Nullumgebung im Raum  $F$ . Wir setzen  $V_i = A_i^{-1}(W)$  und  $V = \bigcap_{i \in J} V_i$ . Wir wollen zeigen, daß  $V$  eine abgeschlossene, konvexe, absorbierende Menge ist. Da alle  $A_i$  stetig sind, und Urbilder abgeschlossener Mengen abgeschlossen sind, ist  $V$  abgeschlossen. Da  $W$  absolut konvex ist und  $A_i$  linear sind, sind  $A_i^{-1}(W)$  absolut konvex. Somit ist auch  $V$  absolut konvex. Um zu sehen, dass  $V$  absorbierend ist, sei  $x \in E$ . Da nach Voraussetzung die Menge  $U(x) = \{A_i(x) \mid i \in J\}$  beschränkt ist, existiert ein  $t_0 > 0$ , so dass für alle  $|t| \leq t_0$  gilt:  $U(tx) = tU(x) \subseteq t_0U(x) \subseteq W$  (siehe Definition 1.44). Somit gilt erst recht  $tx \in A_i^{-1}(U(tx)) \subseteq V_i, i \in J$  und folglich  $tx \in V$ , d.h.  $V$  ist absorbierend. Da  $E$  tonneliert ist, ist  $V$  also eine Nullumgebung. Nach Konstruktion von  $V$  gilt:  $A_i(V) \subseteq W, i \in J$ , d.h. die Menge  $U = \{A_i \mid i \in J\}$  ist gleichgradig stetig. ■

## 3.4 Der Satz von Hahn-Banach

Der Satz von Hahn-Banach befasst sich mit der Fortsetzung von linearen Funktionale, welche auf einem linearen Teilraum definiert sind, auf den ganzen Vektorraum - dies unter Beibehaltung der Norm bzw. verwandter Größen. Im nicht-separablen Fall wird zum Beweis des Lemma von Zorn verwendet.

Ähnlich wie die Sätze im vorigen Kapitel zählt der Satz von Hahn-Banach zu den grundlegenden Sätzen der Funktionalanalysis. Wir geben eine analytische und später eine geometrische Formulierung dieses Satzes.

### Analytische Formulierung des Satzes von Hahn-Banach

**4.1 Satz (von Hahn-Banach, analytische Form, reeller Fall).** *Sei  $V$  ein Vektorraum über  $\mathbb{R}$  und  $p : V \rightarrow \mathbb{R}$  subadditiv und positiv homogen, d.h.  $p(\lambda x) = \lambda p(x)$  für alle  $x \in V$  und  $\lambda \in \mathbb{R}, \lambda > 0$ , und  $p(x + y) \leq p(x) + p(y)$ . Ferner sei  $W$  ein linearer Teilraum von  $V$  und  $\varphi : W \rightarrow \mathbb{R}$  ein lineares Funktional mit*

$$\varphi(x) \leq p(x) \quad \text{für alle } x \in W.$$

*Dann gibt es eine Fortsetzung  $f : V \rightarrow \mathbb{R}$  von  $\varphi$ , d.h.  $f|_W = \varphi$ , so dass*

$$f(x) \leq p(x) \quad \text{für alle } x \in V.$$

- Besonders wichtig ist der Fall, dass  $p(x) = \|x\|$ , wenn  $V$  ein normierter Raum mit Norm  $\|\cdot\|$  ist.
- Nicht jedes beschränkte, lineare Funktional  $\varphi$  auf  $L^\infty(a, b)$  lässt sich in der Form

$$\varphi(z) = \int_a^b g z dx, \quad z \in L^\infty(a, b)$$

mit einer Funktion  $g \in L^1$  darstellen. In anderen Worten:  $L^1(a, b)$  ist ein echter Teilraum von  $(L^\infty)^*$ . Da  $(L^1)^* = L^\infty$ , folgt dass der  $L^1(a, b)$  nicht reflexiv ist.

BEWEIS ( $(L^\infty)^* \neq L^1$ ): Der Raum  $C(a, b)$  ist ein echter Teilraum von  $L^\infty$ . Wir betrachten das Dirac-Funktional  $\varphi = \delta$  mit

$$\delta(z) = z(0), \quad z \in C(a, b)$$

$((a, b)$  enthalte die Zahl 0). Die Voraussetzungen des Satzes von Hahn-Banach mit  $p(z) = \|z\|_\infty$  sind erfüllt, es gibt daher eine (sicherlich nicht eindeutige) Fortsetzung  $f$  von  $\varphi = \delta$  auf ganz  $L^\infty$  mit  $f(z) \leq \|z\|_\infty$  für alle  $z \in L^\infty$ . Die letzte Ungleichung gilt auch für  $-z$ , somit ist  $|f(z)| \leq \|z\|_\infty$  und  $f \in (L^\infty)^*$ . Das Dirac-Funktional lässt sich aber nicht in der Form  $\delta(z) = \int_a^b gz \, dx$  darstellen (vergleiche Abschnitt 3.2), somit auch nicht das Funktional  $f$ . ■

Zum Beweis von Satz 4.1 zeigen wir zunächst, dass das Funktional  $\varphi$  in Satz 4.1 auf einen größeren Raum  $W \oplus \langle x_0 \rangle$  unter Erhaltung der Ungleichung  $\varphi(x) \leq p(x)$  fortgesetzt werden kann.

**4.2 Lemma (Elementarerweiterung).** *Unter den Voraussetzungen von Satz 4.1 gibt es zu jedem  $x_0 \in V \setminus W$  eine lineare Fortsetzung  $f_1 : W \oplus \langle x_0 \rangle \rightarrow \mathbb{R}$  von  $\varphi : W \rightarrow \mathbb{R}$ , so dass*

$$f_1(y) \leq p(y) \quad \text{für alle } y \in W \oplus \langle x_0 \rangle. \quad (4.3)$$

BEWEIS : Wir setzen

$$f_1(x + tx_0) = \varphi(x) + t\alpha,$$

wobei  $\alpha$  noch definiert werden muss. Jedenfalls ist  $f_1$  linear auf  $W \oplus \langle x_0 \rangle$ . Um ein geeignetes  $\alpha$  zu finden, welches die Ungleichung (4.3) erfüllt, beachten wir, dass wegen der positiven Homogenität von  $p$  nur sichergestellt werden muss, dass für alle  $x \in W$

$$\begin{aligned} \varphi(x) + \alpha &\leq p(x + x_0), \\ \varphi(x) - \alpha &\leq p(x - x_0) \end{aligned} \quad (4.4)$$

gilt. Ersetzt man in (4.4)  $x$  durch  $x/t$ ,  $t > 0$ , und multipliziert mit  $t$ , so ergibt sich (4.3) für  $y = x \pm tx_0$ . Um (4.4) zu gewährleisten, muss man offensichtlich

$$\sup_{x \in W} \{\varphi(x) - p(x - x_0)\} \leq \alpha \leq \inf_{x \in W} \{p(x + x_0) - \varphi(x)\}$$

sicherstellen. Eine solche Wahl von  $\alpha$  ist aber möglich, da

$$\varphi(\xi) - p(\xi - x_0) \leq p(x + x_0) - \varphi(x), \quad \forall \xi, x \in W. \quad (4.5)$$

Es gilt nämlich

$$\varphi(x) + \varphi(\xi) = \varphi(x + \xi) \leq p(x + \xi) \leq p(x + x_0) + p(\xi - x_0),$$

woraus (4.5) folgt. ■

Bevor wir Satz 4.1 beweisen, bemerken wir, dass der wichtige Fall von separablen normierten Räumen sowie dem Funktional  $p(x) = K\|x\|$  einfach zu beweisen ist.

**BEWEIS** (Satz 4.1, separabler, normierter Raum,  $p(x) = K\|x\|$ ): Es sei  $\{x_j \mid j \in \mathbb{N}\}$  eine Menge von linear unabhängigen Vektoren aus  $V$ , deren lineare Hülle in  $V$  dicht ist. Sei  $\{x_j \mid j \in \Lambda \subset \mathbb{N}\}$  die Teilmenge von Vektoren  $x_j$ , die nicht in  $\overline{W}$  liegen. Durch sukzessive Anwendung der Elementarerweiterung mit  $x_0 = x_j$ ,  $j \in \Lambda$ , erhalten wir eine Fortsetzung  $f_1$  von  $\varphi$  auf  $W \oplus \langle x_j \mid j \in \Lambda \rangle$  unter Beibehaltung der Ungleichung  $f_1(x) \leq K\|x\|$ . Durch Abschluss erhalten wir eine Fortsetzung auf den ganzen Raum  $V$ , d.h. wir definieren

$$f(x) := \lim_{y_k \rightarrow x} f_1(y_k), \quad y_k \in W \oplus \langle x_j \mid j \in \Lambda \rangle.$$

■

Der Beweis des allgemeinen Satzes von Hahn-Banach geschieht mit Hilfe des *Lemmas von Zorn*. Zur Formulierung dieses Lemmas benötigen wir einige Begriffe: Es sei  $M$  eine Menge mit einer **Halbordnung**  $\prec$ , d.h. für gewisse Paare  $(a, b) \in M \times M$  ist

$$a \prec b \quad \text{oder} \quad b \prec a.$$

Wenn einer der beiden Fälle zutrifft, sagt man „ $a$  und  $b$  sind vergleichbar“. Es gelten die Regeln

- (i)  $a \prec a$ ,
- (ii)  $\{a \prec b \text{ und } b \prec a\} \Rightarrow a = b$ ,
- (iii)  $\{a \prec b \text{ und } b \prec c\} \Rightarrow a \prec c$ .

**Beispiel:** (i)  $M = \mathbb{R}^n$  mit der Halbordnung  $\leq$ , die definiert ist durch

$$a \leq b \Leftrightarrow a_i \leq b_i \quad \text{für die Komponenten } a_i \text{ und } b_i, i = 1, \dots, n.$$

(ii) Die Menge der linearen Teilräume eines Vektorraumes bildet eine Halbordnung bezüglich der mengentheoretischen Inklusion.

Eine **Kette**  $Q$  ist eine Teilmenge einer Menge  $M$  mit Halbordnung, welche **total geordnet** ist, d.h. für je zwei Elemente  $a, b \in Q$  gilt  $a \prec b$  oder  $b \prec a$ .

**Beispiel:** Jede eindimensionale Strecke im  $\mathbb{R}^n$  mit einem Richtungsvektor der nicht-negative Komponenten hat ist bzgl. der Halbordnung  $\leq$  eine Kette.

Eine **obere Schranke** für eine Teilmenge  $S$  einer Menge  $M$  mit Halbordnung  $\prec$  ist ein Element  $a$  mit

$$s \prec a \quad \text{für alle } s \in S.$$

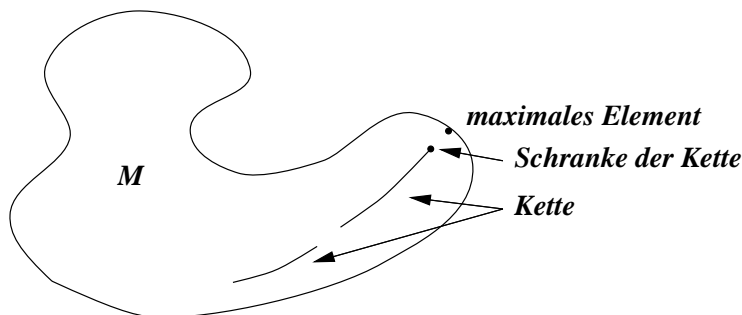
Ein Element  $m \in M$  heißt **maximal** (auf  $M$ ), wenn es von keinem anderen Element  $x \in M$  bezüglich  $\prec$  übertroffen werden kann, d.h. aus

$$m \prec x \quad \text{folgt } x = m.$$

(Es ist aber durchaus möglich, dass  $m$  maximal ist und mit dem Element  $x$  nicht vergleichbar ist.)

**4.6 Lemma (Zorn).** *Sei  $M$  eine nichtleere Menge mit Halbordnung. Jede Kette aus  $M$  besitze eine obere Schranke. Dann besitzt  $M$  ein maximales Element.*

**Beispiel:** Sei  $M \subset \mathbb{R}^2$  beschränkt und abgeschlossen und mit der Halbordnung  $\leq$  aus Beispiel (i) versehen. Man überlegt sich - etwa mit Hilfe komponentenweiser Supremumsbildung - dass jede Kette eine obere Schranke besitzt. Es muss daher ein maximales Element in  $M$  geben.  $M$  braucht durchaus nicht „konvex“ sein.



Das Lemma von Zorn wird aus dem sogenannten *Auswahlaxiom* der Mengenlehre hergeleitet: *Ist  $\mathcal{F}$  eine Familie von nichtleeren Mengen, so gibt es eine Funktion, die jeder Menge  $M$  aus  $\mathcal{F}$  ein Element  $m \in M$  zuordnet.*

Dies erscheint evident, aber die Aussage ist äquivalent zum sogenannten *Wohlordnungssatz*, der zumindest einem Analytiker Unwohlsein erzeugt, wie sogleich erläutert wird.

Der Wohlordnungssatz besagt, dass jede Menge  $M$  wohlgeordnet werden kann, d.h. es gibt eine Ordnungsrelation  $\prec$  in  $M$ , die die Axiome der *Halbordnung* erfüllt, so dass je zwei Elemente  $a, b \in M$  *vergleichbar* sind (d.h. es gilt  $a \prec b$  oder  $b \prec a$ ) und zusätzlich die Eigenschaft besitzt, dass jede nichtleere Teilmenge von  $M$  ein kleinstes Element bezüglich der Ordnung  $\prec$  besitzt. Die letztere Eigenschaft bedeutet, dass die übliche Ordnung in  $\mathbb{R}$ , welche durch das  $\leq$ -Symbol gegeben ist, *keine* Wohlordnung ist, denn offene Intervalle in  $\mathbb{R}$  besitzen *kein* kleinstes Element. In der Tat hat bisher noch niemand eine Wohlordnung der reellen Zahlen konkret konstruiert - man weiss nur, dass sie existiert und aus dem Auswahlaxiom folgt.

