Vorlesung

FUNKTIONALANALYSIS I

gelesen an der Universität Bonn
im Wintersemester 1998/99 von

Prof. Dr. Jens Frehse

ausgearbeitet und korrigiert von

Sonja Goj, Moritz Kassmann und Marco Röder

Stand: 25. April 2002

Inhaltsverzeichnis

1 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.

Ein typischer Satz der Funktionalanalysis ist der *Banachsche Fixpunktsatz* („jede kontrahierende Abbildung eines Banachraumes in sich besitzt einen Fixpunkt").

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 für die Vertiefungsgebiete an unserem Institut 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 von Null verschiedenen Kern haben.

Beispiel: Sei $V = \{\,(c_1, c_2, c_3, \dots) \,|\, c_j \in \mathbb{R}\,\}$ 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 $A\colon V \to V$ sei definiert durch

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

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

b) Ist der Nullraum $N(A)$ einer linearen Abbildung Null, 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

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

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

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

Dies ist nicht möglich, $\sin \neq$ 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 exotisch und für die Analysis nicht zu gebrauchen.)

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

Man hat z.B. die Topologie der punktweisen Konvergenz

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

Ferner gibt es die l^2-Konvergenz

$$c^j \to c \ :\Leftrightarrow \ \|c^j - c\|_{l^2} \to 0 \qquad (j \to \infty)\,.$$

Hierbei ist $\|\omega\| = (\sum\limits_{k=1}^{\infty} |\omega_k|^2)^{1/2}$.

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. $p^j \to p \Longleftrightarrow \|p^j - p\|_\infty \to 0$ $(j \to \infty)$. Hierbei ist $\|q\|_\infty = \max\{ |q(x)| \mid x \in [-2, 2] \}$.

Wir definieren die lineare Abbildung $A\colon V \to V$, 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 $p_n \to 0$ $(n \to \infty)$, aber

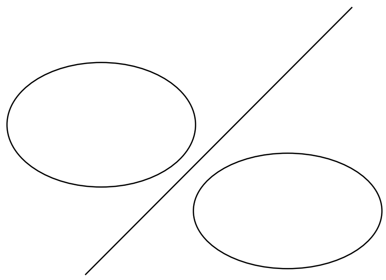
$$\|Ap_n(x)\| = \frac{3^n \cdot x^n}{(2,5)^n} \to \infty\,.$$

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

f) Im $\mathbb{R}^n$ gilt der sogenannte Browersche Fixpunktsatz: „Jede stetige Abbildung der Einheitskugel in sich besitzt einen Fixpunkt.“

In jedem unendlich–dimensionalen Banachraum (Definition später) gibt es eine stetige Abbildung der Einheitskugel in sich *ohne* Fixpunkt. Man sagt: Unendlich–dimensionale Banachräume besitzen nicht die Fixpunkteigenschaft.

g) Im $\mathbb{R}^n$ gilt der anschaulich einleuchtende Satz, daß disjunkte konvexe Mengen durch eine Hyperebene getrennt werden können.



Dies ist z.B. wichtig für die Optimierungstheorie.

Im unendlich–dimensionalen Fall gilt dies nur unter Zusatzbedingungen an die konvexen Mengen (z.B. eine von ihnen muß nicht–leeres Inneres haben).

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^ω, C^α, $L^2(\Omega)$, $L^p(\Omega)$. Man benötigt die vielen Funktionenräume u.a., um Sätze der Funktionalanalysis einer konkreten Situation anzupassen. Will man z.B. erreichen, daß der sogenannte Nemytski-Operator $T_f\colon L^p[a,b] \to L^q[a,b]$

$$(T_f u)(x) := f(u(x))\,, \qquad f \in C(\mathbb{R},\mathbb{R}) \text{ vorgegeben,}$$

stetig bezüglich der in L^p und L^q gegebenen Normen ist, so folgt notwendig

$$|f(\xi)| \leq K|\xi|^{p/q} + K\,.$$

Ist f also z. B. eine vorgegebene Funktion mit polynomialem Wachstum, so muß man den Grundraum L^p mit einem *genügend großem p* wählen. *Man paßt also den Grundraum der Nichtlinearität an.*

In Hinblick auf die Abstimmung mit anderen Vorlesungen beginnt die Vorlesung mit einem Schnellkurs über den Hilbertraum. Danach beschäftigen wir uns mit linearer Funktionalanalysis in Banachräumen und lokalkonvexen Räumen. Es werden grundlegende Definitionen und Prinzipien erläutert. Nach einem anschließenden Exkurs über Spektraltheorie hoffe ich, noch die Theorie monotoner Operatoren zu behandeln.

Teil I

Schnellkurs über Hilberträume

2 Der Hilbertraum

Definition 2.1 *Ein* Prä–Hilbertraum V *(„inner product space“) ist ein Vektorraum über* $\mathbb{R}$ *bzw.* $\mathbb{C}$*, in dem ein* Skalarprodukt $(.,.)$ *gegeben ist, d.h. eine Abbildung von* $H \times H$ *nach* $\mathbb{R}$ *oder* $\mathbb{C}$ *mit den Eigenschaften*

(i) Bilinearität bzw. Sesquilinearität

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

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

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

$$(u, v) = \overline{(v, u)} \qquad \text{für alle } u, v \in V \qquad (\overline{\ \ } = \text{„Konjugiert-Komplex-Zeichen“}).$$

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

$$(u, u) \geq 0 \qquad \text{für alle } u \in V,$$

und $(u, u) = 0$ *genau für* $u = 0$.

Definition 2.2 $\|u\| := \sqrt{(u, u)}$ *ist die durch das Skalarprodukt in* V *definierte Norm.*

Definition 2.3 *Eine Folge* $(u_j)_{j\in\mathbb{N}}$, $u_j \in V$ *konvergiert bezüglich* $\|.\|$ *(man sagt auch „in* V*“) gegen* $u \in V$ *genau dann, wenn*

$$\|u_j - u\| \to 0 \qquad\qquad (j \to \infty)\,.$$

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

Definition 2.4 *Eine Folge* $(u_j)_{j\in\mathbb{N}}$, $u_j \in V$ *heißt Cauchyfolge (bezüglich* $\|.\|$*) 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 $\|u_j - u_k\| \to 0$ für $j, k \to \infty$.)

Definition 2.5 *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 $(.,.)$ sind Hilberträume.

b) $\mathbb{R}^n$ oder $\mathbb{C}^n$ mit dem Skalarprodukt

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

mit einer hermiteschen positiv definiten Matrix A.

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

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

ist Prä–Hilbertscher, aber kein Hilbertscher Raum.

d) l^2 ist ein Hilbertraum. Hierbei ist

$$l^2 = \{\, (c_1, c_2, \ldots,) \mid c_j \in \mathbb{R} \text{ oder } \mathbb{C}\}$$
$$(b, c)_{l^2} = (b, c) = \sum_{j=1}^{\infty} b_j \overline{c}_j$$

e) $L^2[a, b]$, allgemeiner $L^2(\Omega)$, ist ein Hilbertraum. Hierbei ist

$$L^2[a, b] = \left\{\, [u] \,\middle|\, u \text{ ist meßbar und } \int_a^b |u|^2\, dx < \infty \,\right\}$$

und $[u]$ die Klasse aller Funktionen, die sich von u auf einer Menge vom Maß Null unterscheiden. Das Skalarprodukt ist erklärt durch

$$\big([u], [v]\big) = (u, v) = \int_a^b u\overline{v}\, dx\,.$$

Es ist üblich, nicht zwischen $[u]$ und u zu unterscheiden, obwohl dies unpräzise ist.

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

Definition 2.6 *von $\tilde{H}^{1,2}(\Omega)$. Sei $\tilde{H}^{1,2}$ die Menge aller Funktionen aus $C^1(\Omega)$ mit $\int\limits_\Omega \left(|u|^2 + |\nabla u|^2\right) dx < \infty$.*
Bezüglich des (sogenannten $H^{1,2}$–Skalarproduktes)

$$(u,v)_1 = \int\limits_\Omega u\overline{v}\,dx + \int\limits_\Omega \nabla u \cdot \nabla\overline{v}\,dx$$

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

Wir erläutern im Folgenden *drei* Definitionen des Sobolevraumes $H^{1,2}$.

Definition 2.7 (Erste Definition von $H^{1,2}$) *$H^{1,2}(\Omega)$ ist der Restklassenraum der Cauchyfolgen $(u_j \in \tilde{H}^{1,2})_{j\in\mathbb{N}}$ 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 $u_j \to u$ in $L^2(\Omega)$. Die Größe $\mathrm{L}^2\text{-}\lim\limits_{\mathrm{j}\to\infty} \nabla u_j$ kann man sich als verallgemeinerte Ableitung von u vorstellen und schreibt $\nabla u = \lim\limits_{j\to\infty} \nabla u_j$.

Durch die Restklassenbildung sieht man Limites von Cauchyfolgen $(u_j), (u_j')$ als gleich an, wenn $u_j - u_j' \to 0$ in $\tilde{H}^{1,2}$. Diese Art der Definition des Sobolevraums $H^{1,2}$ nennt man die „Definition durch Vervollständigung“.

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

Definition 2.8 *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*

$$(u, D_i\varphi)_{L^2} = -(u_i, \varphi)_{L^2} \qquad \text{für alle } \varphi \in C_0^\infty(\Omega). \tag{2.1}$$

Hierbei ist $C_0^\infty(\Omega)$ der sogenannte *Raum der Testfunktionen*, d.h. $C_0^\infty(\Omega) = \{\varphi \,|\, \varphi$ ist unendlich oft differenzierbar und φ hat kompakten Träger in $\Omega\}$.

Der Träger von φ ist der Abschluß der Menge aller Punkte x mit $\varphi(x) \neq 0$. Es gilt $\varphi = 0$ in $U(\partial\Omega)$. Die Definition (2.1) ist motiviert durch die partielle Integration, falls $u \in C^1(\Omega)$, denn dann gilt

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

Lemma 2.1 (Eindeutigkeit der verallgemeinerten Ableitung) *Es seien u_i, $\tilde{u}_i \in L^2(\Omega)$ verallgemeinerte Ableitungen von $u \in L^2$. Dann gilt $u_i = \tilde{u}_i$ f. ü.*

Definition 2.9 (Zweite Definition von $H^{1,2}$ mit Hilfe verallgemeinerter Ableitungen)

$$H^{1,2}(\Omega) = \left\{ u \in L^2(\Omega) \,\middle|\, u \textit{ besitzt erste verallgemeinerte Ableitungen in } L^2(\Omega) \right\}.$$

Das Skalarprodukt wird dann mit Hilfe der verallgemeinerten Ableitungen definiert durch

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

Satz 2.1 $H^{1,2}(\Omega)$ *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 $u \in L^2(\Omega)$. Wir zeigen, daß u verallgemeinerte Ableitungen in $L^2(\Omega)$ hat. Da $(D_i u_j)$ ebenfalls Cauchy-Folge in $L^2(\Omega)$ ist, gilt $D_i u_j \to \tilde{u}_i \in L^2(\Omega)$ in $L^2(\Omega)$ und $-(D_i \varphi, u_j) = (\varphi, D_i u_j) \to (\varphi, \tilde{u}_i)$ für alle $\varphi \in C(\Omega)$. Daher gilt $\tilde{u}_i = D_i u$.
□

Definition 2.10 (Dritte Definition des Raumes $H^{1,2}(\Omega)$ nach Tonnelli)
Es sei M die Menge aller quadratintegrablen Funktionen f mit der Eigenschaft, daß die durch

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

definierten Restriktionen φ_i bis auf eine Ausnahmemenge von Punkten $(x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ mit dem (n-1)-dimensionalen Maß Null absolutstetig *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 $D_i f \in L^2(\Omega)$, $i = 1, \ldots, n$, 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.