Nimmt man den Wohlordnungssatz als gegeben an, ist der Beweis des Zornschen Lemmas „einfach“. Wir geben ihn an, weil in ihm das Prinzip der *transfiniten Induktion* verwendet wird, welches man aus „erkenntnis-theoretischen“ Gründen einmal gesehen haben sollte.

**BEWEIS** (Lemmas von Zorn mit Hilfe des Wohlordnungssatzes): Es sei  $M$  die im Lemma von Zorn genannte halbgeordnete Menge und  $x_\alpha$ ,  $\alpha \in I$ , eine Wohlordnung der Elemente  $x_\alpha \in M$ , d.h.  $\alpha$  ist eine wohlgeordnete Indexmenge bezüglich einer Ordnungsrelation, die wir mit dem Symbol  $\leq$  bezeichnen und zu jedem  $\alpha \in I$  gehört genau ein  $x_\alpha \in M$ . Wir bestimmen durch transfinite Induktion eine *Kette*  $G$  in  $M$ .

- (i) Sei  $\alpha_0$  der kleinste Index aus  $I$ . Dann gehöre  $x_{\alpha_0}$  zu  $G$ .
- (ii) Ist für  $\beta < \gamma$ ,  $\beta, \gamma \in I$ , entschieden, welche  $x_\beta$  zu  $G$  gehören, so gehöre  $x_\gamma$  genau dann zu  $G$ , wenn  $x_\beta \prec x_\gamma$  für alle  $x_\beta \in G$ ,  $\beta < \gamma$ .

Nach Konstruktion ist  $G$  eine Kette. Nach der Voraussetzung im Lemma von Zorn hat  $G$  eine obere Schranke  $z$ . Da  $z$  eines der  $x_\alpha$  ist, muss  $z$  in  $G$  liegen, ist also das größte Element von  $G$  und maximal in  $M$ . (Wenn es nicht maximal wäre, gäbe es ein  $x'_\alpha \succ z$ , welches aber dann aufgrund der Konstruktion von  $G$  in  $G$  liegen müsste,  $z$  könnte dann nur obere Schranke sein, wenn  $x'_\alpha = z$ .) ■

Der Kürze halber haben wir hier die Indexmenge  $I$  nicht konkretisiert und verweisen auf Bücher aus der Mengenlehre (Kamke).

Nach diesen Vorbereitungen beweisen wir den Satz von Hahn-Banach für reelle Vektorräume in voller Allgemeinheit.

BEWEIS (Satz 4.1): Wir betrachten die Menge

$$M = \{h : D(h) \longrightarrow \mathbb{R} \mid D(h) \text{ linearer Teilraum von } V, W \subset D(h), \\ h \text{ linear, } h|_W = \varphi, h(x) \leq p(x) \forall x \in D(h)\}.$$

In  $M$  besteht die Halbordnung  $\prec$ , welche erklärt ist durch

$$h_1 \prec h_2 \Leftrightarrow D(h_1) \subset D(h_2) \text{ und } h_2 \text{ ist Fortsetzung von } h_1.$$

Da  $\varphi \in M$ , ist  $M \neq \emptyset$ . Ferner erfüllt  $(M, \prec)$  die Voraussetzung des Lemmas von Zorn. Ist nämlich  $K$  eine Kette in  $M$ ,  $K = \{h_i \in M \mid i \in I\}$ , so definiert man

$$D(h) = \bigcup_{i \in I} D(h_i) \text{ und } h(x) = h_i(x), \quad x \in D(h_i).$$

Das so konstruierte Element  $h \in M$  ist dann eine obere Schranke von  $K$ . Nach dem Lemma von Zorn gibt es daher ein maximales Element  $f \in M$  bezüglich  $\prec$ . Wir behaupten, dass  $D(f) = V$ , was den Beweis dann vollenden würde. Angenommen, es wäre  $D(f) \neq V$ . Dann gibt es ein  $x_0 \in V$ ,  $x_0 \notin D(f)$  und wir führen eine Elementarerweiterung von  $f$  nach Lemma 4.2 durch.  $f$  wäre dann nicht maximal. ■

Im *komplexen Fall* sind die Ungleichungen in der reellen Formulierung des Satzes von Hahn-Banach - Satz 4.1 - folgendermaßen abzuändern: Für das lineare Funktional  $\varphi : W \rightarrow \mathbb{C}$  setzt man

$$|\varphi(x)| \leq p(x)$$

voraus und erhält eine Fortsetzung  $f : V \rightarrow \mathbb{C}$  von  $\varphi$  mit  $|f(x)| \leq p(x)$ .

## Geometrische Formulierung des Satzes von Hahn-Banach

Man kann den Satz von Hahn-Banach auch als Satz über die Trennung konvexer Mengen durch Hyperebenen formulieren. Dazu benötigen wir einige Begriffe. Wie im endlich-dimensionalen Fall definiert man Hyperebenen.

**4.7 Definition.** Eine **Hyperebene** in einem Vektorraum  $V$  über  $\mathbb{R}$  ist eine Menge der Gestalt

$$H = \{x \in V \mid \varphi(x) = \alpha\} =: \{\varphi = \alpha\}$$

mit einer linearen Abbildung  $\varphi : V \rightarrow \mathbb{R}$ ,  $\varphi \neq 0$ , und einer Zahl  $\alpha \in \mathbb{R}$ .

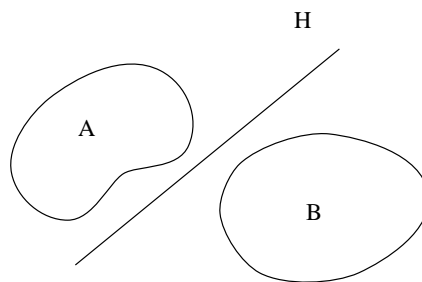
**4.8 Lemma.** Sei  $V$  ein normierter, reeller Vektorraum. Eine Hyperebene  $\{\varphi = \alpha\}$  ist genau dann abgeschlossen, wenn  $\varphi$  stetig ist.

BEWEIS : Übungsaufgabe ■

**4.9 Definition.** Seien  $A, B$  Teilmengen des normierten, reellen Vektorraumes  $V$ . Man sagt, „die Hyperebene  $\{\varphi = \alpha\}$  trenne  $A$  und  $B$ “, wenn

$$\varphi(x) \leq \alpha \leq \varphi(y) \quad \text{für alle } x \in A \text{ und } y \in B. \quad (4.10)$$

• Auf die Reihenfolge von  $A$  und  $B$  kommt es nicht an, d.h. (4.10) bedeutet auch, dass  $\{\varphi = \alpha\}$  die Mengen  $B$  und  $A$  trennt.



Man spricht von **strikt**er Trennung, wenn ein  $\varepsilon > 0$  existiert mit

$$\varphi(x) + \varepsilon \leq \alpha \leq \varphi(y) - \varepsilon, \quad x \in A, y \in B.$$

**4.11 Satz (Hahn-Banach, geometrische Form).** Sei  $V$  ein normierter, reeller Raum und  $A, B$  zwei nichtleere, disjunkte, konvexe Teilmengen von  $V$ . Die Menge  $A$  sei offen. Dann gibt es eine abgeschlossene,  $A$  und  $B$  trennende Hyperebene.

Zum Beweis benötigen wir das **Minkowski-Funktional**  $p_C$  einer konvexen Menge  $C \subset V$ ,  $0 \in C$

$$p_C(x) := \inf\{\alpha > 0 \mid \alpha^{-1}x \in C\}.$$

**4.12 Lemma.** Sei  $C$  eine offene, konvexe Teilmenge des normierten Vektorraumes  $V$ . Es sei  $0 \in C$ . Dann gilt

$$0 \leq p_C(x) \leq M \|x\| \quad \text{mit einer Konstanten } M \quad (4.13)$$

und

$$C = \{x \in V \mid p_C(x) < 1\}. \quad (4.14)$$

Ferner ist  $p_C$  positiv homogen und subadditiv.

BEWEIS : Sei  $B_r \subset C$ ,  $r > 0$ . Dann gilt  $(r - \varepsilon) \frac{x}{\|x\|} \in C$ ,  $x \in V$ , und  $p_C(x) \leq \alpha_0 = \frac{\|x\|}{r}$  nach Definition von  $p_C$ . Daraus folgt (4.13). Die positive Homogenität von  $p_C$  folgt aus

$$\begin{aligned} p_C(\lambda x) &= \inf\{\alpha > 0 \mid \alpha^{-1} \lambda x \in C\} \\ &= \inf\{\beta \lambda \mid \beta^{-1} x \in C\} \\ &= \lambda \inf\{\beta \mid \beta^{-1} x \in C\} = \lambda p_C(x), \quad \lambda > 0. \end{aligned}$$

Wir beweisen nun (4.14). Ist  $x \in C$ , so gilt auch  $(1 + \varepsilon)x \in C$  für genügend kleines  $\varepsilon$  wegen der Offenheit von  $C$ . Daher  $p_C(x) \leq \frac{1}{1+\varepsilon} < 1$ , d.h.  $C \subset \{x \in V \mid p_C(x) < 1\}$ . Ist umgekehrt  $p_C(x) < 1$ , dann folgt  $\alpha^{-1}x \in C$  mit einem  $\alpha > 0$  und somit  $x = \alpha(\alpha^{-1}x) + (1 - \alpha)0 \in C$ . Es verbleibt der Nachweis der Subadditivität von  $p_C$ . Seien  $x, y \in V$ ,  $\varepsilon > 0$ . Da

$$p_C\left(\frac{x}{p_C(x) + \varepsilon}\right) = \frac{1}{p_C(x) + \varepsilon} p_C(x) < 1,$$

folgt

$$\frac{x}{p_C(x) + \varepsilon} \in C \quad \text{und entsprechend} \quad \frac{y}{p_C(y) + \varepsilon} \in C.$$

Aus Konvexitätsgründen ist

$$\frac{tx}{p_C(x) + \varepsilon} + \frac{(1-t)y}{p_C(y) + \varepsilon} \in C, \quad 0 < t < 1.$$

Setzt man

$$t = \frac{p_C(x) + \varepsilon}{p_C(x) + p_C(y) + 2\varepsilon},$$

folgt

$$\frac{x + y}{p_C(x) + p_C(y) + 2\varepsilon} \in C.$$

Dies zusammen mit (4.14) und der positiven Homogenität liefert

$$p_C(x + y) \leq p_C(x) + p_C(y) + 2\varepsilon.$$

Grenzübergang  $\varepsilon \rightarrow 0$  ergibt die Behauptung. ■

**4.15 Lemma (Trennung einer konvexen Menge und eines Punktes).** *Sei  $V$  ein normierter, reeller Vektorraum und  $C \subset V$  offen, konvex und nichtleer. Sei  $x_0 \in V$ ,  $x_0 \notin C$ . Dann gibt es ein beschränktes, lineares Funktional  $f \in V^*$ ,  $f \neq 0$  mit*

$$f(x) < f(x_0) \quad \text{für alle } x \in C.$$

BEWEIS : O.B.d.A. sei  $0 \in C$ . Auf dem eindimensionalen Raum  $\langle x_0 \rangle$  definieren wir das lineare Funktional  $\varphi$  durch

$$\varphi(tx_0) = t.$$

Es gilt  $\varphi(x) \leq p_C(x)$ , für alle  $x \in \langle x_0 \rangle$  da andernfalls für ein  $y \in \langle x_0 \rangle$  gelten würde:

$$p_C(y) < \varphi(y), \quad \text{d.h. } p_C(tx_0) < t \text{ und somit } p_C(x_0) < 1.$$

Dies würde nach Lemma 4.12 aber  $x_0 \in C$  bedeuten. Nach dem Satz von Hahn-Banach in der analytischen Form (Satz 4.1) lässt sich  $\varphi$  zu einem linearen Funktional  $f : V \rightarrow \mathbb{R}$  fortsetzen mit

$$f(x) \leq p_C(x).$$

Insbesondere ist  $f(x_0) = 1$  und andererseits  $f(x) < 1$  für  $x \in C$  wegen (4.14). ■

BEWEIS (Satz 4.11): Die Menge  $C = A \ominus B \equiv \{a - b \mid a \in A, b \in B\}$  ist konvex. Wegen der Darstellung  $C = \bigcup_{y \in B} (A - y)$  ist  $C$  offen. Da  $A \cap B = \emptyset$ , gilt  $0 \notin C$ . Wegen Lemma 4.15 gibt es  $f \in V^*$  (d.h. ist  $f$  stetig und linear) mit

$$\begin{aligned} f(z) &< f(0) = 0, & z \in C, \\ f(x) &< f(y), & x \in A, y \in B. \end{aligned} \quad (4.16)$$

Wir wählen  $\alpha$  so, dass

$$\sup_{x \in A} f(x) \leq \alpha \leq \inf_{y \in B} f(y).$$

Die Hyperebene  $\{f = \alpha\}$  ist wegen Lemma 4.8 abgeschlossen und trennt  $A$  und  $B$ . ■

Im Unendlichdimensionalen lassen sich beliebige konvexe Mengen nicht notwendig trennen.

**Beispiel:**  $V = L^2(\Omega)$ ,  $\Omega$  ein Gebiet des  $\mathbb{R}^n$ .

$$A = \{z \in C(\Omega) \cap L^2 \mid \|z\|_2 < 1\}, \quad B = \{y_0\}$$

mit fester  $L^2$ -Funktion  $y_0 \notin C(\Omega)$ ,  $\|y_0\|_{L^2} = \frac{1}{2}$ . Man beachte, dass das Innere von  $A$  bezüglich der  $L^2$ -Norm leer ist. Ferner gilt  $A \cap B = \emptyset$ . Wir behaupten, dass  $A$  und  $B$  nicht durch eine abgeschlossene Hyperebene getrennt werden können. Gäbe es ein  $f \in (L^2)^*$  mit darstellendem Element  $u_0 \neq 0$ , für das

$$f(z) < f(y_0), \quad z \in C(\Omega) \cap L^2, \quad \|z\|_{L^2} < 1,$$

gilt, so folgt

$$(u_0, z) < (u_0, y_0), \quad \|z\|_{L^2} < 1,$$

und, da  $C(\Omega) \cap B_1$  dicht in  $B_1 = \{w \in L^2 \mid \|w\|_{L^2} \leq 1\}$ , ist

$$\|u_0\|_{L^2} \leq (u_0, y_0) \leq \frac{1}{2} \|u_0\|_{L^2}.$$

Wir notieren eine Variante der geometrischen Form des Satzes von Hahn-Banach.

**4.17 Satz.** Sei  $V$  ein normierter Vektorraum und  $A, C \subset V$  konvexe, disjunkte nicht-leere Teilmengen von  $V$ . Ferner sei  $A$  abgeschlossen und  $C$  kompakt. Dann gibt es eine abgeschlossene Hyperebene, die  $A$  und  $C$  strikt trennt.

BEWEIS : Für  $\varepsilon > 0$  setzen wir  $A_\varepsilon = A + B_\varepsilon(0)$ ,  $C_\varepsilon = C + B_\varepsilon(0)$ . Offensichtlich sind  $A_\varepsilon, C_\varepsilon$  offen, konvex und nichtleer. Falls  $\varepsilon > 0$  klein genug gewählt wurde gilt auch  $A_\varepsilon \cap C_\varepsilon = \emptyset$ . Falls dies nicht gelten würde, gäbe es  $\varepsilon_n \searrow 0$ ,  $x_n \in A$ ,  $y_n \in C$ ,  $\|x_n - y_n\| \leq 2\varepsilon_n$ . Da  $C$  kompakt ist gibt es eine Teilfolge  $\Lambda \subseteq \mathbb{N}$ ,  $y_n \rightarrow y \in A \cap B$ ,  $n \rightarrow \infty$ ,  $n \in \Lambda$ . Dies ist aber ein Widerspruch.

Satz 4.11 liefert die Existenz einer  $A_\varepsilon$  und  $C_\varepsilon$  trennenden, abgeschlossenen Hyperebene, d.h. es gibt ein  $f \in V^*$  und  $\alpha \in \mathbb{R}$  mit

$$f(x + \varepsilon z) \leq \alpha \leq f(y + \varepsilon z), \quad \forall x \in A, y \in C, z \in B_1(0).$$



Somit gilt, da sowohl  $z$  als auch  $-z$  gewählt werden können

$$f(x) + \varepsilon \|f\| \leq \alpha \leq f(y) - \varepsilon \|f\|, \quad \forall x \in A, y \in C,$$

d.h.  $A$  und  $C$  sind strikt getrennt. ■

• Die analytische Form des Satzes von Hahn-Banach lässt sich mit Hilfe der geometrischen Form beweisen - siehe Köthe, Topologische lineare Räume, Kap. 17.2. Der Satz ist dort noch erheblich allgemeiner dargestellt.

Eine Reihe von Sätzen über Hilberträume - z.B. über die normale Lösbarkeit linearer Gleichungen - lässt sich - abgeändert - auf Banach-Räume  $B$  übertragen, wenn man mit dem Raumpaars  $B, B^*$  und der Wirkung  $\langle f, u \rangle$  eines Elementes  $f \in B^*$  auf  $u \in B$  anstelle des Skalarproduktes arbeitet. Hierfür (und für vieles andere) sind die anschließenden Folgerungen aus dem Satz von Hahn-Banach wichtig.

Wir erinnern an die Definition  $\|f\|_{V^*} = \sup_{\substack{\|u\|_V \leq \\ u \in V}} |\langle f, u \rangle|$

**4.18 Lemma.** Sei  $V$  ein normierter Raum und  $W \subset V$  ein linearer Teilraum. Jedes lineare, stetige Funktional  $\varphi \in W^*$  lässt sich zu einem linearen, stetigen Funktional  $f \in V^*$  fortsetzen, so dass

$$\|f\|_{V^*} \leq \|\varphi\|_{W^*}.$$

BEWEIS : Man wendet Satz 4.1 mit  $p(x) = \|\varphi\|_{W^*} \|x\|$  an. ■

**4.19 Lemma.** Sei  $V$  ein normierter Vektorraum. Zu jedem  $u \in V$  existiert ein  $f \in V^*$  mit  $\|f\|_{V^*} = \|u\|_V$  und  $\langle f, u \rangle = \|u\|_V^2$ .

BEWEIS : Man wendet Lemma 4.18 mit  $W = \langle u \rangle$  und  $\varphi(tu) = t\|u\|^2$  an. Es gilt dann

$$\|\varphi\|_{W^*} = \sup \{ |t| \|u\|^2 \mid |t|\|u\| \leq 1 \} = \|u\|.$$

Nach dem Satz von Hahn-Banach gibt es eine Fortsetzung  $f \in V^*$  von  $\varphi$  mit  $\|f\| \leq \|u\|$ . Da  $f|_W = \varphi$  gilt

$$\langle f, u \rangle = \|u\|^2$$

und somit

$$\sup_{\|v\| \leq 1} |\langle f, v \rangle| = \|u\|,$$

denn

$$\langle f, v \rangle \leq \|f\| \|v\| \leq \|u\| \|v\|.$$

und für  $v = \frac{u}{\|u\|}$  wird die Gleichheit angenommen. ■

**4.20 Lemma.** Sei  $V$  normierter Vektorraum und  $u \in V$ . Aus  $\langle f, u \rangle = 0$  für alle  $f \in V^*$  folgt  $u = 0$ .

BEWEIS : Man wähle das in Lemma 4.19 konstruierte  $f$  und erhält  $0 = \langle f, u \rangle = \|u\|^2$ . ■

• Das Element  $f$  in Lemma 4.19 muss nicht eindeutig sein. Dies gilt nur, wenn die Norm in  $V^*$  **strikt konvex** ist, d.h. für jedes Paar  $f_1 \neq f_2 \in V^*$ ,  $\|f_1\| = \|f_2\| = 1$  gilt

$$\|tf_1 + (1-t)f_2\| < 1, \quad 0 < t < 1.$$

Die Abbildung, die jedem  $u \in V$  die Menge der Elemente  $f$  zuordnet, so dass die Aussagen von Lemma 4.19 gelten, nennt man die **Dualitätsabbildung**.

Nach Definition von  $\|f\|_{V^*}$ ,  $f \in V^*$ , gilt

$$\|f\|_{V^*} := \sup \{ |\langle f, u \rangle| \mid \|u\| \leq 1, u \in V \}, \quad (4.21)$$

nach Lemma 4.19 hingegen gilt

$$\|u\| = \max \{ |\langle f, u \rangle| \mid \|f\|_{V^*} \leq 1, f \in V^* \}, \quad (4.22)$$

denn es gibt ein  $f_0 \in V^*$  mit  $\|f_0\|_{V^*} = \|u\|$ ,  $\langle f_0, u \rangle = \|u\|^2$ . Das Element  $\|u\|^{-1}f_0$  realisiert das Maximum in (4.22).

Man kann beweisen, dass auch das „sup“ in (4.21) genau dann angenommen wird, wenn der Raum reflexiv ist.

## 3.5 Anwendungen des Satzes von Hahn-Banach in der konvexen Analysis

### Konjugiert-konvexe Funktionen

Im Folgenden sei  $V$  ein normierter Raum. Wie üblich, verabreden wir: Die Funktion

$$g : V \rightarrow \mathbb{R} \cup \{\infty\}$$

heißt **konvex**, wenn

$$g(tx + (1-t)y) \leq tg(x) + (1-t)g(y) \quad \text{für alle } x, y \in V, 0 \leq t \leq 1.$$

Hierbei ist  $\alpha + \infty = \infty$  gesetzt. Die Menge aller  $x \in V$  mit  $g(x) < \infty$  bezeichnen wir mit  $D(g)$ . Es ist klar, dass  $D(g)$  konvex ist.

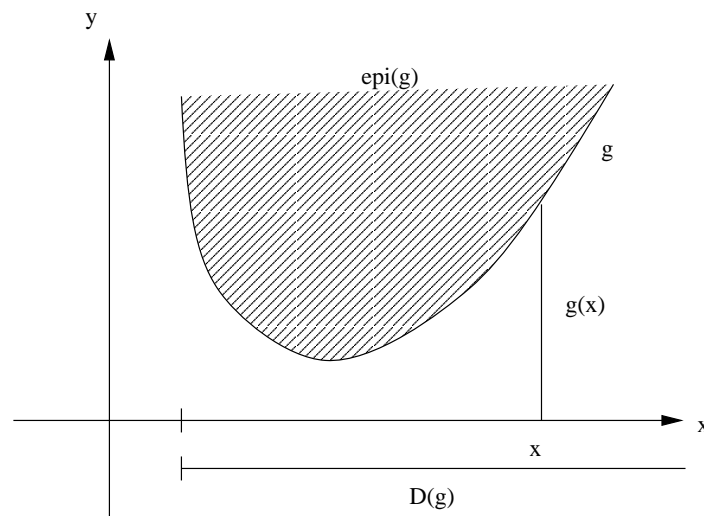
**5.1 Definition.** Die Funktion  $g : D(g) \rightarrow \mathbb{R} \cup \{\infty\}$  heißt **unterhalbstetig**, wenn für alle konvergenten Folgen  $(x_n)$  aus  $V$  mit  $x_n \rightarrow x$  die Ungleichung

$$g(x) \leq \liminf_{n \rightarrow \infty} g(x_n)$$

gilt.

**5.2 Definition.** Der **Epigraph** einer konvexen Funktion  $g : V \rightarrow \mathbb{R} \cup \{\infty\}$  ist die Menge

$$\text{epi}(g) = \{ (x, \lambda) \in V \times \mathbb{R} \mid g(x) \leq \lambda \}.$$



**5.3 Lemma.**  $g$  ist genau dann unterhalbstetig, wenn der Epigraph von  $g$  abgeschlossen ist.

BEWEIS : Übungsaufgabe. ■

• Wenn  $g$  konvex ist, ist auch  $\text{epi}(g)$  konvex; die Umkehrung ist ebenfalls richtig. (Übungsaufgabe)

**5.4 Definition.** Sei  $g : V \rightarrow \mathbb{R} \cup \{\infty\}$  und  $g \not\equiv \infty$ , d.h.  $D(g) \neq \emptyset$ . Die **duale Funktion**  $g^* : V^* \rightarrow \mathbb{R} \cup \{\infty\}$  ist definiert durch

$$g^*(f) = \sup_{u \in V} \{ \langle f, u \rangle - g(u) \}, \quad f \in V^* .$$

• Es ist klar, dass  $g^*$  konvex ist, selbst wenn  $g$  nicht konvex sein sollte.

**Beispiel:**  $V = \mathbb{R}$ ,  $g(x) = \frac{|x|^p}{p}$

$$g^*(\xi) = \sup_{x \in \mathbb{R}} \left\{ \xi x - \frac{|x|^p}{p} \right\} .$$

Das Supremum  $x^* = g^*(\xi)$  wird angenommen; es gilt dann  $\xi - |x^*|^{p-1} \text{sign } x^* = 0$  und somit

$$g^*(\xi) = |x^*|^{p-1} (\text{sign } x^*) x^* - \frac{|x^*|^p}{p} = \left( 1 - \frac{1}{p} \right) |x^*|^p .$$

Wegen  $|x^*| = |\xi|^{\frac{1}{p-1}}$  folgt

$$g^*(\xi) = \frac{|\xi|^q}{q} \quad \text{mit } q = \frac{p}{p-1} .$$

**5.5 Lemma.** Sei  $V$  ein normierter Raum und  $g : V \rightarrow \mathbb{R} \cup \{\infty\}$  konvex, unterhalbstetig und  $g \not\equiv \infty$ . Dann ist  $g^* \not\equiv \infty$ .

BEWEIS : Sei  $x_0 \in D(g)$  und  $\lambda_0 < g(x_0)$ . Wir wenden den Satz von Hahn-Banach in der Formulierung von Satz 4.17 an. Als Grundraum wählen wir  $V \times \mathbb{R}$ , als abgeschlossene, konvexe Menge  $A = \text{epi}(g)$  und  $C = (x_0, \lambda_0)$ . Nach der Definition von  $\text{epi}(g)$  ist  $A \cap C = \emptyset$ .  $A$  ist konvex und abgeschlossen und  $C$  kompakt, da letztere nur aus einem Punkt besteht. Es gibt daher eine trennende abgeschlossene Hyperebene  $H$ , die  $A$  und  $C$  strikt trennt. Die Hyperebene hat die Gestalt

$$H = \{(x, \lambda) \in V \times \mathbb{R} \mid \langle f, x \rangle + k\lambda = \alpha\}$$

mit einem  $f \in V^*$ ,  $k \in \mathbb{R}$ ,  $\alpha \in \mathbb{R}$ . Es gilt hierbei

$$\langle f, x \rangle + k\lambda > \alpha \quad \forall (x, \lambda) \in \text{epi}(g), \quad (5.6)$$

$$\langle f, x_0 \rangle + k\lambda_0 < \alpha. \quad (5.7)$$

Da  $(x_0, g(x_0)) \in \text{epi}(g)$  folgt

$$\langle f, x_0 \rangle + kg(x_0) > \alpha > \langle f, x_0 \rangle + k\lambda_0.$$

Daraus folgt  $k > 0$ . Aus (5.6) folgt mit  $\lambda = g(x)$

$$\left\langle -\frac{1}{k}f, x \right\rangle - g(x) < -\frac{\alpha}{k}, \quad x \in D(g),$$

und damit

$$g^*\left(-\frac{1}{k}f\right) \leq -\frac{\alpha}{k} < \infty. \quad \blacksquare$$

Die **biduale konvexe Funktion** zu  $g$  hat die Gestalt  $g^{**}(x) = \sup_{f \in V^*} \{\langle f, x \rangle - g^*(f)\}$ .