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

$$H_0^{1,2}(\Omega) = \overline{C_0^\infty(\Omega)}^{H^{1,2}} = \left\{ u \in H^{1,2} \,\middle|\, \exists u_j \in C_0^\infty \text{ mit } u_j \to u \text{ in } H^{1,2} \right\}$$

Beispiele von $H^{1,2}$-Funktionen:

Lemma 2.2 *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 $|x|^{-s} \notin H^{1,2}$.

Für $n = 3, 4$ ist also z. B. $\frac{1}{|x|} \notin H^{1,2}$, für $n \geq 5$ gilt die Inklusion jedoch.

Eine unbeschränkte $H^{1,2}$-Funktion für $n = 2$ ist die Funktion

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

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

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

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

und mit Hilfe von Polarkoordinaten erhält man $x_i\, |x|^{-s-2} \sim |x|^{-s-1} \in L^2(\Omega)$.

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

$$\big(D_i(|x|^{-s}\tau), \varphi\big) = \big(|x|^{-s}\tau, D_i\varphi\big) \;\to\; -\big(|x|^{-s}, D_i\varphi\big) \quad (R \to 0)\,.$$

Andererseits ist

$$\big(D_i(|x|^{-s}\tau), \varphi\big) = (-s\, x_i\, |x|^{-s-2}\tau, \varphi) + \big(|x|^{-s}, D_i\varphi\big)\,.$$

Es gilt $|x|^{-s}|D_i\tau| \leq KR^{-s}R^{-1}$ und (für $n \geq 2$) somit $\int_{B_{2R}} |x|^{-s}|D_i\tau| \leq KR^{n/2-1+\delta} \to 0$ $(R \to 0)$, wobei $\delta = \frac{n}{2} - 1 - s$. Daraus erhält man die Identität

$$\big(D_i(|x|^{-s}\tau), \varphi\big) = -\big(|x|^{-s}, D_i\varphi\big)\,.$$

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

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

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

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

$$\|a+b\|^2 = \|a\|^2 + (a,b) + (b,a) + \|b\|^2 = \|a\|^2 + 2\,\mathcal{R}e(a,b) + \|b\|^2$$

($\mathcal{R}e =$ Realteil) gilt, da $(b,a) = \overline{(a,b)}$.

Lemma 2.3 (Cauchy–Schwarzsche Ungleichung) *Sei V ein Prä-Hilbertraum. Dann gilt für alle $u, v \in V$*

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

Beweis für den reellen Fall: Es gilt

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

und somit

$$2\,\mathcal{R}e(u,v) \leq |\alpha|^2\,\|u\|^2 + \frac{1}{|\alpha|^2}\|v\|^2\,. \tag{2.2}$$

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

$$\mathcal{R}e(u,v) \leq \|u\|\,\|v\|\,.$$

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

Beweis für den reellen und komplexen Fall: Es gilt mit $\alpha \in \mathbb{R}$, $\rho \in \mathbb{C}$

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

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

$$\mathcal{R}e(u,\rho v) \leq \|u\|\,\|v\| \text{ und } \mathcal{R}e\big(\overline{\rho}(u,v)\big) \leq \|u\|\,\|v\|$$

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

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

□

Folgerung aus der Cauchy–Schwarzschen Ungleichung:

Lemma 2.4 *Das Skalarprodukt ist bezüglich der Normkonvergenz stetig.*

Beweis: Sei $u_j \to u$, d.h. $\|u_j - u\| \to 0$ $(j \to \infty)$. Dann gilt

$$|(u,w) - (u_j,w)| \leq \|u - u_j\| \, \|w\|$$

und daher $(u_j,w) \to (u,w)$ $(j \to \infty)$. Entsprechend gilt $(w,u_j) \to (w,u)$ $(j \to \infty)$.
□

Lemma 2.5 *Sei V ein Prä–Hilbertscher Raum mit Skalarprodukt $(.,.)$. Die durch $\|u\| = \sqrt{(u,u)}$ definierte Größe erfüllt die Dreiecksungleichung*

$$\|u+v\| \leq \|u\| + \|v\| \,.$$

Beweis:

$$\begin{aligned} \|u+v\|^2 &= \|u\|^2 + 2\,\mathcal{R}e(u,v) + \|v\|^2 \leq \\ &\leq \|u\|^2 + 2\,|(u,v)| + \|v\|^2 \leq \|u\|^2 + 2\,\|u\| \cdot \|v\| + \|v\|^2 = (\|u\| + \|v\|)^2 \,. \end{aligned}$$

□

3 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

$$H = V \oplus V^{\perp} . \tag{3.1}$$

Hierbei ist das *„Orthogonalkomplement von V"* definiert durch

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

Die Gleichung (3.1) 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 (3.1) ist der Hilbertraum $L^2(I)$, $I = (a,b)$ und der Raum $V = C_0^\infty(I)$, wobei

$$C_0^\infty(I) = \{ \varphi \in C^\infty(I) \mid \operatorname{Trg} \varphi \subset I \}$$

mit $\operatorname{Trg} \varphi := \overline{\{ x \in I \mid \varphi(x) \neq 0 \}}$ ($\overline{M}$ = Abschließung von M).

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

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

der Widerspruch $w = 0$. (Hierbei ist $(.,.)$ das L^2-Skalarprodukt.)

Zum Beweis dieser Folgerung beachten wir, daß $C_0^\infty(I)$ in $L^2(I)$ *dicht* ist, d.h. zu $w \in L^2(I)$ existiert eine Folge φ_j mit $\varphi_j \to w$ in $L^2(I)$. Da das Skalarprodukt stetig bezüglich der L^2–Konvergenz ist, gilt für das obige w

$$0 = (w, \varphi_j) \ \to \ (w,w) \,,$$

d.h. $w = 0$, q.e.d..

Für den Beweis des Satzes, daß $C_0^\infty(I)$ *dicht* in $L^2(I)$ ist, wird benutzt, daß

$$\chi(I_\varepsilon) w \to w \text{ in } L^2 \quad (\varepsilon \to 0) \,,$$

wobei $\chi(I_\varepsilon)$ die charakteristische Funktion des Intervalles $I_\varepsilon = [a+\varepsilon, b-\varepsilon]$, $I = [a,b]$ ist. Ferner wird benutzt, daß die Faltung $w_h * \chi(I_\varepsilon)w$ gegen $\chi(I_\varepsilon)w$ in $L^2(I)$ konvergiert.

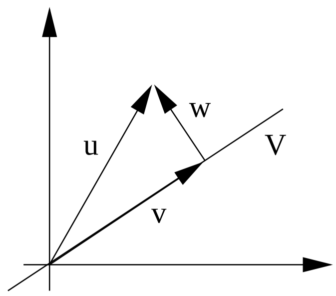
Die Hilbertraum–Theorie startet mit dem Satz, daß die Zerlegung (3.1) für *abgeschlossene* lineare Unterräume $V \subset H$ funktioniert.

Satz 3.1 (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: $H = V \oplus V^{\perp}$.

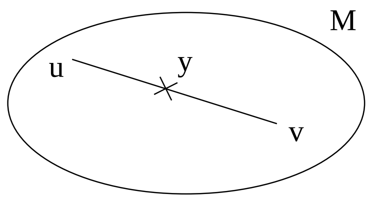
Der Beweis beruht auf der Tatsache, daß das Element v in Satz 3.2 dem Element u „am nächsten" ist, d.h. $\|u - v\| = \inf\big\{\|u - y\| \,\big|\, y \in V\big\}$.



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.

Definition 3.1 *Eine Menge $M \subset H$ heißt* konvex $\Leftrightarrow$ $\forall u, v \in M$ *und* $\forall \alpha \in [0,1]$ *gilt*

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



Satz 3.2 *Es sei M eine nichtleere, abgeschlossene, konvexe Menge $\subset H$ und $u \in H$. Dann gibt es ein eindeutiges Element $v \in M$ mit*

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

Beweis: Wenn $u \in M$, gibt es nichts zu beweisen ($\inf = 0$). Sei also $u \notin M$. Sei (w_i), $w_i \in M$ eine Minimalfolge, d.h. $\|u - w_i\| \to d$ ($i \to \infty$). 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

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

für $j,\, k \to \infty$. 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

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

und somit $\|\frac{1}{2}(w_k - w_j)\| = o(1)$, d.h. (w_i) *ist eine Cauchyfolge* und hat einen Limes $v \in M$. (Beachte, daß M abgeschlossen ist.) Da $w_i \to v$, 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**

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

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

Beweis von Satz 3.1: Aus Satz 3.2 mit $M = V$ folgt die Existenz eines eindeutigen Elementes $v \in V$ mit $\|u - v\|^2 = \inf\left\{ \|u - w\|^2 \,\middle|\, w \in V \right\} =: d^2$. 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 $t \to 0$ ergibt $0 \leq 2\,(u - v, \varphi)$ für alle $\varphi \in V$. Ersetzt man in dieser Ungleichung φ 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

$$0 \le 2\,\mathcal{R}e(u-v, t\varphi) + |t|^2\,\|\varphi\|^2 \tag{3.2}$$

und daraus wie eben

$$0 = \mathcal{R}e(u-v, \varphi)$$

Um auch $\mathcal{I}m(u-v,\varphi) = 0$ zu erhalten, wählt man in (3.2) $t = \mathrm{i}\,\tau$ ($\mathrm{i} = \sqrt{-1}$, $\tau \in \mathbb{R}$, $\tau > 0$) und erhält

$$0 \le 2\,\mathcal{R}e(-\,\mathrm{i}\,\tau(u-v,\varphi)) - \tau^2\,\|\varphi\|^2$$

Kürzen mit τ, und Grenzübergang $\tau \to 0$ ergibt

$$0 \le \mathcal{R}e(-\,\mathrm{i}(u-v,\varphi)) = \mathcal{I}m(u-v,\varphi)$$

Durch Übergang zu $-\varphi$ folgt wieder insgesamt

$$\mathcal{I}m(u-v,\varphi) = 0$$

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

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

$\square$

4 Beschränkte lineare Funktionale und der Riesz-sche Darstellungssatz

Definition 4.1 *Ein* beschränktes lineares Funktional *F auf einem Hilbertraum H ist eine lineare Abbildung $H \to \mathbb{R}$ bzw. $\mathbb{C}$ mit der Eigenschaft, daß eine Konstante K existiert, so daß*

$$|F(u)| \leq K\|u\| .$$

Die Menge der beschränkten linearen Funktionale wird auch mit H^* bezeichnet.

Satz 4.1 *Ein lineares Funktional auf einem Hilbertraum ist genau dann beschränkt, wenn es stetig ist.*

Beweis:

„$\Rightarrow$“ *Beschränktheit impliziert Stetigkeit.*

Dies ist trivial: Gilt $u_m \to u$, so folgt

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

„$\Leftarrow$ “ *Stetigkeit impliziert Beschränktheit.*

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

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

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

□

Definition 4.2 *Die* Norm *eines beschränkten linearen Funktionals $F\colon H \to \mathbb{R}$ ist definiert durch*

$$\|F\| = \inf\left\{K \,\middle|\, |F(u)| \leq K\|u\| \text{ für alle } u \in H\right\}.$$

Äquivalent ist die Definition

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

(Nachweis: Übungsaufgabe)

Lemma 4.1 *Die oben definierte Abbildung* $\|.\|\colon H \to \mathbb{R}$ *ist eine Norm auf* H^*. *(Übungsaufgabe)*

Satz 4.2 (Rieszscher Darstellungssatz) *Sei* H *ein Hilbertscher Raum. 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 $N = \{\, x \in H \,|\, F(x) = 0 \,\}$ der Nullraum von F. N ist abgeschlossener linearer Teilraum von H. Ist nämlich $u_m \in N$, $u_m \to u$, so folgt $0 = F(u_m) \to F(u)$, 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 $F\big(u - \frac{F(u)}{F(w)} w\big) = 0$, also $u - \frac{F(u)}{F(w)} w \in N$. Da $w \in N^{\perp}$, folgt

$$\big(u - \tfrac{F(u)}{F(w)} w, w\big) = 0$$

und

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

somit

$$F(u) = \big(u, \overline{\tfrac{F(w)}{\|w\|^2}}\, w\big)\,.$$

Damit ist das darstellende Element konstruiert. □

Eine wichtige Anwendung des Rieszschen Darstellungssatzes ist der Nachweis eines Existenzsatzes für elliptische Randwertprobleme der Gestalt

$$-\sum_{i,k=1}^{n} D_i\big(a_{ik}(x) D_k u\big) + c(x) u = f \text{ in } \Omega \tag{4.1}$$

mit Randbedingungen für u auf $\partial\Omega$. Hierbei ist Ω eine offene Punktmenge des $\mathbb{R}^n$, $f \in L^2, a_{ik} \in L^\infty(\Omega), c \in L^\infty(\Omega)$, und es gilt die Elliptizitätsbedingung

$$\sum_{i,k=1}^{n} a_{ik}\,\xi_k\,\xi_i \geq \lambda_0|\xi|^2 \qquad \text{für alle } \xi \in \mathbb{R}^n\,. \tag{4.2}$$

mit einer positiven Konstante λ_0 („Elliptizitätskonstante“).

Ferner gelte

$$a_{ik} = a_{ki} \tag{4.3}$$

und

$$c(x) \geq \lambda_1 > 0 \tag{4.4}$$

(Für Nullrandbedingungen könnte man $\lambda_1 = 0$ zulassen.)

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

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

und beschränken uns auf den reellen Fall.

Die Bedingungen (4.2) und (4.3) stellen sicher, daß $(.,.)$ ein Skalarprodukt ist. Diese Vorgehensweise ist typisch für die Anwendungen der Funktionalanalysis - *man paßt das Skalarprodukt dem Problem an.*

Aufgrund von (4.2) und (4.4) erkennt man, daß die Skalarprodukte $(.,.)_{H^{1,2}}$ und $(.,.)_A$ äquivalent sind, d.h. es gibt Konstanten K_1, K_2 mit

$$\|u\|_{H^{1,2}} \leq K_1\,\|u\|_A \leq K_2\|u\|_{H^{1,2}}\,. \tag{4.5}$$

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

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

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

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

Ausgeschrieben bedeutet dies

$$\sum_{i,k=1}^{n} \int_\Omega a_{ik}\,D_k u\,D_i v\,dx + \int_\Omega c\,u\,v\,dx = \int_\Omega f\,v\,dx \tag{4.6}$$

d.h. u ist schwache Lösung der Differentialgleichung (4.1). Außerdem folgt aus (4.6) noch die natürliche Randbedingung

$$\sum_{i,k=1}^{n} a_{ik}\, D_k u\, \nu_i = 0 \quad \text{auf } \partial\Omega,\ \nu = \text{ Normalenvektor} \tag{4.7}$$

Statt $H^{1,2}$ läßt sich auch $H_0^{1,2}$ als Grundraum nehmen, (4.7) entfällt dann.

5 Beschränkte lineare Abbildungen

Definition 5.1 *Seien H_1, H_2 Hilbertsche Räume. Eine lineare Abbildung $A\colon H_1 \to H_2$ heißt* beschränkt *genau dann, wenn eine Zahl $K \in \mathbb{R}$ existiert mit*

$$\|Au\| \leq K\|u\| \qquad \text{für alle } u \in H_1\,. \tag{5.1}$$

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

Satz 5.1 *Eine Abbildung A ist genau dann beschränkt, wenn sie stetig ist.*

Beweis:

„$\Rightarrow$" Sei A beschränkt und $u_m \to u$ in H_1, d.h. $\|u_m - u\| \to 0$ $(m \to \infty)$. Dann folgt aus (5.1)

$$\|Au_m - Au\| \to 0$$

d.h. $Au_m \to Au$.

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

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

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

□

Anmerkung: Satz und Beweis gelten in *beliebigen normierten Räumen.*

Beispiele beschränkter linearer Abbildungen in Hilbertschen Räumen:

a) $H_1 = H_2 = L^2(\Omega)$, Ω offene Teilmenge $\subset \mathbb{R}^n$. Sei $A\colon H_1 \to H_2$ gegeben durch

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

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

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

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

(D_1 = verallgemeinerte partielle Ableitung). A ist beschränkt, da

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

(Auf L^2 wäre D_1 nicht einmal definiert, als Abbildung von $L^2 \cap H^{1,2} \to L^2$ ist D_1 unbeschränkt.)

c) Sei $H_1 = H_2 = L^2(\Omega)$ und $K\colon \Omega\times\Omega \to \mathbb{R}$, K meßbar und $\int\limits_\Omega \int\limits_\Omega K^2(x,y)\,dx\,dy =: C < \infty$. Dann ist die durch

$$(Au)(x) := \int\limits_\Omega K(x,y)\,u(y)\,dy \tag{5.2}$$

definierte Abbildung beschränkt.

Beweis: $\|Au\|_{L^2}^2 = \int\limits_\Omega \big(\int\limits_\Omega K(x,y)\,u(y)\,dy\big)^2\,dx \leq \int\limits_\Omega \big(\int\limits_\Omega K^2(x,y)\,dy \int\limits_\Omega u^2(y)\,dy\big)\,dx$
$= \int\limits_\Omega \int\limits_\Omega K^2(x,y)\,dy\,dx \int\limits_\Omega u^2(y)\,dy = C\|u\|^2.$ □

Die Abbildung in (5.2) heißt Integraloperator mit Kern K. Es gibt Kerne, die so singulär sind, daß $\int\int K^2(x,y)\,dx\,dy = \infty$, sie aber dennoch einen beschränkten Integraloperator erzeugen.

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

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

Man kann beweisen:

$$\|\tilde{R}u\|_{L^2(\partial\Omega)} \leq K\|u\|_{H^{1,2}(\Omega)}\,. \tag{5.3}$$

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

$$R\colon H^{1,2}(\Omega) \to L^2(\partial\Omega)$$

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 Ω ein Würfel ist, reicht es, zu verlangen, daß $\partial\Omega$ sich in endlich viele C^1–Flächen zerlegen läßt. (Genauer: $\partial\Omega$ ist Lipschitz–Mannigfaltigkeit.) Die Konstruktion von $\tilde{R}$ geschieht folgendermaßen: Zu $u \in H^{1,2}(\Omega)$ existieren Funktionen $u_j \in C^1(\overline{\Omega})$ mit $u_j \to u$ in H^1 $(j \to \infty)$. (Beweisbedürftig, da $u_j \in C^1(\overline{\Omega})$ und nicht bloß $\in C^1(\Omega)$ verlangt wird - die Glattheit von $\partial\Omega$ wird verwendet.) Wegen (5.3) gilt $\|\tilde{R}u_j - \tilde{R}u_k\|_{L^2(\partial\Omega)} \to 0$ $(j,k \to \infty)$, d.h. $\tilde{R}$ ist Cauchyfolge mit Limes $v \in L^2(\partial\Omega)$. *Man definiert nun* $Ru = v$.

Es folgen nun einige Lösbarkeitssätze für lineare Gleichungen $Au = f$ mit beschränkten linearen Abbildungen $A\colon H \to H$, H ein Hilbertscher Raum.

Satz 5.2 *Sei A eine beschränkte lineare Abbildung eines Hilbertschen Raumes H in sich. Es gelte die Koerzitivitätsbedingung*

$$\mathcal{R}e(Au,u) \geq c_0\|u\|^2 \qquad \text{für alle } u \in H\,, \tag{5.4}$$

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 $A(H)$ von H *abgeschlossen* ist. Aus (5.4) folgt $\|Au\| \geq c_0\|u\|$. Ist nun $Au_j \to z$, so folgt $c_0\|u_j - u_k\| \leq \|Au_j - Au_k\| \to 0$ $(j,k \to \infty)$, d.h. u_j ist eine Cauchyfolge mit Limes u. Wäre nun $A(H)$ ein echter Teilraum von H, so gäbe es ein Element $w \neq 0$, $w \perp A(H)$. Aus (5.4) folgt der Widerspruch

$$0 = \mathcal{R}e(Aw,w) \geq c_0\|w\|^2 > 0\,,$$

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: $H = L^2(\Omega)$, $Au := u - \int\limits_\Omega K(x,y)\,u(y)\,dy$ unter der Voraussetzung

$$\int\limits_\Omega \int\limits_\Omega |K(x,y)|^2\,dx\,dy \leq q < 1,$$

denn mit Hilfe der Cauchy–Schwarzschen Ungleichung folgt

$$\big|\big(u, \int\limits_\Omega K(.,y)u(y)\,dy\big)\big| \leq \|u\|_{L^2}^2 \cdot \int\limits_\Omega \int\limits_\Omega |K|^2\,dx\,dy.$$

Die Konstante c_0 in (5.4) ist dann $1 - q$.

Eine der wichtigsten Anwendungen von Satz 5.2 liegt jedoch in der Theorie der elliptischen partiellen Differentialgleichungen. Hierzu benötigen wir den Begriff der beschränkten Bilinearform (bzw. Sesqilinearform) und das Lemma von Lax–Milgram.

Definition 5.2 *Eine beschränkte* Bilinearform (Sesquilinearform) *ist eine Abbildung* $[.,.]$ *von* $H \times H$ *nach* $\mathbb{R}$ *bzw.* $\mathbb{C}$ *mit den Eigenschaften*

a) $[u, v]$ *ist linear in* u *und* v *(bzw. antilinear in* v*).*

b) Es existiert eine Konstante K *mit*

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

Definition 5.3 *Eine Bilinearform (Sesquilinearform) heißt* koerzitiv*, wenn eine Konstante* $c_0 > 0$ *existiert mit* $|[u, u]| \geq c_0\|u\|^2$ *für alle* u*. Man beachte, daß wir keine Symmetrie für* $[u, v]$ *vorausgesetzt haben.*

Lemma 5.1 (von Lax–Milgram) *Sei* H *ein Hilbertscher Raum und* $[.,.]$ *eine koerzitive beschränkte Bilinearform auf* H*. Dann gibt es zu dem beschränkten linearen Funktional* $f \in H^*$ *ein eindeutiges* $u \in H$ *mit*

$$[v, u] = f(v) \text{ für alle } v \in H\,. \tag{5.5}$$

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

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

Die Abbildung A ist linear, denn es gilt $\big(v, A(\lambda x + \beta y) - \lambda A(x) - \beta A(y)\big) = 0$ 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 $[.,.]$ beschränkte Bilinearform (Sesqilinearform) ist, gilt $\|Au\|^2 = [Au, u] \leq K\|Au\| \, \|u\|$, 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 koerzitiv. Das lineare Funktional f stellen wir nach dem Rieszschen Darstellungssatz durch ein Element f_0 dar:

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

Die Gleichung (5.5) übersetzt sich daher in die Form

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

oder, äquivalent,

$$Au = f_0\,.$$

Da wir gezeigt haben, daß A beschränkt und koerzitiv ist, folgt mit Satz 5.2 die eindeutige Lösbarkeit von (5.5). $\square$

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

$$\sum_{i,k=1}^{n} a_{ik}(x)\,\xi_i\,\xi_k \geq \lambda_0 |\xi|^2 \qquad \text{für alle } \xi \in \mathbb{R}^n\,. \tag{5.6}$$

Ferner sei $c \in L^\infty(\Omega)$, $c \geq \lambda_1 > 0$. Sei $H = H^{1,2}(\Omega)$ und die Bilinearform $[.,.]$ 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 leicht, daß $[.,.]$ die Voraussetzungen des Satzes 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 beschränkt). Nach dem Lemma von Lax–Milgram gibt es daher eine eindeutig bestimmte Funktion u aus $H^{1,2}$ mit

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

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

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

mit natürlichen Randbedingungen (s. Skript Infini III).