Hierbei ist  $\alpha - \infty = -\infty$  gesetzt.

**5.8 Satz.** Sei  $V$  ein normierter Raum und  $g : V \rightarrow \mathbb{R} \cup \{\infty\}$  konvex, unterhalbstetig und  $g \not\equiv \infty$ . Dann gilt  $g^{**} = g$ .

BEWEIS : (i) Sei zunächst  $g \geq 0$  und  $g^*(f) < \infty$  (nur solche  $f$  sind für  $g^{**}$  relevant). Da  $g(x) + g^*(f) \geq \langle f, x \rangle$  folgt  $g(x) \geq \langle f, x \rangle - g^*(f)$  und durch Übergang zum Supremum<sup>1</sup>

$$g(x) \geq g^{**}(x).$$

Angenommen, es würde  $g^{**} \neq g$  gelten. Dann gäbe es ein  $x_0$  mit

$$g^{**}(x_0) < g(x_0). \quad (5.9)$$

( $g(x_0) = \infty$  ist zugelassen, aber dann beinhaltet die Widerspruchsannahme  $g^{**}(x_0) < \infty$ .) Wir wenden den Satz von Hahn-Banach an und trennen die Menge  $\text{epi}(g)$  und den Punkt  $(x_0, g^{**}(x_0))$ , denn wegen (5.9) liegt  $(x_0, g^{**}(x_0))$  nicht in  $\text{epi}(g)$ . Ähnlich wie beim Beweis des letzten Lemmas erhält man dann ein  $f \in V^*$  und  $k \in \mathbb{R}$  sowie  $\alpha \in \mathbb{R}$  mit

$$\langle f, x \rangle + k\lambda > \alpha \quad \forall (x, \lambda) \in \text{epi}(g), \quad (5.10)$$

$$\langle f, x_0 \rangle + kg^{**}(x_0) < \alpha. \quad (5.11)$$

<sup>1</sup>Dies gilt auch ohne die Voraussetzung  $g \geq 0$ .

Lässt man in (5.10)  $\lambda$  gegen  $\infty$  gehen, folgt  $k \geq 0$ . Aus (5.10) schließt man mit  $\lambda = g(x) \geq 0$ , für  $\varepsilon > 0$

$$\langle f, x \rangle + (k + \varepsilon)g(x) > \alpha, \quad x \in D(g),$$

und damit

$$\left\langle -f \frac{1}{k+\varepsilon}, x \right\rangle - g(x) < -\frac{\alpha}{k+\varepsilon},$$

und durch Übergang zum Supremum in  $x$

$$g^*\left(-f \frac{1}{k+\varepsilon}\right) \leq -\frac{\alpha}{k+\varepsilon}.$$

Weiterhin folgt nach Definition von  $g^{**}(x_0)$

$$\begin{aligned} g^{**}(x_0) &\geq \left\langle -f \frac{1}{k+\varepsilon}, x_0 \right\rangle - g^*\left(-f \frac{1}{k+\varepsilon}\right) \\ &\geq \left\langle -f \frac{1}{k+\varepsilon}, x_0 \right\rangle + \frac{\alpha}{k+\varepsilon}, \end{aligned}$$

und somit

$$\langle f, x_0 \rangle + (k + \varepsilon)g^{**}(x_0) \geq \alpha.$$

Der Grenzübergang  $\varepsilon \rightarrow 0$  führt zu einem Widerspruch mit (5.11). Dies war der Fall  $g \geq 0$ .

(ii) Den allgemeinen Fall erhält man durch Abändern von  $g$  mit einem  $f_0 \in D(g^*)$

$$\bar{g}(x) := g(x) - \langle f_0, x \rangle + g^*(f_0) \geq 0.$$

Für  $\bar{g}$  weiß man, dass  $\bar{g}^{**} = \bar{g}$ . Es gilt

$$\begin{aligned} \bar{g}^*(f) &= \sup_{x \in V} \{ \langle f, x \rangle - \bar{g}(x) \} \\ &= \sup_{x \in V} \{ \langle f + f_0, x \rangle - g(x) - g^*(f_0) \} \\ &= g^*(f + f_0) - g^*(f_0) \\ \bar{g}^{**}(x) &= \sup_{f \in V^*} \{ \langle f + f_0, x \rangle - g^*(f + f_0) \} + g^*(f_0) - \langle f_0, x \rangle \\ &= \sup_{\tilde{f} \in V^*} \{ \langle \tilde{f}, x \rangle - g^*(\tilde{f}) \} - \langle f_0, x \rangle + g^*(f_0) \\ &= g^{**}(x) - \langle f_0, x \rangle + g^*(f_0) \end{aligned}$$

Daraus folgt  $g(x) = g^{**}(x)$ . ■

## Orlicz-Räume

**5.12 Definition.** Es sei  $g : \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+$  eine nichtnegative, konvexe Funktion und  $\Omega \subset \mathbb{R}^n$  messbar. Mit  $\tilde{L}_g$  (**Orliczklasse**) bezeichnen wir die Menge aller messbaren Funktionen  $f : \Omega \rightarrow \mathbb{R}$  mit

$$G(f) := \int_{\Omega} g(|f(x)|) dx < \infty.$$

• Die Lebesgue-Räume  $L^p$ ,  $1 < p < \infty$ , sind Spezialfälle der Orliczklassen  $\tilde{L}_g$ . Man setzt  $g(t) = ct^p$ , mit einer beliebigen positiven Konstanten  $c$ . Weitere Beispiele von Orliczklassen sind durch die Funktionen

$$\begin{aligned} g_1(t) &= t \ln^+ t, \\ g_2(t) &= (1+t) \ln(1+t) - t, \\ g_3(t) &= e^t - t - 1, \\ g_4(t) &= e^{t^2} - 1. \end{aligned}$$

gegeben. Die ersten beiden Funktionen  $g_1, g_2$  werden oft benutzt um den Raum  $L^1$  zu ersetzen und die Funktionen  $g_3, g_4$  um den Raum  $L^\infty$  zu ersetzen. Beide Räume  $L^1, L^\infty$  sind Ausnahmefälle innerhalb der Lebesgue-Räume  $L^p$ .

**5.13 Definition.** Wir sagen, dass  $g$  eine **Young-Funktion** ist, falls

$$g(t) = \int_0^t \gamma(s) ds, \quad t \geq 0.$$

Die Funktion  $\gamma : \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+$  hat die Eigenschaften

- (i)  $\gamma(0) = 0$ ,
- (ii)  $\gamma(s) > 0$ , für  $s > 0$ ,
- (iii)  $\gamma$  ist rechtsseitig stetig für alle  $s \geq 0$ ,
- (iv)  $\gamma$  ist nichtfallend auf  $(0, \infty)$ ,
- (v)  $\lim_{s \rightarrow \infty} \gamma(s) = \infty$ .

**5.14 Lemma.** Eine Young-Funktion  $g$  ist stetig, nichtnegativ, strikt wachsend und konvex auf  $[0, \infty)$ . Darüber hinaus gilt:

$$\begin{aligned} g(0) &= 0, & \lim_{t \rightarrow \infty} g(t) &= \infty, \\ \lim_{t \rightarrow 0^+} \frac{g(t)}{t} &= 0, & \lim_{t \rightarrow \infty} \frac{g(t)}{t} &= \infty, \\ g(\alpha t) &\leq \alpha g(t), & \alpha &\in [0, 1], t \geq 0, \\ g(\beta t) &\geq \beta g(t), & \beta &> 1, t \geq 0. \end{aligned}$$

BEWEIS : Übungsaufgabe ■

• Mit Hilfe von Lemma 5.14 kann man zeigen, dass

$$\tilde{L}_g(\Omega) \subseteq L^1(\Omega),$$

falls  $g$  eine Young-Funktion ist und  $\Omega$  ein beschränktes Gebiet.

• In den obigen Beispielen ist  $g_1$  keine Young-Funktion.

**5.15 Definition.** Sei  $g$  eine Young-Funktion. Der **Orlicz-Raum**  $L_g$  ist die Menge aller Äquivalenzklassen  $[f]$  von Funktionen  $f$ , für die

$$\|f\|_g = \sup \left\{ \int_{\Omega} |\varphi(x)f(x)| dx \mid G^*(\varphi) \leq 1 \right\} < \infty$$

gilt. Hierbei ist  $[f]$  die Klasse aller Funktionen, die sich von  $f$  nur auf einer Menge vom Mass Null unterscheiden, und  $G^*$  ist die durch  $G^*(\varphi) = \int_{\Omega} g^*(|\varphi(x)|) dx$  gegebene Funktion, wobei  $g^*$  die duale, konvexe Funktion zu der durch Null auf  $(-\infty, 0)$  fortgesetzten Funktion  $g$  ist. Die Größe  $\|f\|_g$  nennt man die **Orlicz-Norm** von  $f$  zur Funktion  $g$ .

Es gilt folgende Erweiterung der Young-Ungleichung:

**5.16 Lemma.** Für alle  $u, v \in [0, \infty)$  gilt:

$$uv \leq g(u) + g^*(v),$$

wobei  $g$  eine Young-Funktion ist, und  $g^*$  die zu  $g$  duale Funktion.

Mit Hilfe dieses Lemmas kann man die Hölder-Ungleichung auf Orlicz-Räume erweitern.

**5.17 Lemma.** Es sei  $g$  eine Young-Funktion und  $g^*$  die zu  $g$  duale Funktion. Sei  $u \in L_g(\Omega)$  und  $v \in L_{g^*}(\Omega)$ . Dann ist  $uv \in L^1(\Omega)$  und es gilt

$$\int_{\Omega} |uv| dx \leq \|u\|_g \|v\|_{g^*}. \quad (5.18)$$

## Der Kuhn-Tucker-Satz der konvexen Optimierung

Die Funktionen  $F : \mathbb{R}^n \rightarrow \mathbb{R}$  und  $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$  seien stetig und konvex. Wir betrachten die Optimierungsaufgabe. Gesucht ist  $x^* \in \mathbb{R}^n$ , so dass

$$F(x^*) \leq F(x) \quad \text{für alle } x \in \mathbb{R}^n \text{ mit } g(x) \leq 0.$$

Man schreibt auch

$$F(x) = \min!, \quad g(x) \leq 0, \quad x \in \mathbb{R}^n. \quad (5.19)$$

Die Menge  $M = \{x \in \mathbb{R}^n \mid g(x) \leq 0\}$  heißt **zulässige Menge** (der Optimierungsaufgabe).

Ähnlich wie in der Theorie der Lagrange-Multiplikatoren für Extrema mit Nebenbedingungen und differenzierbaren Funktionen lässt sich ein entsprechender Satz für konvexe, nicht differenzierbare Funktionen  $F$  und  $g$  herleiten. Dies geschieht - wenn man will - über den Satz von Hahn-Banach.

Die **Lagrange-Funktion** zum Problem (5.19) lautet

$$L(x, \lambda) = F(x) + \lambda \cdot g(x), \quad x \in \mathbb{R}^n, \quad \lambda \in \mathbb{R}^m.$$

Für  $\lambda \geq 0$  (d.h. die Komponenten  $\lambda_i$  von  $\lambda$  sind  $\geq 0$ ) ist  $L(x, \lambda)$  konvex in  $x$ .

**5.20 Definition.** Ein **Sattelpunkt** der Lagrange-Funktion  $L$  zu (5.19) ist ein Paar  $x^*, \lambda^* \in \mathbb{R}^n \times \mathbb{R}^m$  mit  $\lambda^* \geq 0$ , und

$$L(x^*, \lambda) \leq L(x^*, \lambda^*) \leq L(x, \lambda^*) \quad (5.21)$$

für alle  $x \in \mathbb{R}^n$ ,  $\lambda \in \mathbb{R}^m$ ,  $\lambda \geq 0$ .

Wir werden mit Hilfe des Satzes von Hahn-Banach unter einer Zusatzvoraussetzung an  $M$  einen Sattelpunkt konstruieren. Der Zusammenhang mit der Optimierungsaufgabe (5.19) ergibt sich durch den folgenden Satz:

**5.22 Satz.** Ist  $(x^*, \lambda^*)$  ein Sattelpunkt von  $L$ , so ist  $x^*$  eine Lösung der Aufgabe (5.19). Sind überdies  $F$  und  $g$  differenzierbar, so gilt

$$\nabla_x F(x^*) + \sum_{j=1}^m \lambda_j^* \nabla g_j(x^*) = 0. \quad (5.23)$$

• Offensichtlich entspricht (5.23) der aus der Analysis unter anderen Voraussetzungen bekannten Aussage zur Existenz von Lagrange-Multiplikatoren.