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

Definition 5.4 *Die* Norm *einer beschränkten linearen Abbildung $A\colon H_1 \to H_2$, H_1, H_2 Hilbertsche Räume, ist die Größe*

$$\|A\| = \sup\big\{\|Au\| \,\big|\, \|u\| = 1\big\}$$

oder äquivalent

$$\|A\| = \inf\big\{K \in \mathbb{R} \big|\ \|Au\| \leq K\|u\| \qquad \textit{für alle } u \in H\big\}\,.$$

Man beweist leicht

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

sowie für $A\colon H_1 \to H_2,\ B\colon H_2 \to H_3$

$$\|AB\| \leq \|A\|\,\|B\|\,.$$

Satz 5.3 („Neumannsche Reihe“) *Sei H ein Hilbertscher Raum und $A\colon H \to H$ linear und beschränkt mit*

$$\|A\| \leq q < 1\,. \tag{5.7}$$

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

Beweis: Wir wollen die Gleichung $(I - A)u = f$ lösen. Mit der Bedingung (5.7) zeigt man leicht, daß die Reihe $Sf = f + Af + A^2f + A^3f + \ldots$ in H konvergiert. („Die Neumannsche Reihe $S = \sum\limits_{j=0}^{\infty} A^j$ konvergiert, wenn $\|A\| < 1$“) und ähnlich wie bei der Berechnung der geometrischen Reihe erhält man

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

und somit $(I - A)^{-1} = \sum\limits_{j=0}^{\infty} A^j$, $A^0 = I$.

Die Eindeutigkeit und Stetigkeit der Inversen folgt aus

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

□

6 Adjungierte Abbildungen

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

$$\varphi_u(v) = (A^*u, v)\,. \tag{6.1}$$

Definition 6.1 *Sei $A\colon H_1 \to H_2$ linear und beschränkt. Die vermöge (6.1) definierte Abbildung $A^*\colon H_2 \to H_1$ heißt die zu A* adjungierte Abbildung.

Satz 6.1 *A^* ist linear und beschränkt.*

Beweis:

$$\begin{aligned}\big(A^*(\alpha u + \beta w), v\big) &= (\alpha u + \beta w, Av) = \bar\alpha(u, Av) + \bar\beta(w, Av)\\ &= \bar\alpha(A^*u, v) + \bar\beta(A^*w, v) = (\alpha A^*u + \beta A^*w, v) \qquad \text{für alle } v \in H_2\end{aligned}$$

$\Rightarrow A^*(\alpha u + \beta w) = \alpha A^*u + \beta A^*w$.
Die Beschränktheit schließt man aus

$$\begin{aligned} &|(A^*u, v)| = |(u, Av)| \le \|u\|\;\|A\|\;\|v\|\\ \Rightarrow\quad &\sup_{\|v\|=1} |(A^*u, v)| = \|A^*u\| \le \|A\|\;\|u\|\,.\end{aligned}$$

□

Satz 6.2 *$A^{**} = A$, falls $A\colon H_1 \to H_2$ beschränkt, ferner ist $\|A\| = \|A^*\|$. (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.

Satz 6.3 *Es seien H_1, H_2 Hilbertsche Räume und $A\colon H_1 \to H_2$ 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 Nullraum von A^*.

Folgerung 6.1 (Fredholmsche Alternative) *Unter der Voraussetzung von Satz 6.3 gilt: Die Gleichung $Au = f$ ist genau dann lösbar, wenn $f \perp N(A^*)$.*

Beweis: 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 $(y, Au) = (A^*y, u) = 0$ für alle u, d.h. $y \perp A(H_1)$. In 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 $\|y - y_2\| = (y_1, y - y_2) = 0$, d.h. $y = y_2 \in V$. Damit folgt $N(A^*) \subset V$.

„$\supset$": Sei umgekehrt $y \in V$. Dann ist $v \perp A(H_1)$ und $0 = (y, Av) = (A^*y, v)$ 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 6.3 bewiesen. □

Beweis der Folgerung:

„$\Leftarrow$": Sei $f \perp N(A^*)$. Nach Satz 6.3 ist $f = f_1 + f_2\,,\ f_1 \in A(H)\,,\ f_2 \in N(A^*)$, somit $\|f - f_1\|^2 = (f_2, f - f_1) = 0$, d.h. $f = f_1 \in A(H_1)$, d.h. es gibt ein $u \in H_1$ mit $Au = f$.

„$\Rightarrow$": Ist f nicht orthogonal zu $N(A^*)$, so gibt es eine Zerlegung $f = f_1 + f_2$, $f_1 \in A(H_1)$, $f_2 \in N(A^*)$, $f_2 \neq 0$. Hätte f die Gestalt Au, wäre $f \perp N(A^*)$, und es folgte
$0 = (f - f_1, f_2) = \|f_2\|^2$, d.h. $f_2 = 0$

□

Beispiele von adjungierten Operatoren:

1. $H_1 = H_2 = L^2(\Omega)$, Ω offene Punktmenge des R^n,

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

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

$$A^*u = \int\limits_\Omega \overline{K(y,x)}\, u(y)\, dy\,.$$

2. $H_1 = H_2 = L^2(\Omega)$, $(Au)(x) := g(x)\, u(x)$. $A^*u(x) = \overline{g(x)}\, u(x)$.
3. $H_1 = H^{1,2}(I)$, $I = [a,b] \subset \mathbb{R}$, $H_2 = L^2(I)$, $Au = u'$.
   Nach Definition ist

$$(A^*v, u)_{H_1} = (v, u')_{L^2} \qquad \text{für alle } u \in H_1\,, \text{ d.h.}$$

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

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

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

d.h. y ist schwache Lösung der Differentialgleichung

$$-y'' + y + \int\limits_a^t v(s)\, ds = 0 \tag{6.2}$$

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

Beispiel zur Fredholmschen Alternative bei gewöhnlichen Differentialgleichungen:

Sei $H_1 = H_2 = H$, $H = H^{1,2}(I) =$ Sobolevraum über einem Intervall $I = [a,b]$. Sei A die vermöge des Rieszschen Darstellungssatzes definierte Abbildung $A\colon H \to H$, die definiert ist durch

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

Die Gleichung

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

besagt, daß u eine schwache Lösung der gewöhnlichen Differentialgleichung

$$-u'' + c\,u = f \qquad \text{a.e. in } I$$

ist mit den natürlichen Randbedingungen $u'(a) = u'(b) = 0$.
Da der Bildraum von A abgeschlossen ist (nicht trivial, beweisbedürftig!), 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

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

Dies bedeutet aber, daß v eine schwache Lösung der Differentialgleichung

$$-v'' + c(x)\,v = 0 \quad \text{a.e. in } I$$

mit den natürlichen Randbedingungen $v'(a) = v'(b) = 0$ ist.

Wir stellen fest:
Die Gleichung

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

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

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

ist.

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

Übungsaufgabe: Erläutere die Fredholmsche Alternative für das Randwertproblem

$$-u'' + du' + cu = f\,, \quad u'(a) = u'(b) = 0\,.$$

Hinweis: $N(A^*)$ besteht aus den Lösungen von

$$-v'' - (bv)' + cv = 0\,, \qquad v'(a) + d(a)\,v(a) = v'(b) + d(b)\,v(d) = 0.$$

Es fehlt noch der Nachweis, daß $A(H)$ abgeschlossen ist. Dies folgt aus der Koezitivitätsungleichung

$$(Au, u) \geq \alpha \, \|u'\|_{L^2}^2 - d \, \|u\|_{L^2}^2 \,, \quad u \in H^{1,2}$$

mit Zahlen $\alpha > 0$ und $d \in \mathbb{R}$. Der Beweis läßt sich abstrakt führen, wenn der Begriff der Kompaktheit und die kompakte Einbettung von L^2 nach $H^{1,2}$ behandelt ist. Alternativ läßt sich mit Hilfe von Fourieranalysis argumentieren.

Adjungierte Abbildungen unbeschränkter Operatoren

Faßt man, wie eben, die Differentation $Au = u'$ 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 $A^* = -u'$ ist. Dies erreicht man, indem man lineare, nicht notwendig beschränkte Abbildungen betrachtet, die nicht auf dem gesamten zugrundeliegenden Hilbertraum erklärt sind.

Beispiel: $H = L^2(\mathbb{R})$, $Au = u'$, $D(A) =$ Definitionsbereich von $A = C_0^\infty(\mathbb{R})$ oder auch

$$S = \big\{ u \in C^\infty \,\big|\, |\nabla^m u(x)| \cdot |x|^s \to 0 \text{ für } |x| \to \infty, \, m \in \mathbb{N}, \, s \in \mathbb{N} \,\big\}.$$

A ist ein *unbeschränkter* Operator.

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

Definition 6.2 (adjungierte Abbildung) *Sei $A\colon D(A) \to H$ eine lineare Abbildung, wobei $D(A)$ ein linearer dichter Teilraum von H ist. Dann hat die adjungierte Abbildung $A^*\colon D(A^*) \to H$ den Definitionsbereich*

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

A^*v ist dasjenige durch den Rieszschen Darstellungssatz gegebene Element von H, für welches

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

gilt.

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

$$|\varphi_v(u_j - u_k)| = \big(v, A(u_j - u_k)\big) \leq K \|u_j - u_k\|, \tag{6.3}$$

läßt sich φ_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 $u_j \to u$ in H und (u_j) ist Cauchyfolge. Wegen (6.3) konvergiert dann $\varphi(u_j)$ und der Limes wird als $\varphi(u)$ 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 S), $Au = u'$. Sei $v \in D(A^*)$, d.h. $v \in L^2(\mathbb{R})$ und

$$|(v, u')| \leq K_u \|u\|_{L^2} \quad \text{für alle } u \in L^2(\mathbb{R})\,. \tag{6.4}$$

Man kann dann zeigen, daß v absolutstetig ist und $v' \in L^2(\mathbb{R})$ liegt, also $v \in H^{1,2}$ (s. Analysis IV–Skript, Teil über Variationsrechnung). Umgekehrt gilt (6.4) genau für $v \in H^{1,2}$. Daher gilt

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

und damit

$$-(v', u) = (v, u')\,, \tag{6.5}$$

d.h. $A^*u = -u'$. Die Randterme in (6.5) fallen weg, da $u, v \in L^2(\mathbb{R})$ ist.

Definition 6.3 *Eine lineare Abbildung $A\colon D(A) \subset H \to H$, $D(A)$ linearer dichter Teilraum von H, heißt* selbstadjungiert, *wenn*

1. $D(A) = D(A^*)$ *und*
2. $Au = A^*u$ *für alle $u \in D(A)$.*

Demzufolge ist die Abbildung

$$Au = \tfrac{1}{\sqrt{-1}}\, u'\,, \qquad D(A) = C_0^\infty(\mathbb{R}) \text{ oder } S$$

nicht selbstadjungiert, da $D(A)$ zu klein ist, $D(A) \subset L^2(\mathbb{R})$. Die Abbildung $Au = \frac{1}{\sqrt{-1}}\, u'$ (a.e.) mit $D(A) = H^{1,2}(\mathbb{R})$ ist jedoch 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.

Definition 6.4 *Eine unbeschränkte lineare Abbildung $A\colon D(A) \subset H \to H$, $D(A)$ dicht in H, heißt* abgeschlossen, *wenn aus $u_j \to u$, $u_j \in D(A)$ und $Au_j \to f$ die Implikation $u \in D(A)$, $Au = f$ folgt.*

Eine Abbildung ist also abgeschlossen, wenn der Definitionsbereich nicht „unnötig klein“ ist.

7 Separable Hilberträume und Orthogonalsysteme

Wir erinnern:

Definition 7.1 *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 $u_j \to u$.*

Definition 7.2 *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. $L^2(\Omega), H^{1,2}(\Omega)$.*

Definition 7.3 *Ein* Orthogonalsystem *ist eine Menge $\{\, \varphi_j \in H, \varphi_j \neq 0 \,|\, j \in \mathbb{N}\}$ mit*

$$(\varphi_k, \varphi_j) = 0, \qquad k \neq j.$$

Das Orthogonalsystem heißt Orthonormalsystem*, wenn zusätzlich $\|\varphi_j\| = 1,\ j \in \mathbb{N}$. Ein Orthogonalsystem heißt* vollständig *in H, wenn* $\operatorname{span}\{\, \varphi_j \,|\, j \in \mathbb{N}\}$ *dicht in H ist.*

Lemma 7.1 *Es sei $M = \{\, v_j \in H \,|\, j \in \mathbb{N}\}$. Dann gibt es Orthonormalsysteme $\{\, \varphi_j \in H, \varphi_j \neq 0 \,|\, j \in \mathbb{N}\}$ mit* $\operatorname{span} M = \operatorname{span}\{\varphi_j\}$.

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 φ_j schon konstruiert. Man setze

$$\tilde{\varphi}_{j+1} = v_{j+1} - \sum_{l=1}^{j} (v_{j+1}, \varphi_l)\, \varphi_l \tag{7.1}$$

$$\varphi_{j+1} = \frac{\tilde{\varphi}_{j+1}}{\|\tilde{\varphi}_{j+1}\|} \tag{7.2}$$

Offensichtlich ist $\varphi_{j+1} \perp \varphi_l,\ 1 \leq l \leq j$. Damit sind die φ_j durch vollständige Induktion definiert. Durch die Konstruktion erreicht man

$$\operatorname{span}\{\, \varphi_j \,|\, j \in \mathbb{N}\,\} = \operatorname{span}\{M\}\,.$$

□

Aus den Definitionen und Lemma 7.1 folgt offensichtlich

Satz 7.1 *Es sei H ein separabler Hilbertraum. Dann gibt es ein vollständiges Orthonormalsystem.*

Lemma 7.2 *Sei $\{\varphi_j\}$ ein Orthonormalsystem in einem Hilbertraum H. Dann ist für jedes $w \in H,\ N \in \mathbb{N}$*

$$\Big\| \sum_{j=1}^{N} (w, \varphi_j)\, \varphi_j \Big\|^2 = \sum_{j=1}^{N} |(w, \varphi_j)|^2 \leq \|w\|^2 .$$

Folgerung: $\sum\limits_{j=1}^{\infty} |(w, \varphi_j)|^2 < \infty$.

Beweis: Es gilt

$$\begin{aligned} 0 &\leq \Big\| w - \sum_{j=1}^{N} (w, \varphi_j)\, \varphi_j \Big\|^2 = \\ &= \|w\|^2 + \sum_{j=1}^{N} |(w, \varphi_j)|^2 - \sum_{j=1}^{N} \big[\big(w, (w, \varphi_j)\, \varphi_j\big) + \big((w, \varphi_j)\, \varphi_j, w\big) \big] = \\ &= \|w\|^2 + \sum_{j=1}^{N} |(w, \varphi_j)|^2 - \sum_{j=1}^{N} \Big[\overline{(w, \varphi_j)}\, (w, \varphi_j) + (w, \varphi_j)\, \overline{(w, \varphi_j)} \Big] = \\ &= \|w\|^2 - \sum_{j=1}^{N} |(w, \varphi_j)|^2 . \end{aligned}$$

Hieraus folgt die Aussage des Lemmas. □

Satz 7.2 *Sei $\{\varphi_j\}$ ein vollständiges Orthonormalsystem in einem Hilbertraum H. Dann läßt sich jedes Element $u \in H$ als konvergente, verallgemeinerte Fourierreihe*

$$u = \sum_{j=1}^{\infty} (u, \varphi_j)\, \varphi_j \tag{7.3}$$

darstellen, und es gilt

$$\|u\|^2 = \sum_{j=1}^{\infty} |(u, \varphi_j)|^2 . \tag{7.4}$$

Beweis: Nach Satz 7.1 gibt es zu $u \in H$ eine Folge $u^N = \sum_{j=1}^{N} \mu_j^N \varphi_j \to u$. Es ist $\mu_j^N = (u^N, \varphi_j)$, wie man sich durch Multiplikation im Sinne des Skalarproduktes mit φ_j überlegt. Es gilt wegen Lemma 7.2

$$\Big\| \sum_{j=1}^{N} (u, \varphi_j)\, \varphi_j - u^N \Big\|^2 = \Big\| \sum_{j=1}^{N} (u - u^N, \varphi_j)\, \varphi_j \Big\|^2 \leq \big\| u - u^N \big\|^2 \to 0\,.$$

Da $u^N \to u$, folgt somit

$$\Big\| \sum_{j=1}^{N} (u, \varphi_j)\, \varphi_j - u \Big\| \to 0\,. \tag{7.5}$$

Aus (7.5) folgt

$$\sum_{j=1}^{N} |(u, \varphi_j)|^2 - \sum_{j=1}^{N} (u, \varphi_j)\,(\varphi_j, u) - \sum_{j=1}^{N} (u, \varphi_j)\,\overline{(u, \varphi_j)} + \|u\|^2 \to 0\,,$$

d.h.

$$\|u\|^2 - \sum_{j=1}^{N} |(u, \varphi_j)|^2 \to 0\,.$$

Da

$$\|u\|^2 - \sum_{j=1}^{N} |(u, \varphi_j)|^2 \geq 0$$

nach Lemma (7.2), gilt

$$\|u\|^2 - \sum_{j=1}^{N} |(u, \varphi_j)|^2 \to 0\,.$$

Damit ist (7.4) bewiesen. Man kann sich überlegen, daß die Vollständigkeit eines Orthonormalsystems äquivalent zu Relation (7.4) ist. □

Die verallgemeinerten Fourierkoeffizienten besitzen eine *Minimaleigenschaft*. Die Lösung des Minimumproblems „Suche $\mu_1, \dots, \mu_N \in \mathbb{R}$ oder $\mathbb{C}$, so daß

$$\|u - \sum_{j=1}^{N} \mu_j \varphi_j\|$$

minimal ist," lautet $\mu_j = (u, \varphi_j)$.

Satz 7.3 *Jeder separable Hilbertraum H ist isometrisch–isomorph zum Hilbertraum l^2.*

Beweis: Sei $\{\varphi_j\}$ ein vollständiges Orthonormalsystem in H. Jedes u besitzt eine Darstellung

$$u = \sum_{j=1}^{\infty} (u, \varphi_j)\, \varphi_j \, .$$

Dem Element u ordnen wir den Vektor

$$Tu = \big((u, \varphi_1), (u, \varphi_2), (u, \varphi_3), \ldots\big)$$

zu. Offensichtlich ist T eine umkehrbar eindeutige Abbildung von H auf l^2. Insbesondere ist

$$(u, v) = \sum_{j=1}^{\infty} (u, \varphi_j)\, \overline{(v, \varphi_j)} = (Tu, Tv)_{l^2} \, .$$

$\square$

Satz 7.3 ist nützlich, da gewisse Sätze über Hilberträume in dem recht anschaulichen Raum l^2 bewiesen werden können. Ein Beispiel werden wir in dem nächsten Kapitel sehen.

8 Schwache Konvergenz und schwache Kompaktheit

Neben der Normkonvergenz gibt es in Hilberträumen einen weiteren, sehr wichtigen Konvergenzbegriff, nämlich die *schwache* Konvergenz.

Definition 8.1 *Eine Folge $(u_j)_{j\in\mathbb{N}}$, $u_j \in H$ in einem Hilbertraum H* konvergiert schwach *gegen ein Element u, in Zeichen*

$$u_j \rightharpoonup u \quad \textit{oder} \quad u_j \rightharpoonup u \textit{ schwach in } H \ (j \to \infty)$$

genau dann, wenn für alle $v \in H$

$$(u_j, v) \to (u, v) \qquad (j \to \infty)\,.$$

Im Unterschied hierzu bezeichnet man die Konvergenz bezüglich der Norm als starke *Konvergenz.*

Beispiel: Sei $\{\varphi_j\}$ ein Orthonormalsystem in H. Dann konvergiert

$$\varphi_j \rightharpoonup 0 \quad \text{schwach in } H\,.$$

Beweis: Sei $v \in H$. Es gilt

$$v = \sum_{j=1}^{\infty} (v, \varphi_j)\, \varphi_j\,.$$

Da nach Satz 7.3 $\sum\limits_{j=1}^{\infty} |(v,\varphi_j)|^2$ konvergiert, gilt $(v, \varphi_j) \to 0$ □

Da $\|\varphi_j\| = 1$, kann (φ_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.

Satz 8.1 („Schwache Kompaktheit beschränkter Mengen im Hilbertraum“)

Sei $(u_m)_{m\in\mathbb{N}}$, $u_m \in H$ eine beschränkte Folge in einem Hilbertschen Raum H. Dann gibt es eine schwach konvergente Teilfolge $(u_m)_{m\in\Lambda}$, $\Lambda \subset \mathbb{N}$.

Beweis: Wir führen das Problem auf den Fall des separablen Hilbertraumes zurück, indem wir die Abschließung H_0 von $\operatorname{span}\{u_m \,|\, m \in \mathbb{N}\}$ 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 vorangegangenes Kapitel), und Satz 8.1 übersetzt sich in die Aussage:

„Sei $c^l = (c_1^l, c_2^l, \dots) \in l^2$, $\sum\limits_{j=1}^{\infty} |c_j^l|^2 \le K$, $l \in \mathbb{N}$. Dann gibt es eine Teilfolge $\Lambda \in \mathbb{N}$ und ein Element $c = (c_1, c_2, \dots) \in l^2$ mit

$$\sum_{j=1}^{\infty} c_j^l\, \overline{\mu}_j \to \sum_{j=1}^{\infty} c_j\, \overline{\mu}_j \qquad (l \to \infty,\ l \in \Lambda)$$

für alle $\mu = (\mu_1, \mu_2, \dots) \in l^2$. “

Der Beweis geschieht folgendermaßen: Nach dem Weierstraßschen Doppelfolgensatz gibt es eine Teilfolge $\Lambda \subset \mathbb{N}$, so daß

$$c_j^l \to c_j \qquad (l \to \infty,\ l \in \Lambda)$$

für alle $j \in \mathbb{N}$. (Man setzt das bekannte Diagonalverfahren ein.) Diese Folge Λ ist bereits unsere gesuchte Teilfolge, denn es gilt zunächst

$$\sum_{j=1}^{N} |c_j|^2 = \lim_{l \to \infty} \sum_{j=1}^{N} |c_j^l|^2 \le K\,,$$

und somit

$$\sum_{j=1}^{N} |c_j|^2 \le K\,,$$

d.h. $c \in l^2$. Weiterhin gilt

$$\Big|\sum_{j=1}^{\infty} (c_j - c_j^l)\, \overline{\mu}_j\Big| \le \underbrace{\Big|\sum_{j=1}^{N} (c_j - c_j^l)\, \overline{\mu}_j\Big|}_{=:\,A} + \underbrace{\Big|\sum_{j=N+1}^{\infty} (c_j - c_j^l)\, \overline{\mu}_j\Big|}_{=:\,B}\,.$$

Der Term A geht gegen Null für $l \to \infty$, $l \in \Lambda$, N fest, und B läßt sich abschätzen durch

$$\begin{aligned} B &\le \Big(\sum_{j=N+1}^{\infty} |c_j - c_j^l|^2\Big)^{1/2} \Big(\sum_{j=N+1}^{\infty} |\overline{\mu}_j|^2\Big)^{1/2} \le \big(\|c\| + \|c^l\|\big)\Big(\sum_{j=N+1}^{\infty} |\overline{\mu}_j|^2\Big)^{1/2} \\ &\le 2K\Big(\sum_{j=N+1}^{\infty} |\overline{\mu}_j|^2\Big)^{1/2} \end{aligned}$$

Da $\sum\limits_{j=1}^{\infty} |\mu_j|^2$ konvergiert, läßt sich für $N \geq N(\varepsilon)$, N fest, erreichen, daß $B < \varepsilon$ ist. Damit erhält man

$$\left|(c,\mu) - (c^l,\mu)\right| < o(1) + \varepsilon \qquad (l \to \infty,\ l \in \Lambda)\,,$$

und die schwache Konvergenz $c^l \to c$, $l \in \Lambda$, ist bewiesen.