BEWEIS <sup>2</sup>: Die Ungleichung (5.21) bedeutet

$$F(x^*) + \lambda \cdot g(x^*) \leq F(x^*) + \lambda^* \cdot g(x^*) \leq F(x) + \lambda^* \cdot g(x) \quad (5.24)$$

für  $\lambda \in \mathbb{R}^m$ ,  $\lambda \geq 0$  und  $x \in \mathbb{R}^n$ . Aus der ersten Ungleichung folgt

$$(\lambda - \lambda^*) \cdot g(x^*) \leq 0, \quad \lambda \geq 0, \quad \lambda \in \mathbb{R}^m.$$

Lässt man die Komponenten von  $\lambda$  gegen  $+\infty$  konvergieren, folgt  $g(x^*) \leq 0$ , d.h.  $x^* \in M$ , die Nebenbedingung ist für  $x^*$  erfüllt. Da  $\lambda^* \geq 0$ , folgt

$$\lambda^* \cdot g(x^*) \leq 0,$$

und setzt man in der ersten Ungleichung in (5.24)  $\lambda = 0$ , folgt

$$F(x^*) \leq F(x^*) + \lambda^* \cdot g(x^*),$$

also  $\lambda^* \cdot g(x^*) \geq 0$ . Damit folgt

$$\lambda^* \cdot g(x^*) = 0. \quad (5.25)$$

Aus der zweiten Ungleichung in (5.24) folgt dann

$$F(x^*) \leq F(x) + \lambda^* \cdot g(x) \leq F(x),$$

da  $\lambda^* \geq 0$ ,  $g(x) \leq 0$ . Der Punkt  $x^*$  ist also tatsächlich Minimalstelle. ■

<sup>2</sup>Der Satz von Hahn-Banach wird hier noch nicht verwendet.



Die Beziehung (5.25) wird als besonderes Lemma festgehalten:

**5.26 Lemma („Gleichgewichtssatz“).** Sei  $(x^*, \lambda^*)$  Sattelpunkt der Lagrange-Funktion  $L$  von (5.19). Dann gilt

$$\lambda^* \cdot g(x^*) = 0.$$

Die Bedingung (5.23) heisst auch „dritte Kuhn-Tucker-Bedingung“, die Bedingung (5.25) „zweite Kuhn-Tucker-Bedingung“. Die Nebenbedingung  $g(x) \leq 0$  und die Vorzeichenrestriktion  $\lambda \geq 0$  nennt man die „erste Kuhn-Tucker-Bedingung“.

Um zu einer Lösung  $x^*$  des Optimierungsproblems (5.19) ein  $\lambda^* \in \mathbb{R}^m$ ,  $\lambda^* \geq 0$ , zu konstruieren, so dass  $(x^*, \lambda^*)$  Sattelpunkt ist, benötigt man eine Zusatzbedingung, die manchmal als **Slater-Bedingung** in den Optimierungsbüchern herumgeistert:

$$\text{Es gibt einen „inneren Punkt“ der Menge } M. \quad (5.27)$$

Analytisch heisst dies: Es existiert  $x_0 \in \mathbb{R}^n$  mit  $g_j(x_0) < 0$ ,  $j = 1, \dots, m$ .

**5.28 Satz (Kuhn-Tucker-Satz der konvexen Optimierung).** Sei  $F : \mathbb{R}^n \rightarrow \mathbb{R}$  und  $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$  konvex und stetig. Die Menge  $M = (g \leq 0)$  besitze einen inneren Punkt, d.h. es gibt ein  $x_0$  mit  $g(x_0) < 0$ . Dann ist ein Punkt  $x^* \in \mathbb{R}^n$  genau dann Lösung der Optimierungsaufgabe (5.19), wenn es ein  $\lambda^* \in \mathbb{R}^m$ ,  $\lambda^* \geq 0$  gibt, so dass  $(x^*, \lambda^*)$  Sattelpunkt der Lagrange-Funktion  $L$  zu (5.19) ist.

BEWEIS : Die eine Richtung des Satzes ist bereits in Satz 5.22 festgehalten. Für die Umkehrung müssen wir den Multiplikator  $\lambda^*$  konstruieren. Hierzu betrachten wir die Mengen

$$A = \left\{ y = (y_0, \dots, y_m) \in \mathbb{R}^{m+1} \mid \begin{array}{l} y_0 \geq F(x), \ y_j \geq g_j(x), \ j = 1, \dots, m \\ \text{für mindestens ein } x \in \mathbb{R}^n \end{array} \right\}$$

sowie

$$B = \left\{ y = (y_0, \dots, y_m) \in \mathbb{R}^{m+1} \mid y_0 < F(x^*), \ y_j < 0, \ j = 1, \dots, m \right\}.$$

Die Mengen  $A$  und  $B$  sind konvex. (Für  $B$  ist dies offensichtlich, für  $A$  beachte man, dass  $F$  und  $g$  konvex sind.) Die Mengen  $A$  und  $B$  sind disjunkt, da es sonst ein  $y \in \mathbb{R}^{m+1}$  gäbe mit  $F(x) \leq y_0 < F(x^*)$ ,  $g_j(x) \leq y_j < 0$  für ein  $x \in \mathbb{R}^n$ , so dass  $x^*$  nicht minimal wäre.  $B$  ist offen,  $A$  und  $B$  sind offensichtlich nicht leer. Damit lässt sich der Satz von Hahn-Banach über die Trennung konvexer Mengen (Satz 4.11, vgl. (4.16)) anwenden und es gibt  $\mu \in \mathbb{R}^{m+1}$ ,  $\mu \neq 0$ , so dass

$$\mu \cdot y > \mu \cdot z \quad \text{für alle } y \in A, \ z \in B. \quad (5.29)$$

Lässt man die Komponenten von  $z$  nach  $-\infty$  konvergieren (beachte die Definition von  $B$ ), ergibt sich aus (5.29)

$$\mu \geq 0.$$

Aus (5.29) folgt für  $y \in A, z \in \overline{B}$  (Abschluss von  $B$ )

$$\mu \cdot y \geq \mu \cdot z \quad (5.30)$$

Aus (5.30) ergibt sich mit

$$z = (F(x^*), 0, \dots, 0) \in \overline{B}$$

und

$$y = (F(x), g_1(x), \dots, g_m(x)) \in A.$$

$$\mu_0 F(x) + \sum_{j=1}^m \mu_j g_j(x) \geq \mu_0 F(x^*), \quad x \in \mathbb{R}^n. \quad (5.31)$$

Daraus entnimmt man  $\mu_0 > 0$ . Andernfalls wäre  $\sum_{j=1}^m \mu_j g_j(x) \geq 0$ ,  $x \in \mathbb{R}^n$ , und insbesondere

$$\sum_{j=1}^m \mu_j g_j(x_0) \geq 0$$

mit dem Punkt  $x_0$  aus der Slater-Bedingung  $g(x_0) < 0$ . Da  $\mu_j \geq 0$ , ergäbe sich ein Widerspruch. Wir setzen

$$\lambda^* = \frac{1}{\mu_0} (\mu_1, \dots, \mu_m).$$

Es gilt  $\lambda^* \geq 0$ , und wegen (5.31)

$$F(x) + \lambda^* \cdot g(x) \geq F(x^*), \quad x \in \mathbb{R}^n. \quad (5.32)$$

Wir setzen  $x = x^*$  und erhalten  $\lambda^* \cdot g(x^*) \geq 0$ , da  $x^*$  zulässig, gilt  $g(x^*) \leq 0$ , somit  $\lambda^* \cdot g(x^*) \leq 0$  und

$$\lambda^* \cdot g(x^*) = 0. \quad (5.33)$$

Aus (5.32) und (5.33) ergibt sich die eine der beiden Ungleichungen der Sattelpunktbedingung

$$F(x) + \lambda^* \cdot g(x) \geq F(x^*) + \lambda^* \cdot g(x^*).$$

Die weitere Ungleichung

$$F(x^*) + \lambda^* \cdot g(x^*) = F(x^*) \geq F(x^*) + \lambda \cdot g(x^*).$$

ist trivial, da  $g(x^*) \leq 0$  und  $\lambda \geq 0$  verlangt ist. ■

## Alternativsätze für lineare Ungleichungssysteme

**5.34 Satz.** Sei  $A$  eine  $m \times n$ -Matrix und  $b \in \mathbb{R}^m$ . Entweder besitzt

$$Ax = b, \quad x \geq 0 \quad (5.35)$$

eine Lösung  $x \in \mathbb{R}^n$ , oder

$$A^T y \geq 0, \quad b \cdot y < 0 \quad (5.36)$$

besitzt eine Lösung  $y \in \mathbb{R}^m$ .

BEWEIS : (i) (5.35) und (5.36) können nicht gleichzeitig lösbar sein, denn andernfalls existierte  $x \in \mathbb{R}^n$  und  $y \in \mathbb{R}^m$  mit

$$0 > b \cdot y = (Ax) \cdot y = x^T \cdot A^T y \geq 0.$$

(ii) Ist (5.35) unlösbar, so liegt  $b$  nicht in dem von den Vektoren  $a^{(1)}, \dots, a^{(n)} \in \mathbb{R}^m$  aufgespannten abgeschlossenen Kegel

$$K = \left\{ \sum_{j=1}^n \xi_j a^{(j)} \mid \xi_j \geq 0 \right\}.$$

Es gibt dann auch eine abgeschlossene Kugel  $B_\varepsilon(b)$  um  $b$ , so dass

$$K \cap B_\varepsilon(b) = \emptyset,$$

und für den von  $B_\varepsilon(b)$  aufgespannten Kegel

$$K_b = \{ \alpha y \mid y \in B_\varepsilon(b), \alpha \geq 0 \}$$

gilt

$$\text{int } K_b \cap K = \emptyset.$$

Da  $K$  konvex und abgeschlossen,  $\text{int } K_b$  konvex und offen, lässt sich der Satz von Hahn-Banach (Satz 4.11) anwenden. Es gibt daher eine Hyperebene, welche  $K$  und  $\text{int } K_b$  trennt, d.h. es existiert  $a \in \mathbb{R}^m$ ,  $a \neq 0$  mit (vgl. (4.16))

$$a \cdot y < a \cdot z, \quad y \in K, \quad z \in \text{int } K_b. \quad (5.37)$$

Da  $0 \in K$  und  $b \in \text{int } K_b$ , folgt

$$a \cdot b > 0. \quad (5.38)$$

Da  $0 \in \overline{\text{int } K_b} = K_b$ , folgt

$$a \cdot y \leq 0, \quad y \in K,$$

insbesondere

$$a \cdot a^{(j)} \leq 0, \quad j = 1, \dots, n.$$

Dies bedeutet

$$A^T a \leq 0. \quad (5.39)$$

Mit  $-a = y$  ergibt sich aus (5.38) und (5.39)

$$A^T y \geq 0, \quad y \cdot b < 0,$$

d.h. eine Lösung von (5.36). ■

### 3.6 Adjungierte lineare Operatoren in Banach-Räumen

Es seien  $E$  und  $F$  Banachräume. Wir betrachten nicht notwendig beschränkte, lineare Operatoren  $A : D(A) \rightarrow F$  mit einem in  $E$  dichten linearen Unterraum  $D(A)$ . Wie in der internationalen Literatur üblich, wollen wir den Bildbereich von  $A$  mit  $R(A)$  bezeichnen:

$$R(A) := A(D(A)) \subset F,$$

der Kern ist definiert durch

$$N(A) := \{x \in D(A) \mid Ax = 0\}.$$

Der Vektorraum der beschränkten, linearen Funktionale auf  $E$  bzw.  $F$  wird wieder mit  $E^*$  bzw.  $F^*$  bezeichnet.

**6.1 Definition.** Sei  $A : D(A) \subset E \rightarrow F$  ein linearer Operator und sei  $D(A)$  ein dichter linearer Unterraum von  $E$ . Sei

$$D(A^*) := \{v \in F^* \mid \exists c \geq 0, \text{ so dass } |\langle v, Au \rangle| \leq c \|u\| \text{ für alle } u \in D(A)\}.$$

Ist nun  $v \in D(A^*)$ , so definiert  $\varphi(u) = \langle v, Au \rangle$  ein **beschränktes lineares Funktional** auf  $D(A)$ , welches durch Abschluss auf ganz  $E$  zu einem Element  $f \in E^*$  fortgesetzt werden kann:

$$\langle f, u \rangle = \varphi(u) = \langle v, Au \rangle \quad u \in D(A).$$

Man definiert den **adjungierten Operator**  $A^* : D(A^*) \subseteq F^* \rightarrow E^*$  durch

$$A^*v := f.$$

Es ist klar, dass  $f$  eindeutig ist, denn aus  $\langle f, u \rangle = \langle \tilde{f}, u \rangle = \varphi(u)$  für  $u \in D(A)$  folgt  $f = \tilde{f}$ , da  $u$  dicht in  $E$  ist und  $f, \tilde{f}$  stetig sind. Ebenso ist die Linearität von  $A^*$  und  $D(A^*)$  klar.

Für das Paar  $A, A^*$  gelten ähnliche Alternativsätze zur Lösbarkeit der Gleichung  $Au = f$  wie im Fall des Hilbertraumes. Wir behandeln sogar nicht notwendig beschränkte, jedoch wenigstens *abgeschlossene* lineare Operatoren  $A$ . Für die meisten Anwendungen reicht dies aus. Wir bereiten dies durch einen Satz und ein Lemma vor. Hierfür benötigen wir die folgende Definitionen:

**6.2 Definition.** Sei  $A : D(A) \subseteq E \rightarrow F$  linear und sei  $D(A)$  dicht in  $E$ . Dann heißt  $A$  **abgeschlossen**, wenn der Graph

$$G(A) \equiv \{(v, Av) \mid v \in D(A)\}$$

abgeschlossen in  $E \times F$  ist.