Damit gilt $(v, u_m) \to (v, u)$ $(m \to \infty,\ m \in \Lambda)$, $u = \sum c_j\,\varphi_j$, $v \in H_0$ und damit auch $(w, u_m) \to (w, u)$, $w = w_1 + w_2$, $w_1 \in H_0$, $w_2 \in H_0^{\perp}$. □

Es folgen weitere Betrachtungen zur schwachen Konvergenz sowie ihrem Verhältnis zur starken Konvergenz.

Satz 8.2 *Sei H ein Hilbertscher Raum und (u_m), $u_m \in H$, schwach konvergent. Dann ist (u_m) beschränkt.*

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, also „punktweise“ beschränkt, und das Prinzip der gleichmäßigen Beschränktheit ergibt $\sup\limits_{\|v\|=1} |\varphi_m(v)| \leq K$, woraus $\|u_m\| \leq K$ folgt.

Satz 8.3 *Sei $A\colon H_1 \to H_2$ beschränkt und linear, H_1, H_2 Hilbertsche Räume. Dann ist A stetig bezüglich der schwachen Konvergenz von H_1 und H_2.*

Beweis: Sei $u_m \rightharpoonup u$ schwach in H_1. Dann gilt $(Au_m, v) = (u_m, A^*v) \to (u, A^*v) = (Au, v)$. □

Die Umkehrung von Satz 8.3 können wir im Augenblick nicht beweisen.

Die Umgebungsbasis der schwachen Topologie

Die Umgebungsbasis eines Elementes $u \in H$ in der *starken* Topologie sind die ε-Umgebungen $U_\varepsilon(u) = \bigl\{x \in H \bigm| \|u - x\| < \varepsilon\bigr\}$, $\varepsilon > 0$. In der schwachen Topologie sind es die Mengen

$$U_{\varepsilon,\varphi_1,\varphi_2,\dots,\varphi_m}(u) = \bigl\{x \in H \bigm| |(\varphi_j, u - x)| < \varepsilon,\ j = 1, \dots, m\bigr\}.$$

Hierbei sind alle $\varepsilon > 0$ und alle $m \in \mathbb{N}$ zu verwenden. Im separablen Fall könnte man sich auf abzählbar viele φ_m beschränken. Dann wäre das sogenannte erste Abzählbarkeitsaxiom, d.h. jedes Element hat eine abzählbare Umgebungsbasis, erfüllt (ε 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 zunächst zwischen folgenkompakt und überdeckungskompakt unterscheiden.

Ein weiterer, grundlegender, für die Variationsrechnung und Optimierung sehr nützlicher Satz ist

Satz 8.4 (Banach–Saks) *Es sei H ein Hilbertscher Raum und (u_m), $u_m \in H$, eine beschränkte Folge. Dann gibt es eine Teilfolge $\Lambda \subset \mathbb{N}$, so daß die arithmetischen Mittel $\frac{1}{N}\sum\limits_{j=1}^{N} u_{m_j}$ stark in H konvergieren.*

Beweis: O.B.d.A. nehmen wir an, daß bereits $u_m \rightharpoonup u$ bezüglich der *schwachen* Konvergenz (s. Satz 8.1) und $u = 0$ 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

$$(u_m, u_{m_l}) \to 0 \quad (m \to \infty, l = 1, \dots, k)$$

aus und erhalten so den Index m_{k+1}.
Die Indexfolge (m_j) ist die gewünschte, denn es gilt

$$\begin{aligned} \Big\| \frac{1}{N} \sum_{j=1}^{N} u_{m_j} \Big\|^2 &= \frac{1}{N^2} \sum_{j=1}^{N} \|u_{m_j}\|^2 + \frac{2}{N^2} \sum_{k=1}^{N} \sum_{j=1}^{k-1} \mathcal{R}e(u_{m_j}, u_{m_k}) \\ &\leq \frac{N}{N^2} \cdot K + \frac{2}{N^2} \sum_{k=1}^{N} \sum_{j=1}^{k-1} \frac{1}{j^2} \leq \frac{K}{N} + \frac{2}{N^2} N \cdot \frac{\pi^2}{6} \quad \to 0 \quad (N \to \infty)\,. \end{aligned}$$

□

Wir geben zwei Anwendungen des Satzes von Banach–Saks an:

Satz 8.5 *Sei (u_m) eine schwach konvergente Folge vektorwertiger Funktionen aus $L^2(\Omega; \mathbb{R}^N)$, Ω offen, und C eine abgeschlossene konvexe Menge des $\mathbb{R}^N$, so daß $u_m(x) \in C$ für alle $x \in \mathbb{R}^N$ mit Ausnahme einer Menge vom Maß Null. Dann gilt für den schwachen Limes u der u_m, daß $u(x) \in C$ für alle $x \in \mathbb{R}$ mit Ausnahme einer Menge vom Maß Null.*

Anmerkung: Wenn C *nicht* konvex ist, ist die entsprechende Aussage i. A. falsch.

Beweis: Ist $u_m(x) \in C$, so gilt auch für die arithmetischen Mittel

$$\frac{1}{N}\sum_{j=1}^{N} u_{m_j}(x) \in C\,.$$

Da es sich durch Auswahl von Teilfolgen erreichen läßt, daß $\frac{1}{N}\sum\limits_{j=1}^{N} u_{m_j}(x) \to u$ in L^2 konvergiert $(N \to \infty)$, erhält man durch Auswahl einer weiteren Teilfolge, daß $\frac{1}{N}\sum\limits_{j=1}^{N} u_{m_j}$ fast überall konvergiert. Daraus folgt für den Limes u, daß $u(x) \in C$ für alle $x \in \mathbb{R}^n - E,\ \mu(E) = 0$ gilt. □

Allgemeiner zeigt man mit der gleichen Methode

Satz 8.6 *Es sei C eine konvexe, abgeschlossene Menge in einem Hilbertraum H und*
$u_m \in C,\ u_m \rightharpoonup u$ schwach in H $(m \to \infty)$. Dann gilt $u \in C$.

Interessant ist in diesem Zusammenhang der folgende Satz, den wir ohne Beweis angeben.

Satz 8.7 *Sei $f\colon \mathbb{R} \to \mathbb{R}$ stetig, und die durch $Tu(x) = f(u(x))$ definierte Abbildung bilde $L^2(\mathbb{R}^n)$ in sich ab und sei bezüglich der schwachen Topologie im Bild und Urbild stetig. Dann ist f eine lineare Funktion, d.h. $f(x) = a \cdot x + \beta,\ a \in \mathbb{R}^n,\ \beta \in \mathbb{R}$.*

Wir präsentieren noch eine Anwendung des Satzes von Banach-Saks in der Optimierung:

Satz 8.8 *Es sei C eine nichtleere abgeschlossene Menge eines Hilbertschen Raumes H und $f\colon C \to \overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$ stetig in der starken Topologie von H. Ferner sei f konvex und koerzitiv auf C, d.h. $f(u) \to \infty$ für $\|u\| \to \infty,\ u \in C$. Dann existiert ein Minimum von f auf C.*

Beweis: O.B.d.A. sei $f \neq \text{const} = \infty$. Sei (u_m), $u_m \in C$ eine Minimalfolge, d.h.

$$f(u_m) \to \inf_C f < \infty.$$

Wegen der Koerzivität von f ist $\|u_m\| \leq K$ und es gibt eine Teilfolge, so daß, nach Umnummerierung,

$$u_m \rightharpoonup u \text{ schwach} \qquad \text{und} \qquad \tfrac{1}{N}\sum_{j=1}^{N} u_j \to u \text{ stark.}$$

Es ist $\frac{1}{N}\sum_{j=1}^{N} u_j \in C$, also auch $u \in C$. Ferner gilt $\frac{1}{N}\sum_{j=1}^{N} f(u_j) \to \inf f$ und

$$f(u) \leq \varepsilon + f\left(\frac{1}{N}\sum_{j=1}^{N} u_j\right) \leq \varepsilon + \frac{1}{N}\sum_{j=1}^{N} f(u) \leq 2\varepsilon + \inf f\,.$$

Anschließender Grenzübergang $\varepsilon \to 0$ ergibt die Behauptung. $\square$

9 Kompakte lineare Abbildungen

Definition 9.1 *Eine Menge M eines Hilbertraumes H heißt* relativ folgenkompakt*, wenn jede Folge (u_m), $u_m \in M$ eine Teilfolge besitzt, welche in H konvergiert.*

Unpräziserweise sagt man auch, daß M kompakt ist, meint aber relativ folgenkompakt.

Definition 9.2 *Eine Menge M eines Hilbertraumes H heißt* folgenkompakt*, wenn sie relativ folgenkompakt ist und alle Häufungspunkte 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_ι, $\iota \in I$, eine endliche Teilüberdeckung besitzt, d.h. es gibt eine *endliche* Indexmenge $I_1 \subset I$ mit $M \subset \bigcup\limits_{\iota \in I_1} O_\iota$. 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 Kapitel sieben bereits diskutiert haben, sind beschränkte Mengen im Hilbertraum nicht notwendig relativ folgenkompakt - für die schwache Topologie gilt dies jedoch.

Fundamental für die Anwendungen der Funktionalanalysis ist der Begriff des kompakten Operators (Abbildung).

Definition 9.3 *Eine Abbildung $A\colon H_1 \to H_2$ mit Hilberträumen H_1, H_2 heißt* kompakt*, wenn A beschränkte Mengen in relativ kompakte überführt.*
Die Abbildung $A\colon H_1 \to H_2$ heißt vollstetig*, wenn sie schwach konvergente Folgen in stark konvergente überführt.*

Bei diesen Definitionen muß A nicht notwendig linear sein.

Lemma 9.1 *Seien H_1, H_2 Hilberträume. Die lineare Abbildung $A\colon H_1 \to H_2$ ist genau dann kompakt, wenn sie vollstetig ist. Insbesondere ist A beschränkt.*

Beweis:

„$\Leftarrow$“ Sei A vollstetig und M beschränkt. Sei (u_m) eine Folge mit $u_m \in M$. Zu zeigen ist, daß (Au_m) eine konvergente Teilfolge hat. Da beschränkte Mengen schwach kompakt sind, gibt es eine Teilfolge $\Lambda \subset \mathbb{N}$, so daß $u_m \rightharpoonup u$ $(m \in \Lambda,\ m \to \infty)$ schwach in H_1. Da A vollstetig ist, konvergiert $Au_m,\ m \in \Lambda$.

„$\Rightarrow$“ 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 $\|Av_m\| \to \infty$. Dies widerspricht der Forderung, daß (Av_m) relativ kompakt ist.

Nach den Ergebnissen von Kapitel 7 sind beschränkte 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 Λ 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 $Au_m \to f$ stark in H_2. Es folgt $\|f - Au\| > \varepsilon$.
Andererseits gilt $(v, Au_m) = (A^*v, u_m) \to (A^*v, u) = (v, Au)$ und $(v, Au_m) \to (v, f)$. Damit folgt $Au = f$. Es muß daher die Konvergenz $Au_m \to Au$ stattfinden.

□

Beispiele kompakter linearer Abbildungen:

Typische Beispiele sind Integraloperatoren: Sei $\Omega \subset \mathbb{R}^n$ beschränkt und offen, $K \in L^2(\Omega \times \Omega)$ und

$$Au(x) := \int_\Omega K(x,y)\, u(y)\, dy \tag{9.1}$$

Lemma 9.2 *Die durch (9.1) definierte Abbildung $A : L^2(\Omega) \to L^2(\Omega)$ ist vollstetig.*

Beweis: Wir beweisen die Aussage zunächst für Kerne $K \in C\big(\overline{\Omega} \times \overline{\Omega}\big)$. Es gilt für $u_m \in L^2$, $\|u_m\| \le const$

$$\begin{aligned}\|Au_m(x) - Au_m(z)\|_\infty &\le \sup\left|\int_\Omega \big(K(x,y) - K(z,y)\big)\, u(y)\, dy\right| \le \\ &\le \sup_{y\in\Omega} |K(x,y) - K(z,y)| \cdot \int_\Omega |u(y)|\, dx \le \\ &\le \sup_{y\in\Omega} |K(x,y) - K(z,y)| \cdot |\Omega|^{1/2}\, \|u\|_{L^2(\Omega)}\,.\end{aligned}$$

Da K gleichmäßig stetig ist, gibt es für $\varepsilon > 0$ ein $\delta > 0$, so daß $\sup\limits_{y\in\Omega} |K(x,y) - K(z,y)| < \varepsilon$ ist für $\|x - z\| < \delta$.

Die Funktionen Au_m sind daher *gleichgradig* stetig – beschränkt in $L^2(\Omega)$ sind sie ebenfalls. Nach dem Satz von Arzela–Ascoli gibt es daher eine gleichmäßig konvergente Teilfolge. Konvergenz in $L^\infty(\Omega)$ impliziert L^2–Konvergenz bei beschränktem Gebiet. 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 der Raum der gleichmäßig stetigen Funktionen dicht in L^2 ist, gibt es eine Folge von Funktionen $K_m \in C(\overline{\Omega} \times \overline{\Omega})$ mit

$$\int\limits_\Omega \int\limits_\Omega |K - K_m|^2 \, dx \, dy \to 0 \qquad (m \to \infty)\,.$$

Wir setzen $A_m u(x) := \int\limits_\Omega K_m(x,y)\, u(y)\, dy$ und erhalten

$$\begin{aligned} \|A - A_m\|^2 &= \sup_{\|u\|_{L^2}=1} \int\limits_\Omega \left| \int\limits_\Omega \big(K(x,y) - K_m(x,y)\big)\, u(y)\, dy \right|^2 dx \le \\ &\le \int\limits_\Omega \int\limits_\Omega |K(x,y) - K_m(x,y)|^2 \, dx\, dy \to 0 \end{aligned}$$

Da A_m vollstetig ist, impliziert der folgende Satz die Aussage von Lemma 9.2. □

Satz 9.1 *Seien H_1, H_2 Hilbertsche Räume und $A_m\colon H_1 \to H_2$ linear und vollstetig. Sei*

$$\|A - A_m\| \to 0 \qquad (m \to \infty)\,.$$

Dann ist A vollstetig.

Interessanterweise führt also die Konvergenz bezüglich der Operatornorm nicht aus der Menge der vollstetigen Abbildungen heraus.

Beweis: Sei $u_j \rightharpoonup u$ schwach in H_1. Es gilt $\|u_j\| \le K$ gleichmäßig. Sei $\varepsilon > 0$ vorgegeben. Dann gilt

$$\|A - A_m\| < \frac{\varepsilon}{4K}$$

für mindestens ein m. Es konvergiert $A_m u_j \to A_m u$ $(j \to \infty)$ stark in H_2. Daraus folgt

$$\begin{aligned}\|Au_j - Au\| &= \|A_m u_j - A_m u + (A - A_m)\, u_j - (A - A_m)\, u\| \leq \\ &\leq \|A_m u_j - A_m u\| + |\!|\!|A - A_m|\!|\!|\, \|u_j\| + |\!|\!|A - A_m|\!|\!|\, \|u\| < \\ &< o(1) + \frac{\varepsilon}{2} < \varepsilon\end{aligned}$$

□

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.)

Satz 9.2 *Sei $J\colon H_1 \to L^2$ 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 $u_m \to u$ 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

$$|u_m(\eta) - u_m(\xi)| = \left|\int_\xi^\eta u_m'\, ds\right| \leq |\eta - \xi|^{1/2} \left(\int_I |u_m'|^2\, dt\right)^{1/2} \leq K|\eta - \xi|^{1/2}\,,$$

d.h. die u_m sind gleichgradig stetig. Die Beschränktheit folgt, da

$$\left|\fint_I u_m(\eta)\, d\eta - u_m(\xi)\right| \leq \fint_I |u_m(\eta) - u(\xi)|\, d\eta \leq K \fint_I |\eta - \xi|^{1/2}\, d\eta \leq K \cdot |I|^{\frac{1}{2}}$$

($\fint_I$ = Mittelwert über I), und $|\int u_m(\eta)\, d\eta| \leq |I|^{1/2}\|u_m\|_{L^2}$. Nach dem Satz von Arzela-Ascoli gibt es daher eine Teilfolge, die gleichmäßig und damit auch stark in L^2 konvergiert.

(ii) Der Fall $H_1 = H_0^{1,2}(\Omega)$

Wir werden noch andere Beweise kennenlernen; diesmal machen wir es über Fourieranalysis.

Da $H_0^{1,2}(\Omega)$ der Abschluß von $C_0^\infty(\Omega)$ in der $H^{1,2}(\Omega)$–Norm ist, können wir Ω als Teilraum eines Würfels $Q = [-\pi L, \pi L]^n$ betrachten. Die $H_0^{1,2}(\Omega)$–Funktionen

werden periodisch auf ganz $\mathbb{R}^n$ fortgesetzt mit Q als Periodizitätswürfel. Sei nun $u_m \in H_0^{1,2}(Q)$ beschränkt.
u_m besitzt eine Fourierentwicklung

$$u_m = \sum_k c_k^m e^{ikx/L} ,$$

wobei die Summation über die Multiindizes $k = (k_1, k_2, \dots , k_n), -\infty < k_j < \infty$, läuft. Die Beschränktheit der Folge $\|u_m\|_{H^{1,2}}$ zieht die Beschränktheit von

$$\sum_k |k|^2 |c_k^m|^2 \le K , \qquad |c_{0,0,\dots,0}^m|^2 \le K \tag{9.2}$$

nach sich. Aus (9.2) 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 $c^m \to c$ *komponentenweise*, so folgt aus (9.2) $\sum\limits_k |k|^2 \, |c_k|^2 \le K$ und

$$\begin{aligned}\sum_k |c_k^m - c_l|^2 &\le \sum_{|k|\le N} |c_k^m - c_k|^2 + \frac{1}{N^2} \sum_{|k|\ge N} |k|^2 \, |c_k^m - c_k^2| \\ &\le o(1) + \frac{K}{N^2} < \varepsilon \qquad \text{für } m \ge m(\varepsilon) .\end{aligned}$$

Auf die Fouriersche Reihe übertragen, impliziert die Konvergenz $c^m \to c$ in l^2 die Konvergenz $u^m \to u$ stark in L^2, wobei c der Vektor $\in l^2$ der Fourierkoeffizienten der u definierenden Reihe ist.

□

Satz 9.2 läßt sich auch folgendermaßen interpretieren:

Lemma 9.3 *Sei $J\colon H_0^{1,2}(\Omega)$ bzw. $H^{1,2}(I) \to H^{1,2}(\Omega)$ die durch den Rieszschen Darstellungssatz gegebene Abbildung mit*

$$(Ju, v)_{H^{1,2}} = (u, v)_{L^2} \qquad \textit{für alle } u, v \in H_0^{1,2} \textit{ bzw. } H^{1,2}. \tag{9.3}$$

Dann ist J kompakt.

Beweis: Sei $\|u_m\|_{H^{1,2}} \le K$. Dann folgt aus (9.3) mit $v = Ju_m$

$$\|Ju_m\|_{H^{1,2}} \le \|u_m\|_{H^{1,2}} \le K .$$

Die Ju_m sind also beschränkt. Andererseits gibt es wegen Satz 9.2 eine Teilfolge Λ, so daß $(u_m)_{m\in\Lambda}$ eine Cauchy–Folge in L^2 ist. Daraus folgt

$$\sup_{\|v\|_{L^2}=1} (Ju_m - Ju_k, v)_{H^{1,2}} = \|u_m - u_k\|_{L^2} \to 0 \qquad (m, k \to \infty,\ m, k \in \Lambda)\,,$$

woraus $\|Ju_m - Ju_k\|_{H^{1,2}} \to 0$ folgt. □

Als erste Anwendung der Begriffe werden wir in diesem Kapitel die Endlichdimensionalität des Nullraumes von Operatoren der Gestalt $A + K$, A koerzitiv und beschränkt, K kompakt, und verwandte Fragen behandeln. Zuvor benötigen wir noch

Satz 9.3 *Sei $K\colon H_1 \to H_2$ lineare, vollstetige Abbildung eines Hilbertraumes H_1 in einen Hilbertraum H_2. Dann ist die adjungierte Abbildung $K^*\colon H_2 \to H_1$ vollstetig.*

Beweis: Sei $v_m \rightharpoonup v$ schwach in H_2. Nach Definition des Begriffs „Supremum“ gibt es ein Element u_m mit $\|u_m\| = 1$ und Zahlen $\varepsilon_m \to 0$, so daß

$$\|K^*v_m - K^*v\| = \sup_{\|u\|=1} (K^*v_m - K^*v, u) = (K^*v_m - K^*v, u_m) + \varepsilon_m\,.$$

Daraus folgt

$$\|K^*v_m - K^*v\| = (v_m - v, Ku_m) + \varepsilon_m\,. \tag{9.4}$$

Da die Folge (Ku_m) kompakt ist und $v_m - v \rightharpoonup 0$ schwach konvergiert, gilt

$$(v_m - v, Ku_m) \to 0.$$

In der Tat: Angenommen, es wäre $|(v_m - v, Ku_m)| > \varepsilon$ für eine Teilfolge Λ. Dann gäbe es eine weitere Teilfolge $\Lambda_1 \subset \Lambda$ und ein $f \in H_2$ mit $Ku_m - f \to 0$ stark. Man hat dann den Widerspruch

$$|(v_m - v, Ku_m)| = |(v_m - v, Ku_m - f) + o(1)| \le \|v_m - v\|\,\|Ku_m - f\| + o(1) \to 0\,.$$

Mit (9.4) folgt daher die Behauptung. □

Satz 9.4 *Es sei $A\colon H_1 \to H_2$ eine beschränkte lineare Abbildung zwischen den Hilberträumen H_1 und H_2. A genüge der Koerzivitätsungleichung*

$$\|Au\| \ge c\|u\|\,, \qquad u \in H_1$$

mit einer positiven Konstanten c. Ferner sei $K\colon H_1 \to H_2$ linear und vollstetig. Dann ist der Bildraum $(A + K)(H_1)$ abgeschlossen und der Nullraum von $A + K$ endlichdimensional.

Beweis:

(i) Angenommen, es wäre

$$\dim N(A+K) = \infty\,. \tag{9.5}$$

Dann gibt es nach den Ergebnissen von Kapitel 6 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 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

$$\|A\varphi_{j_l} - A\varphi_{j_k}\| \geq c\sqrt{2}\,, \qquad l \neq k\,.$$

Die Folge $(A\varphi_{j_l})$ kann daher nicht konvergieren. Die Annahme (9.5) wurde daher zum Widerspruch geführt.

(ii) Es sei $(A+K)u_m \to f$ in H_2. Wir müssen zeigen, daß ein $u \in H_1$ existiert mit $(A+K)u = f$. Da $\dim N(A+K) < \infty$, ist der Nullraum $N(A+K)$ von $A+K$ abgeschlossen. Jedes u_m hat somit eine eindeutige Zerlegung

$$u_m = v_m + w_m, \qquad w_m \in N(A+K),\ v_m \perp N(A+K).$$

Wir werden zeigen, daß die v_m konvergieren. Nach dem im Anschluß bewiesenen Lemma gibt es eine positive Konstante c_0, so daß

$$\|(A+K)v\| \geq c_0\|v\| \qquad \text{für alle } v \in N(A+K)^\perp\,. \tag{9.6}$$

Da $(A+K)u_m = (A+K)v_m$, gilt auch

$$(A+K)v_m \to f \qquad (m \to \infty)$$

und $\big((A+K)v_m\big)$ ist eine Cauchyfolge. Wegen (9.6) ist daher auch (v_m) eine Cauchyfolge (wende (9.6) mit $v = v_m - v_k$ an). Daher folgt $f = (A+K)v$, und Satz 9.4 ist bewiesen.