**6.3 Definition.** Sei  $B$  ein Banachraum. Das **Orthogonalkomplement** einer Menge  $M \subset B$  ist die Menge

$$M^\perp := \{f \in B^* \mid \langle f, m \rangle = 0 \quad \text{für alle } m \in M\}.$$

Das **Biorthogonalkomplement**  $M^{\perp\perp}$  ist die Menge

$$M^{\perp\perp} = \{u \in B \mid \langle f, u \rangle = 0 \quad \text{für alle } f \in M^\perp\}.$$

**6.4 Satz.** Sei  $B$  ein Banachraum und  $W \subset B$  ein linearer Teilraum. Dann gilt

$$W^{\perp\perp} = \overline{W}.$$

BEWEIS : Es ist klar, dass  $W \subset W^{\perp\perp}$ , denn ist  $w \in W$ , so gilt  $\langle f, w \rangle = 0$  für alle  $f \in W^\perp$ . Dies impliziert  $w \in W^{\perp\perp}$ . Ferner ist klar, dass  $W^{\perp\perp}$  abgeschlossen ist, da  $M^\perp$  für jede Menge  $M$  abgeschlossen ist. In der Tat, wenn  $f_j \rightarrow f$ ,  $f_j \in M^\perp$ , folgt  $0 = \langle f_j, m \rangle \rightarrow \langle f, m \rangle$ , d.h.  $f \in M^\perp$ , und entsprechend für  $M^{\perp\perp}$ .

Da  $W \subset W^{\perp\perp}$ , gilt somit  $\overline{W} \subset W^{\perp\perp}$ . Angenommen,  $\overline{W}$  wäre echt in  $W^{\perp\perp}$  enthalten. Dann gibt es ein  $x_0 \in W^{\perp\perp}$  mit  $x_0 \notin \overline{W}$ . Nach dem Satz von Hahn-Banach (Satz 4.17) gibt es eine abgeschlossene Hyperebene, die  $x_0$  und  $\overline{W}$  trennt, d.h. es existiert  $f \in B^*$  und  $\alpha \in \mathbb{R}$  mit

$$\langle f, x \rangle < \alpha < \langle f, x_0 \rangle, \quad x \in \overline{W}.$$

Da  $\overline{W}$  ein linearer Teilraum ist, folgt  $\langle f, x \rangle = 0$  für alle  $x \in \overline{W}$ , d.h.  $f \in \overline{W}^\perp$ . Da andererseits  $x_0 \in W^{\perp\perp}$ , gilt dann  $\langle f, x_0 \rangle = 0$ . Dies ist ein Widerspruch, und der Satz ist bewiesen. ■

**6.5 Lemma.** Sei  $A : D(A) \subset E \rightarrow F$  linear,  $\overline{D(A)} = E$ . Dann ist  $A^*$  abgeschlossen, d.h. der Graph  $G(A^*) := \{(v, A^*v) \mid v \in D(A^*)\}$  ist abgeschlossen in  $F^* \times E^*$ .

BEWEIS : Sei  $v_n \rightarrow v$ ,  $A^*v_n \rightarrow f$ ,  $v_n \in D(A^*)$ , in  $F^*$  bzw.  $E^*$ . Wir müssen zeigen, dass

$$v \in D(A^*) \quad \text{und} \quad A^*v = f.$$

Es gilt aber:

$$\langle v_n, Au \rangle = \langle A^*v_n, u \rangle, \quad u \in D(A),$$

nach Grenzübergang  $n \rightarrow \infty$

$$\langle v, Au \rangle = \langle f, u \rangle,$$

woraus man  $v \in D(A^*)$  und  $A^*v = f$  abliest. ■

**6.6 Satz.** Sei  $A : D(A) \subset E \rightarrow F$  abgeschlossen und linear mit  $\overline{D(A)} = E$ . Dann gilt:

- (i)  $N(A^*) = R(A)^\perp$ ,
- (ii)  $N(A) = R(A^*)^\perp$ ,
- (iii)  $N(A^*)^\perp \supset R(A)$ ,
- (iv)  $N(A)^\perp \supset R(A^*)$ .

BEWEIS : Ein genialer Trick, den man in dem Buch „Analyse fonctionelle“ von Brezis findet, macht den Beweis einfach: (i) Wir betrachten den Graphen

$$G = G(A) = \{(u, Au) \in E \times F \mid u \in D(A)\}$$

und den Raum  $L = E \times \{0\} \subset E \times F$ . Man beachte, dass  $G$  und  $L$  in  $E \times F$  abgeschlossen sind. Die Beweisidee besteht darin,  $(G + L)^\perp$  auszurechnen:

$$\begin{aligned} (G + L)^\perp &= \{(u + v, Au) \in E \times F \mid u \in D(A), v \in E\}^\perp \\ &= \{(w, Au) \in E \times F \mid u \in D(A), w \in E\}^\perp \\ &= \{(0, f^*) \in E^* \times F^* \mid f^* \perp Au, u \in D(A)\} \\ &= \{0\} \times R(A)^\perp. \end{aligned}$$

Andererseits gilt ganz allgemein für die Summe zweier Räume

$$(G + L)^\perp = G^\perp \cap L^\perp,$$

denn

$$(G + L)^\perp = \{z + y \mid z \in G, y \in L\}^\perp = \{\varphi^* \mid \langle \varphi^*, z \rangle + \langle \varphi^*, y \rangle = 0, z \in G, y \in L\}.$$

Setzt man  $z$  bzw.  $y = 0$ , erhält man, dass das linear beschränkte Funktional  $\varphi^*$  sowohl in  $G^\perp$  als auch in  $L^\perp$  liegt. Umgekehrt erfüllen diese  $\varphi^* \in G^\perp \cap L^\perp$  die verlangte Orthogonalität zu  $z + y$ .

Wir erhalten somit

$$\begin{aligned} \{0\} \times R(A)^\perp &= G^\perp \cap L^\perp \\ &= \{(e^*, f^*) \in E^* \times F^* \mid \langle e^*, u \rangle + \langle f^*, Au \rangle = 0 \text{ und} \\ &\quad \langle e^*, v \rangle + \langle f^*, 0 \rangle = 0, u \in D(A), v \in E\} \end{aligned}$$

Falls aber  $\langle e^*, v \rangle + \langle f^*, 0 \rangle = 0$  für alle  $v \in E$ , folgt  $e^* = 0$ , und die letzte Menge  $\{\dots\}$  ist gleich der Menge

$$\{(0, f^*) \in E^* \times F^* \mid \langle f^*, Au \rangle = 0, u \in D(A)\}$$

und damit gleich

$$\{(0, f^*) \mid A^* f^* = 0\} = \{0\} \times N(A^*).$$

Damit erhalten wir

$$R(A)^\perp = N(A^*), \tag{6.7}$$

d.h. die Behauptung ist damit bewiesen.<sup>3</sup>

<sup>3</sup>Die folgende simple Überlegung reicht nicht zum Nachweis von (6.7). Falls  $f = Au \in R(A)$ , folgt für  $f^* \in N(A^*)$

$$\langle f^*, f \rangle = \langle f^*, Au \rangle = \langle A^* f^*, u \rangle = 0,$$

d.h.  $R(A) \perp N(A^*)$ . Daraus folgt nur

$$N(A^*) \subset R(A)^\perp.$$

Es könnte ja noch andere Elemente außerhalb von  $N(A^*)$  geben, die orthogonal zu  $R(A)$  sind.

(ii) Wir betrachten den Ausdruck  $(G^\perp + L^\perp)^\perp$ ,  $G$  und  $L$  wie im Beweis für (i). Einerseits ist

$$\begin{aligned} (G^\perp + L^\perp)^\perp &= G^{\perp\perp} \cap L^{\perp\perp} = G \cap L \\ &= \{(u, Au) \mid u \in D(A)\} \cap \{(v, 0) \mid v \in E\} \\ &= \{(u, Au) \mid Au = 0, u \in D(A)\} \\ &= N(A) \times \{0\}, \end{aligned} \tag{6.8}$$

da  $G$  und  $L$  abgeschlossen sind. Andererseits gelten

$$L^\perp = \{(0, h^*) \in E^* \times F^* \mid h^* \in F^*\},$$

$$G^\perp + L^\perp = \{(e^*, f^*) + (0, h^*) \in E^* \times F^* \mid \langle e^*, u \rangle + \langle f^*, Au \rangle = 0, u \in D(A), h^* \in F^*\}.$$

Da alle  $h^* \in F^*$  zugelassen sind, folgt

$$G^\perp + L^\perp = \{(e^*, g^*) \mid g^* \in F^*, \text{ zu } e^* \text{ existiert } f^*, \text{ mit } \langle e^*, u \rangle + \langle f^*, Au \rangle = 0, u \in D(A)\}.$$

Wenn für ein Paar  $(e^*, f^*)$  die Gleichung  $\langle e^*, u \rangle + \langle f^*, Au \rangle = 0$  für alle  $u \in D(A)$  gilt, so ist  $f^* \in D(A^*)$  und

$$\langle e^* + A^* f^*, u \rangle = 0, \quad u \in D(A).$$

Daraus folgt  $e^* = -A^* f^*$ . Umgekehrt impliziert die letzte Gleichung, dass

$$\langle e^*, u \rangle + \langle f^*, Au \rangle = 0.$$

Daraus folgt

$$\begin{aligned} G^\perp + L^\perp &= \{(-A^* f^*, g^*) \mid f^* \in D(A^*), g^* \in F^*\} \\ &= R(A^*) \times F^*. \end{aligned}$$

Anwendung der  $\perp$ -Operation ergibt

$$(G^\perp + L^\perp)^\perp = R(A^*)^\perp \times \{0\}.$$

Zusammen mit (6.8) folgt

$$N(A) \times \{0\} = R(A^*)^\perp \times \{0\}$$

und damit  $N(A) = R(A^*)^\perp$ .

(iii) Aus (i) folgt

$$N(A^*)^\perp = R(A)^{\perp\perp},$$

und nach Satz 6.5 gilt  $R(A)^{\perp\perp} = \overline{R(A)} \supset R(A)$ .

Ähnlich schliesst man für (iv). ■

Wie im Fall des Hilbertraumes (vergleiche Kapitel 2, Folge 5.6) lässt sich auch die Aussage des vorigen Satzes verschärfen, wenn man die Abgeschlossenheit des Bildbereiches  $R(A)$  fordert:

**6.9 Satz.** Sei  $A : D(A) \subset E \rightarrow F$  abgeschlossen, linear und  $D(A)$  dicht in  $E$ ,  $E$  und  $F$  Banachräume. Dann sind die folgenden Eigenschaften äquivalent:

(i)  $R(A)$  ist abgeschlossen.

(ii)  $R(A^*)$  ist abgeschlossen.

(iii)  $R(A) = N(A^*)^\perp$ .

(iv)  $R(A^*) = N(A)^\perp$ .

• Die Behauptung (iii) lässt sich so lesen:  $Au = f$  ist genau dann lösbar, wenn  $\langle v, f \rangle = 0$  für alle  $v \in N(A^*)$ .

BEWEIS : Wir begnügen uns damit, aus der Abgeschlossenheit von  $R(A)$  zu folgern, dass  $R(A) = N(A^*)^\perp$  ist. Dies folgt einfach aus Satz 6.6, (i), durch Anwendung von  $\perp$ , welches  $R(A)^{\perp\perp} = N(A^*)^\perp$  ergibt. Wenn  $R(A)$  abgeschlossen, ist  $R(A)^{\perp\perp} = R(A)$ , woraus (iii) folgt. Analog folgt aus (ii) die Behauptung (iv). Die fehlenden Implikationen können in Brezis: „Analyse fonctionelle“ gefunden werden. ■

**Beispiel:** Sei  $I = [\alpha, \beta]$  ein beschränktes Intervall des  $\mathbb{R}$  und  $E = L^p(I)$ ,  $F = L^p(I)$ , mit  $p \in (1, \infty)$ . Sei  $D(A) = H^{2,p}(I) \cap H_0^{1,p}(I)$ . Offensichtlich ist  $D(A)$  dicht in  $L^p$  (bezüglich der  $L^p$ -Norm). Der Operator  $A : D(A) \rightarrow L^p$  sei definiert durch

$$Au = -u'' + cu$$

mit einer vorgegebenen  $L^\infty$ -Funktion  $c$ . Für  $u \in H^{2,p}$  und  $c \in L^\infty$  ist  $cu \in L^p$ , so dass der Störterm  $cu$  die Inklusion  $Au \in L^p$  nicht verfälscht.

Man überlegt sich leicht, dass der Graph von  $A$  abgeschlossen ist: Aus  $Au_m \rightarrow f$  in  $L^p$ ,  $u_m \rightarrow u$  in  $L^p$  folgert man leicht, dass  $u_m$  in  $H^{2,p}$  konvergent ist. Wegen der Vollständigkeit von  $H^{2,p}$  folgt  $u \in H^{2,p}$  und anschliessend  $Au_m \rightarrow Au$ . Wir können damit aussagen, dass der Kern des adjungierten Operators orthogonal zu  $R(A)$  ist und dass die Elemente  $N(A^*)$  auch die einzigen sind, die zu  $R(A)$  orthogonal sind.

Wir wollen den Kern von  $A^*$  bestimmen. Wir wissen, dass

$$A^* : D(A^*) \subset L^q(I) \rightarrow L^q(I), \quad q = \frac{p}{p-1},$$

abbildet und dass  $N(A^*)$  aus allen Elementen  $g \in L^q$  besteht, für die

$$\int_{\alpha}^{\beta} g(-u'' + cu) dt = 0 \quad \text{für alle } u \in D(A).$$

Ähnlich wie im Hilbertraum schließt man damit zunächst, dass  $g$  zweite verallgemeinerte Ableitungen hat und dass  $g'' \in L^q$ ,  $q = (p-1)/p$ . Es folgt dann weiter

$$-g'' + cg = 0 \quad \text{in } I \tag{6.10}$$



und

$$g(\alpha) = g(\beta) = 0. \quad (6.11)$$

Der Kern von  $A^*$  besteht also aus den Lösungen des Randwertproblems (6.10), (6.11). Weiter gilt: Ist das Randwertproblem

$$-u'' + cu = f \in L^p, \quad u(\alpha) = u(\beta) = 0 \quad (6.12)$$

lösbar, so muss  $f$  die Orthogonalitätsrelation („Paarung von  $(L^q, L^p)$ “) erfüllen.

$$\int_{\alpha}^{\beta} fg \, dx = 0$$

erfüllen. Dass dies auch notwendig für die Lösbarkeit von (6.12) ist, folgt aus Satz 6.9,(iii), wofür man die Abgeschlossenheit von  $R(A)$  beweisen muss. Der Nachweis geschieht ähnlich wie im Hilbertraum, indem  $A$  als Summe eines koerziven und eines kompakten Operators geschrieben wird.

## 3.7 Schwache Topologie und reflexive Räume

Sei  $E$  ein Banachraum und  $E^*$  sein Dualraum. Dann gibt es im Raum  $E^*$  zwei verschiedene schwache Konvergenzen für eine Folge  $(f_n) \subseteq E^*$ , nämlich

$$f_n \rightharpoonup f \quad \text{schwach in } E^*,$$

d.h. für alle  $g \in E^{**}$  gilt:  $\langle g, f_n \rangle_{E^{**}, E^*} \rightarrow \langle g, f \rangle_{E^{**}, E^*}$  und

$$f_n \xrightarrow{*} f \quad * \text{-schwach in } E^*,$$

d.h. für alle  $x \in E$  gilt:  $\langle f_n, x \rangle_{E^*, E} \rightarrow \langle f, x \rangle_{E^*, E}$ .

Wir werden nun die dazugehörigen Topologien untersuchen und die Zusammenhänge zwischen  $E^{**}$  und  $E$  genauer beleuchten.

**7.1 Konstruktion von Topologien.** Sei  $X$  ein Vektorraum und seien  $(Y_i)_{i \in I}$  topologische Räume. Für alle  $i \in I$  seien  $\varphi_i : X \rightarrow Y_i$  Abbildungen. Wir suchen die *größte* Topologie  $\tau$  auf  $X$ , so daß alle  $\varphi_i$  stetig sind, d.h. die Topologie, die am wenigsten offene Mengen enthält. Seien  $\omega_i \subseteq Y_i$  offene Mengen, dann sind notwendigerweise alle  $\varphi_i^{-1}(\omega_i)$  Elemente von  $\tau$ .

Wir bezeichnen die Familie aller solcher Mengen mit  $(U_\lambda)_{\lambda \in \Lambda}$ . Wir suchen nun das kleinste Mengensystem  $\tau$ , daß  $(U_\lambda)_{\lambda \in \Lambda}$  enthält und abgeschlossen bzgl. endlicher Durchschnitte und beliebiger Vereinigungen ist. Dazu bilden wir *zuerst* alle möglichen endlichen Durchschnitte von Mengen aus  $(U_\lambda)_{\lambda \in \Lambda}$ , d.h.  $\bigcap_{\lambda \in \Gamma} U_\lambda, \Gamma \subseteq \Lambda, \Gamma$  endlich.

Dieses System bezeichnen wir mit  $\Phi$ . Danach bilden wir beliebige Vereinigungen von Mengen aus  $\Phi$ . Dieses neue System sei  $\mathcal{F}$ . Es ist klar, daß  $\mathcal{F}$  abgeschlossen bzgl. beliebigen Vereinigungen ist. Es ist nicht völlig offensichtlich, daß  $\mathcal{F}$  auch abgeschlossen bzgl. endlicher Durchschnitte ist. Aber es gilt:

**7.2 Lemma.** *Das System  $\mathcal{F}$  ist abgeschlossen bzgl. endlicher Durchschnitte.*

BEWEIS : Übungsaufgabe aus der Mengentheorie. ■

Somit ist die gesuchte größte Topologie  $\tau$  gegeben durch endliche Durchschnitte von Mengen der Form  $\varphi_i^{-1}(\omega_i)$ ,  $\omega_i$  offen in  $Y_i$ , und beliebigen Vereinigungen solcher Mengen. In Termen von Umgebungen ausgedrückt heißt dies: sei  $x \in X$ , dann ist eine Umgebungsbasis von  $x$  bzgl.  $\tau$  gegeben durch endliche Durchschnitte der Form  $\varphi_i^{-1}(V_i)$ ,  $V_i$  Umgebung von  $\varphi_i(x_i)$  in  $Y_i$ .

Die Konvergenz von Folgen in der Topologie  $\tau$  ist vollständig durch die Abbildungen  $\varphi_i$ ,  $i \in I$  charakterisiert.

**7.3 Lemma.** *Sei  $(x_n)$  eine Folge in  $X$ . Die Folge  $x_n$  konvergiert gegen  $x \in X$  bzgl.  $\tau$  genau dann wenn  $\varphi_i(x_n) \rightarrow \varphi_i(x)$  für alle  $i \in I$ .*

BEWEIS : Aus  $x_n \rightarrow x$  bzgl.  $\tau$  folgt sofort  $\varphi_i(x_n) \rightarrow \varphi_i(x)$ , da alle  $\varphi_i$  bzgl.  $\tau$  stetig sind. Umgekehrt, sei  $U$  eine Umgebung von  $x$ . Aufgrund der Konstruktion von  $\tau$  können wir annehmen, daß  $U$  die Form  $U = \bigcap_{i \in J} \varphi_i^{-1}(V_i)$ ,  $J \subset I$ ,  $J$  endlich,  $V_i$  Umgebung von  $\varphi_i(x)$  in  $Y_i$ . Für alle  $i \in J$  existiert  $N_i$  mit  $\varphi_i(x_n) \in V_i$  für alle  $n \geq N_i$ . Sei  $N = \max N_i$ . Dann ist  $x_n \in U$  für alle  $n \geq N$ . ■

Wir wenden diese allgemeine Konstruktion nun in verschiedenen Situationen an.

Sei  $E$  ein Banachraum und sei  $f \in E^*$ . Wir definieren  $\varphi_f : E \rightarrow \mathbb{R}$  durch

$$\varphi_f(x) \equiv \langle f, x \rangle$$

und betrachten die Familie  $(\varphi_f)_{f \in E^*}$ .

**7.4 Definition.** *Sei  $E$  ein Banachraum und  $E^*$  sein Dualraum. Die **schwache Topologie**  $\tau(E, E^*)$  ist die größte Topologie bzgl. derer alle  $(\varphi_f)_{f \in E^*}$  stetig sind. (Wir setzen also  $X = E, Y_i = \mathbb{R}, I = E^*$  in 7.1).*

• Lemma 7.3 zeigt, daß die Konvergenz bzgl. der schwachen Topologie  $\tau(E, E^*)$  und die schwache Konvergenz im Sinne von Definition 2.5 identisch sind.

**7.5 Lemma.** *Sei  $x_0 \in E$ . Eine Umgebungsbasis von  $x_0$  bzgl.  $\tau(E, E^*)$  ist durch die Mengen der Form*

$$V = \{x \in E \mid |\langle f_i, x - x_0 \rangle| < \varepsilon, i \in J\},$$

wobei  $J$  endlich ist,  $f_i \in E^*$  und  $\varepsilon > 0$ , gegeben.

BEWEIS : Folgt sofort aus der allgemeinen Charakterisierung der Umgebungsbasis und dem Fakt, daß in  $\mathbb{R}$  die Mengen  $B_\varepsilon(y)$ ,  $\varepsilon > 0$ , eine Umgebungsbasis von  $y$  sind. ■

**7.6 Satz.** *Für eine Folge  $(x_n) \subseteq E$  gilt:*

(i)  $x_n \rightharpoonup x$  schwach bzgl.  $\tau(E, E^*)$  genau dann, wenn  $\langle f, x_n \rangle \rightarrow \langle f, x \rangle$  für alle  $f \in E^*$ ,

(ii)  $x_n \rightarrow x$  stark in  $E$ , dann gilt:  $x_n \rightharpoonup x$  schwach in  $E$ ,

(iii)  $x_n \rightharpoonup x$  schwach in  $E$ , dann ist die Folge  $(\|x_n\|)$  beschränkt und es gilt:

$$\|x\| \leq \liminf_{n \rightarrow \infty} \|x_n\| .$$

BEWEIS : (i) folgt aus Lemma 7.3, (ii) und (iii) werden wie im Hilbertraum bewiesen (vergleiche Abschnitt 2.7). ■

Auf dem Dualraum  $E^*$  eines normierten Raumes kann man neben der starken Topologie (gegeben durch die Norm in  $E^*$  (vergleiche Definition 2.1)) und der schwachen Topologie  $\tau(E^*, E^{**})$  noch die  $*$ -schwache Topologie  $\tau(E^*, E)$  definieren. Für  $x \in E$  definieren wir  $\varphi_x : E^* \rightarrow \mathbb{R}$  durch

$$\varphi_x(f) \equiv \langle f, x \rangle$$

und betrachten die Familie  $(\varphi_x)_{x \in E}$ .

**7.7 Definition.** Die  $*$ -schwache Topologie  $\tau(E^*, E)$  auf den Dualraum  $E^*$  eines Banachraumes  $E$  ist die grösste Topologie bzgl. derer alle  $(\varphi_x)_{x \in E}$  stetig sind. (Wir setzen  $X = E^*$ ,  $Y_i = \mathbb{R}$ ,  $I = E$  in 7.1).

Völlig analog zu Lemma 7.5 und Satz 7.6 beweist man.

**7.8 Lemma.** Sei  $f_0 \in E^*$ . Eine Umgebungsbasis von  $f_0$  bzgl.  $\tau(E^*, E)$  ist durch Mengen der Form

$$V = \{f \in E^* \mid |\langle f - f_0, x_i \rangle| < \varepsilon, i \in J\},$$

wobei  $J$  endlich ist,  $x_i \in E$  und  $\varepsilon > 0$ , gegeben.

**7.9 Satz.** Für eine Folge  $(f_n) \subseteq E^*$  gilt:

(i)  $f_n \xrightarrow{*} f$   $*$ -schwach bzgl.  $\tau(E^*, E)$  genau dann, wenn  $\langle f_n, x \rangle \rightarrow \langle f, x \rangle$  für alle  $x \in E$ ,

(ii)  $f_n \rightarrow f$  stark in  $E^*$ , dann gilt:  $f_n \xrightarrow{*} f$   $*$ -schwach in  $E^*$ ,

(iii)  $f_n \xrightarrow{*} f$   $*$ -schwach in  $E^*$ , dann ist die Folge  $(\|f_n\|)$  beschränkt und es gilt:

$$\|f\| \leq \liminf_{n \rightarrow \infty} \|f_n\| .$$

Sei  $E$  ein Banachraum,  $E^*$  sein Dualraum mit der Norm

$$\|f\|_{E^*} = \sup_{\|x\|_E \leq 1} |\langle f, x \rangle_{E^*, E}|$$

und  $E^{**}$  sein Bidualraum mit der Norm

$$\|g\|_{E^{**}} = \sup_{\|f\|_{E^*} \leq 1} |\langle g, f \rangle_{E^{**}, E^*}| .$$

Die **kanonische Isometrie**  $J : E \rightarrow E^{**}$  wird wie folgt definiert: sei  $x \in E$  fest, die Abbildung  $f \mapsto \langle f, x \rangle$  von  $E^*$  nach  $\mathbb{R}$  ist ein beschränktes, lineares Funktional auf  $E^*$ , d.h. ein Element von  $E^{**}$ , das mit  $Jx$  bezeichnet wird. Wir haben also

$$\langle Jx, f \rangle_{E^{**}, E^*} = \langle f, x \rangle_{E^*, E} \quad \forall x \in E, f \in E^*. \quad (7.10)$$

Es ist klar, daß  $J$  linear und eine Isometrie ist, d.h.  $\|Jx\|_{E^{**}} = \|x\|_E$  für alle  $x \in E$ . In der Tat gilt:

$$\|Jx\| = \sup_{\|f\| \leq 1} |\langle Jx, f \rangle| = \sup_{\|f\| \leq 1} |\langle f, x \rangle| = \|x\|$$

aufgrund von (4.22). Allerdings ist  $J$  nicht notwendig surjektiv. Man kann aber immer mit Hilfe von  $J$  den Raum  $E$  mit einem abgeschlossenen Unterraum von  $E^{**}$  identifizieren.

**Beispiel:** Wir haben gezeigt, daß  $(L^1(\Omega))^* = L^\infty$  und  $(L^\infty(\Omega))^* \supsetneq L^1(\Omega)$ , d.h.  $L^1(\Omega) \subsetneq (L^1(\Omega))^{**}$ .

**7.11 Definition.** Sei  $E$  ein Banachraum und sei  $J : E \rightarrow E^{**}$  die durch (7.10) definierte kanonische Isometrie von  $E$  nach  $E^{**}$ . Der Raum  $E$  heißt **reflexiv**, wenn  $J(E) = E^{**}$ , d.h.  $J$  ist surjektiv.

• Aufgrund der Definition von reflexiv und schwacher bzw. \*-schwacher Konvergenz erhält man sofort, daß in reflexiven Banachräumen schwache und \*-schwache Konvergenz übereinstimmen.

**7.12 Lemma.** Sei  $E$  ein Banachraum. Dann gelten:

- (i) Jeder abgeschlossene Unterraum von  $E$  ist reflexiv, falls  $E$  reflexiv ist.
- (ii) Sei  $A : E \rightarrow F$  ein Isomorphismus. Dann ist  $E$  reflexiv genau dann, wenn  $F$  reflexiv ist.
- (iii)  $E$  ist reflexiv genau dann, wenn  $E^*$  reflexiv ist.
- (iv)  $E^*$  ist separabel, dann ist auch  $E$  separabel.

BEWEIS : (i) Sei  $G \subseteq E$  ein abgeschlossener Unterraum. Für  $g \in G^{**}$  definieren wir  $\tilde{g} \in E^{**}$  durch

$$\langle \tilde{g}, f \rangle_{E^{**}, E^*} \equiv \langle g, f|_G \rangle_{G^{**}, G^*}, \quad f \in E^*. \quad (7.13)$$

Wir zeigen, dass  $J_E^{-1}\tilde{g} \in G$ . Falls  $J_E^{-1}\tilde{g} \notin G$  gelten würde, gäbe es nach Satz 4.17 ein  $f \in E^*$  und  $\varepsilon > 0$  mit

$$\langle f, u \rangle > \varepsilon + \langle f, J_E^{-1}\tilde{g} \rangle = \varepsilon + \langle g, f|_G \rangle, \quad \forall u \in G. \quad (7.14)$$

Da die rechte Seite der Ungleichung ein fester Wert ist und  $G$  ein Unterraum von  $E$  ist, folgern wir sofort  $\langle f, u \rangle = 0, u \in G$ , d.h.  $f|_G = 0$ . Dies wiederum in (7.14) eingesetzt ergibt einen Widerspruch.

Wir zeigen nun, dass  $J_G(J_E^{-1}\tilde{g}) = g$  gilt. Sei  $f \in G^*$  und  $\tilde{f} \in E^*$  eine Fortsetzung nach dem Satz von Hahn-Banach. Dann gilt (vgl. (7.13), (7.10)):

$$\begin{aligned} \langle g, f \rangle_{G^{**}, G^*} &= \langle g, \tilde{f}|_G \rangle_{G^{**}, G^*} = \langle \tilde{g}, \tilde{f} \rangle_{E^{**}, E^*} \\ &= \langle \tilde{f}, J_E^{-1}\tilde{g} \rangle_{E^*, E} \\ &= \langle f, J_E^{-1}\tilde{g} \rangle_{G^*, G} = \langle J_G J_E^{-1}\tilde{g}, f \rangle_{G^{**}, G^*}, \end{aligned}$$

d.h.  $g = J_G(J_E^{-1}\tilde{g})$  und somit ist  $J_G$  surjektiv.

(ii) Sei  $E$  reflexiv und  $g \in F^{**}$ . Dann ist durch

$$\langle h, u \rangle_{E^{**}, E^*} \equiv \langle g, u \circ A^{-1} \rangle_{F^{**}, F^*}, \quad u \in E^*$$

ein  $h \in E^{**}$  definiert. Für alle  $v \in F^*$  gilt:

$$\begin{aligned} \langle g, v \rangle_{F^{**}, F^*} &= \langle h, v \circ A \rangle_{E^{**}, E^*} \\ &= \langle v \circ A, J_E^{-1}h \rangle_{E^*, E} \\ &= \langle v, AJ_E^{-1}h \rangle_{F^*, F} \\ &= \langle J_F AJ_E^{-1}h, v \rangle_{F^{**}, F^*}, \end{aligned}$$

d.h.  $g = J_F AJ_E^{-1}h$  und somit ist  $J_F$  surjektiv.

(iii) Sei  $E$  reflexiv. Für alle  $\varphi \in E^{***}$  ist  $\varphi \circ J_E \in E^*$  und für alle  $g \in E^{**}$  gilt:

$$\begin{aligned} \langle \varphi, g \rangle_{E^{***}, E^{**}} &= \langle \varphi \circ J_E, J_E^{-1}g \rangle_{E^*, E} \\ &= \langle g, \varphi \circ J_E \rangle_{E^{**}, E^*} \\ &= \langle J_{E^*} \circ \varphi \circ J_E, g \rangle_{E^{***}, E^{**}}. \end{aligned}$$

Somit ist  $\varphi = J_{E^*}(\varphi \circ J_E)$ , d.h.  $J_{E^*}$  ist surjektiv. Sei umgekehrt  $E^*$  reflexiv. Nach dem gerade gezeigten ist  $E^{**}$  reflexiv. Da  $J_E(E)$  ein abgeschlossener Teilraum von  $E^{**}$  ist, ist nach (i)  $J_E(E)$  reflexiv und nach (ii) ist also  $E$  reflexiv.

(iv) Sei  $\{f_n \mid n \in \mathbb{N}\}$  dicht in  $E^*$ . Wir wählen  $u_n \in E$  mit  $\|u_n\| = 1$  und

$$|\langle f_n, u_n \rangle| \geq \frac{1}{2} \|f_n\|.$$

Wir setzen  $G = \overline{\text{span}\{u_n \mid n \in \mathbb{N}\}}$  und wollen zeigen, dass  $G = E$ . Sei  $u_0 \in E \setminus G$  mit  $\|u_0\| = 1$ . Nach dem Satz von Hahn-Banach (Satz 4.17) existiert ein  $f \in E^*$  mit

$$\langle f, u_0 \rangle > \varepsilon_0 + \langle f, u \rangle, \quad u \in G. \quad (7.15)$$

Da  $G$  ein abgeschlossener Unterraum ist, ist  $f|_G = 0$ . Es gilt:

$$\begin{aligned} \|f - f_n\| &\geq |\langle f - f_n, u_n \rangle| \\ &= |\langle f_n, u_n \rangle| \\ &\geq \frac{1}{2} \|f_n\| \geq \frac{1}{2} (\|f\| - \|f_n - f\|), \end{aligned}$$

und somit

$$\|f\| \leq 3 \|f_n - f\|.$$

Dies und (7.15) implizieren

$$\varepsilon_0 < |\langle f, u_0 \rangle| \leq \|f\| \leq 3 \|f_n - f\| \rightarrow 0 \quad (n \rightarrow \infty),$$

was ein Widerspruch ist. ■

Nun können wir den Hauptsatz über reflexive Räume beweisen.

**7.16 Satz.** *Sei  $E$  ein reflexiver Banachraum. Dann ist die abgeschlossene Einheitskugel in  $E$*

$$B_E = \{u \in E \mid \|u\| \leq 1\}$$

*schwach folgenkompakt.*

BEWEIS : Sei  $(u_n)$  eine Folge aus  $B_E$  und sei

$$G \equiv \overline{\text{span}\{u_n \mid n \in \mathbb{N}\}}.$$

Nach Lemma 7.12 ist auch  $G$  reflexiv und natürlich separabel. Somit ist auch  $J_E(G) = G^{**}$  separabel und also auch  $G^*$  (Lemma 7.12 (iv)). Wir können also Satz 2.6 auf  $G^*$  und die Folge  $(J_G u_n) \subseteq G^{**}$  anwenden. Es gibt also ein Element  $g \in G^{**}$ , so dass für eine Teilfolge

$$J_G u_n \xrightarrow{*} g \quad \text{in } G^{**}.$$

Für  $u = J_G^{-1}g$  und alle  $f \in G^*$  gilt:

$$\langle f, u_n \rangle_{G^*, G} = \langle J_G u_n, f \rangle_{G^{**}, G^*} \rightarrow \langle g, f \rangle_{G^{**}, G^*} = \langle f, u \rangle_{G^*, G}.$$

Da für beliebige  $F \in E^*$  die Restriktion  $f = F|_G$  ein Element von  $G^*$  ist, folgt somit auch

$$\langle F, u_n \rangle = \langle f, u_n \rangle \rightarrow \langle f, u \rangle = \langle F, u \rangle.$$

d.h.  $u_n \rightharpoonup u$  schwach in  $E$ . ■

**Beispiel:** (i) Jeder Hilbertraum ist reflexiv, da  $H^* \cong H$  und somit  $H^{**} \cong H^* \cong H$ .  
(ii) Die Lebesgue-Räume  $L^p, 1 < p < \infty$ , sind reflexiv, da  $(L^p)^* \cong L^{p'}$ , mit  $p' = \frac{p}{p-1}$  und somit

$$(L^p)^{**} = (L^{p'})^* = L^p.$$

(iii) Die Sobolevräume  $H^{m,p}, m \in \mathbb{N}, 1 < p < \infty$  sind reflexiv.

(iv)  $L^1$  und  $L^\infty$  sowie  $C^0(\overline{\Omega})$  sind nicht reflexiv.

Man kann zeigen, dass die Bedingung aus Satz 7.16 auch hinreichend für die Reflexivität eines Banachraumes ist. Es gilt:

**7.17 Satz.** *Sei  $E$  ein Banachraum. Dann ist  $E$  reflexiv genau dann, wenn die abgeschlossene Einheitskugel in  $E$  bzgl. der starken Topologie*

$$B_E = \{u \in E \mid \|u\| \leq 1\}$$

*kompakt bzgl. der schwachen Topologie  $\tau(E, E^*)$  ist.*

## 3.8 Konvexität und schwache Topologie

Aus der Hilbertraum-Theorie kennen wir den Satz von Banach-Saks, der besagt, dass man aus den Elementen einer *schwach* konvergenten Folge stark konvergente konvexe Linearkombinationen (nämlich arithmetische) mit dem gleichen Limes bilden kann. Unter anderem konnte man damit beweisen, dass der Begriff „abgeschlossen und konvex“ unabhängig davon ist, ob die starke oder die schwache Topologie verwendet wird. Diese Aussagen bilden ein sehr allgemeines Prinzip in der Funktionalanalysis, welches sogar in lokalkonvexen Räumen gilt. Wir beschränken uns hier aber auf die starke und schwache Topologie in Banachräumen. Im Folgenden sei also  $B$  ein Banachraum und  $B^*$  sein Dualraum.

**8.1 Satz.** *Eine konvexe Menge  $C$  eines Banachraumes  $B$  ist genau dann bzgl. der starken Topologie abgeschlossen, wenn sie bzgl. der schwachen Topologie abgeschlossen ist.*

BEWEIS : Die Richtung

„ $C$  ist schwach abgeschlossen“  $\Rightarrow$  „ $C$  ist stark abgeschlossen“

ist trivial, da die schwache Topologie gröber ist als die starke (vgl. Satz 7.6 (ii)). Sei also  $C$  stark abgeschlossen. Wir müssen zeigen, dass das Komplement von  $C$  bezüglich der schwachen Topologie offen ist:

Sei  $x_0 \notin C$ . Nach dem Satz von Hahn-Banach existiert eine  $x_0$  und  $C$  strikt trennende abgeschlossene Hyperebene, d.h. es existiert ein  $f \in B^*$  und ein  $\alpha \in \mathbb{R}$  mit

$$\langle f, x_0 \rangle < \alpha < \langle f, y \rangle, \quad y \in C.$$

Die Menge (vgl. Lemma 7.5)

$$V = \{x \in B \mid \langle f, x \rangle < \alpha\}$$

ist aber eine bezüglich der schwachen Topologie offene Menge, somit eine offene, mit  $C$  disjunkte Umgebung von  $x_0$ . Damit ist  $C$  abgeschlossen bezüglich der schwachen Topologie. ■

**8.2 Folgerung (Satz von Mazur).** *Es sei  $(u_m)$ ,  $u_m \in B$ ,  $m \in \mathbb{N}$  eine schwach konvergente Folge in einem Banachraum  $B$ . Dann gibt es eine Folge  $(v_j)$  von konvexen Linearkombinationen*

$$v_j = \sum_{k=1}^{N_j} c_k^j u_{m_{kj}}, \quad c_k^j \geq 0, \quad \sum_{k=1}^{N_j} c_k^j = 1, \quad m_{kj} \in \mathbb{N}, \quad N_j \in \mathbb{N}, \quad (8.3)$$

die stark gegen den schwachen Limes der  $u_m$  konvergiert.

BEWEIS : Wir bilden die Menge  $C$  aller endlichen konvexen Linearkombinationen von Elementen  $u_m$ .  $C$  ist offensichtlich konvex, ebenso der Abschluss  $\overline{C}$  bezüglich der starken Topologie. Ist nun  $u_m \rightharpoonup u$  schwach, so muss wegen Satz 8.1  $u$  ebenfalls in  $\overline{C}$  liegen. Nach Definition von  $\overline{C}$  gibt es dann aber eine Folge von Elementen aus  $C$ , die stark gegen  $u$  geht. Die Elemente aus  $\overline{C}$  haben aber gerade die Gestalt (8.3). ■

**8.4 Lemma.** *Jedes konvexe, koerzive unterhalbstetige Funktional  $f$  auf einer abgeschlossenen, konvexen Menge  $C$  eines reflexiven Banachraumes besitzt ein Minimum. Die Begriffe „abgeschlossen“ und „unterhalbstetig“ sind bezüglich der starken Topologie gemeint.*

BEWEIS : Sei  $(u_m)$ ,  $u_m \in C$  Minimalfolge von  $f$ , d.h.

$$f(u_m) \rightarrow \inf_C f \quad (m \rightarrow \infty).$$

Aufgrund der Koerzivität von  $f$  muss die Folge  $(u_m)$  beschränkt sein. Dann gibt es aufgrund von Satz 7.16 eine schwach konvergente Teilfolge  $(u_m)_{m \in \Lambda}$ ,  $u_m \rightharpoonup u$ ,  $m \in \Lambda$ , und wegen Satz 8.1 gilt  $u \in C$ . Ferner gibt es eine Folge  $(v_j)$  von konvexen Linearkombinationen der  $u_m$ , die stark gegen  $u$  konvergieren. Wegen der Konvexität folgt

$$f(v_j) \leq \sum_{k=1}^{N_j} c_k^j f(u_{m_{k_j}}) \leq \sum_{k=1}^{N_j} c_k^j (\inf_C f + \varepsilon) = \inf_C f + \varepsilon.$$

(Die  $m_{k_j}$  können so gewählt werden, dass  $f(u_{m_{k_j}}) = \inf_C f + e$ ,  $|e| < \varepsilon$ , indem man  $m_{1j}$  genügend groß wählt.) Da  $f$  unterhalbstetig ist und  $v_j \rightarrow u$  stark, gilt

$$f(u) \leq \liminf_{j \rightarrow \infty} f(v_j) \leq \inf_C f + \varepsilon.$$

Grenzübergang  $\varepsilon \rightarrow 0$  ergibt  $f(u) = \inf_C f$ , d.h. das Minimum wird angenommen. ■