□

Lemma 9.4 *Sei $A\colon H_1 \to H_2$ linear und beschränkt, $K\colon H_1 \to H_2$ linear und vollstetig, und es gelte*

$$\|Au\| \geq c\,\|u\| \qquad \textit{für alle } u \in H_1 \tag{9.7}$$

mit einer positiven Konstanten c. Dann gibt es eine positive Konstante c_0, so daß

$$\|Au + Ku\| \geq c_0\,\|u\| \qquad \textit{für alle } u \in N(A+K)^\perp\,.$$

Beweis: Angenommen, die Behauptung wäre falsch. Dann gibt es eine Folge $u_m \in N(A+K)^\perp$ mit $\|u_m\| = 1$, $\|Au_m + Ku_m\| \to 0$. Da K kompakt und (u_m) beschränkt, ist die Folge (Ku_m) relativ kompakt, somit auch die Folge Au_m. (In der Tat, jede Teilfolge $\Lambda \subset \mathbb{N}$ besitzt eine Teilfolge $\Lambda_1 \subset \Lambda$, so daß $Ku_m - Ku_k \to 0$ $(m, k \in \Lambda,\ m, k \to \infty)$.) Da $Au_m + Ku_m \to 0$, konvergiert

$$Au_m - Au_k + Ku_m - Ku_k \to 0 \qquad (m, k \in \Lambda,\ m, k \to \infty)\,,$$

somit auch $Au_m - Au_k \to 0$. Aus (9.7), angewandt für $u_m - u_k$, folgt, daß $(u_m)_{m\in\Lambda}$ für eine Teilfolge eine Cauchy-Folge ist. Somit gilt $u_m \to u$ $(m \to 0,\ m \in \Lambda)$, 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: $0 = (u_m, v) \to (u, v)$.) Damit gilt $(A+K)u \neq 0$, da $N(A+K)^\perp \cap N(A+K) = \{0\}$, und es ergibt sich der Widerspruch

$$(A+K)u_m \to 0 \quad \text{und} \quad (A+K)u \neq 0\,.$$

□

Die Endlichdimensionalität des Defektraumes $\big((A+K)(H_1)\big)^\perp$ erhält man analog zu Satz 9.4 durch eine Voraussetzung an A^*.

Satz 9.5 *Es sei $A\colon H_1 \to H_2$ eine beschränkte lineare Abbildung zwischen den Hilberträumen H_1, H_2. Die adjungierte Abbildung A^* erfülle*

$$\|A^* v\| \geq c\,\|v\|\,, \quad v \in H_1$$

mit einer positiven Konstanten c. Ferner sei $K\colon H_1 \to H_2$ linear und vollstetig. Dann ist

$$\dim\big((A+K)H_1\big)^\perp < \infty.$$

Beweis: Nach Satz 9.4 und Satz 9.3 ist $\dim N(A^* + K^*) < \infty$. Sei nun $v \in \big((A+K)H_1\big)^\perp$, d.h. $\big(v, (A+K)u\big) = 0$ für alle $u \in H_1$. Daraus folgt $\big((A+K)H_1\big)^\perp \subset N(A^* + K^*)$. Da $\dim N(A^* + K^*) < \infty$ und $\big((A+K)H_1\big)^\perp$ linear ist, folgt die Abgeschlossenheit von $\big((A+K)H_1\big)^\perp$ und die Endlichkeit der Dimension von $\big((A+K)H_1\big)^\perp$. □

Ist $H_1 = H_2 = H$, so folgen die Bedingungen $\|Au\| \geq c\,\|u\|$, $\|A^*u\| \geq c\,\|u\|$ beide aus der Bedingung

$$(Au, u) \geq c\,\|u\|^2\,. \tag{9.8}$$

Im Folgenden bringen wir eine bedeutende Anwendung der Ergebnisse von Kapitel 5 und 8 auf Randwertprobleme elliptischer partieller Differentialgleichungen und auch gewöhnlicher Differentialgleichungen. Entscheidend ist, daß man für Randwertprobleme

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

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:

Satz 9.6 *Sei Ω eine beschränkte offene Punktmenge $\subset \mathbb{R}^n$ 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 sei definiert durch

$$Q(u,v) = \int_\Omega \Big[\sum_{i,k=1}^{n} a_{ik}\, D_k u\, D_i v + \sum_{i=1}^{n} b_i\, D_i u\, v + c\, u \cdot v \Big]\, dx\,.$$

Dann gilt die Gårdingsche Ungleichung

$$Q(u,u) \geq c_0\, \|u\|^2_{H^{1,2}} - \lambda_0\, \|u\|^2_{L^2} \text{ für alle } u \in H^{1,2}(\Omega)$$

mit einer positiven Konstanten $c_0 > 0$ sowie einer Konstanten λ_0.

Für gewöhnliche Differentialgleichungen lassen wir eine noch etwas allgemeinere Bilinearform zu:

Satz 9.7 *Sei $I = [a,b]$, $p \in L^\infty(I)$, $p \geq c_0 = \text{const} > 0$, $r \in L^\infty(I)$, $q \in L^\infty(I)$, $\alpha, \beta \in \mathbb{R}$ 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

$$Q(u,u) \geq c_0\, \|u\|^2_{H^{1,2}} - \lambda_0\, \|u\|^2_{L^2} \text{ für alle } u \in H^{1,2}(I),$$

mit einer positiven Konstanten $c_0 > 0$ und einer Konstanten λ_0.

Der Beweis von Satz 9.6 ist sehr einfach:

Beweis von Satz 9.6 Wegen der Elliptizitätsbedingung schätzt man

$$\int_\Omega \sum_{i,k=1}^n a_{ik} D_k u\, D_i u\, dx \geq \alpha \int_\Omega |\nabla u|^2\, dx = \alpha\, \|u\|^2_{H^{1,2}} - \alpha\, \|u\|^2_{L^2}$$

ab. Den zweiten Summand schätzt man nach unten ab,

$$\int_\Omega \sum_{i=1}^n b_i\, D_i u\, u\, dx \geq -\frac{\alpha}{2} \int_\Omega |\nabla u|^2\, dx - \frac{\|b\|^2_\infty}{2\alpha} \int_\Omega |u|^2\, dx\,,$$

woraus sich

$$\begin{aligned} Q(u,u) &\geq \alpha \int_\Omega |\nabla u|^2\, dx - \frac{\alpha}{2} \int_\Omega |\nabla u|^2\, dx - \frac{\|b\|^2_\infty}{2\alpha} \int_\Omega |u|^2\, dx - \|c\|_\infty \int_\Omega |u|^2\, dx = \\ &= \frac{\alpha}{2} \Big(\int_\Omega |\nabla u|^2\, dx + \int_\Omega |u|^2\, dx \Big) - \Big(\frac{\|b\|^2_\infty}{2\alpha} + \frac{\alpha}{2} + \|c\|_\infty \Big) \int_\Omega |u|^2\, dx \end{aligned}$$

ergibt. □

Der Beweis von Satz 9.7 verläuft analog, nur beachtet man, daß

$$u^2(b) - u^2(\xi) = \int_\xi^b (u^2)'\, dx \leq \varepsilon \int_a^b |u'|^2\, dx + K_\varepsilon \int_a^b |u|^2\, dx$$

ist. Um die Gårdingsche Ungleichung zu verwerten, setzen wir – nach dem Rieszschen Darstellungssatz –

$$(Au, v)_{H^{1,2}} = Q(u, v)$$

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

$$(Au, u)_H \geq c\, \|u\|^2_H - \lambda_0\, \|u\|^2_{L^2}\,. \tag{9.9}$$

Es sei J die - nach den vorangegangenen Überlegungen – *vollstetige* lineare Abbildung von H nach H, für die gilt:

$$(Ju, v)_H = (u, v)_{L^2}\,.$$

Aus (9.9) folgt

$$\big((A + \lambda_0 J)u, u\big)_H \geq c\, \|u\|^2_H\,,$$

und wir können Satz 9.4, 9.5 und die Bemerkung im Zusammenhang mit (9.8) 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.*

Satz 9.8 *Unter den Voraussetzungen von Satz 9.6 oder Satz 9.7 ist das Randwertproblem*
„gesucht ist $u \in H_0^{1,2}$ bzw. $H^{1,2}$, so daß

$$Q(u,v) = (f,v)_{L^2} \qquad \text{für alle } v \in H_0^{1,2} \text{ bzw. } H^{1,2} \text{ “}$$

lösbar genau dann, wenn

$$(f,v)_{L^2} = 0 \qquad \text{für alle } v \in N(A^*)\,,$$

$$N(A^*) = \left\{v \in H_0^{1,2} \text{ bzw. } H^{1,2} \,\middle|\, Q(u,v) = 0 \text{ für alle } u \in H_0^{1,2}\right\}.$$

Es gilt $\dim N(A^*) < \infty$ *und* $\dim N(A) < \infty$ *mit*

$$N(A) = \left\{u \in H_0^{1,2} \text{ bzw. } H^{1,2} \,\middle|\, Q(u,v) = 0 \text{ für alle } v \in H_0^{1,2} \text{ bzw. } H^{1,2}\right\}.$$

Im Fall, daß $H^{1,2}$ der gewählte Grundraum ist, muß zusätzlich vorausgesetzt werden, daß der Einbettungsoperator J vollstetig ist.

Erläuterung: Im Fall der Grundräume $H_0^{1,2}(I)$ und $H^{1,2}(I)$ ist J immer kompakt, im Fall $H^{1,2}(\Omega)$ dagegen nur, wenn $\partial\Omega$ genügend regulär ist. Wir haben daher diese benötigte Voraussetzung in die letzten beiden Zeilen von Satz 9.8 aufgenommen.

10 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 Nullraumes *endlich–dimensional* sind. Ferner werden die Eigenvektoren über ein Maximum–Minimum–Prinzip erhalten. Wir werden diese Aussage im Folgenden beweisen, zum Teil unter schwächeren Voraussetzungen.

Der Vollständigkeit halber erinnern wir an die

Definition 10.1 *Sei $A\colon H \to H$. Ein Element $u \in H$, $u \neq 0$, heißt* Eigenvektor *oder* Eigenelement *von H zum* Eigenwert $\lambda \in \mathbb{R}$ *oder $\mathbb{C}$, falls*

$$Au = \lambda u\,.$$

Allgemeiner läßt sich auch die Situation $A\colon D(A) \to H$, $D(A)$ dicht in H, betrachten.

Lemma 10.1 *Sei H ein Hilbertscher Raum über $\mathbb{C}$ und $D(A)$ dichter linearer Teilraum von H. Die lineare Abbildung $A\colon D(A) \to H$ 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 $(u, Au) = \overline{\lambda}\,(u,u)$, und, da A selbstadjungiert ist, $A = A^*$ und $u \in D(A^*)$, d.h.

$$\overline{\lambda}\,(u,u) = (u, Au) = (A^*u, u) = (Au, u) = \lambda\,(u,u).$$

Daraus folgt

$$\lambda = \overline{\lambda}\,.$$

Gilt $Au = \lambda u,\ Av = \mu v$, mit $\lambda \neq \mu,\ u, v \neq 0$, so folgt

$$\lambda\,(u,v) = (Au, v) = (u, Av) = (u, \mu v) = \overline{\mu}\,(u,v) = \mu\,(u,v),$$

woraus $(u,v) = 0$ folgt. $\square$

Lemma 10.2 *Sei H ein Hilbertscher Raum und $K\colon H \to H$ vollstetig und linear. Sei $\lambda \in \mathbb{C}$ ein beliebiger, von Null verschiedener Eigenwert. Dann ist der Eigenraum V zu λ*

$$V := \{u \in H | \, Ku = \lambda u\}$$

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 $\{\varphi_j | \, j = 1, \ldots, \infty\}$. Da K vollstetig ist, K kompakt und die Menge der φ_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 φ_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

Lemma 10.3 *Es sei H ein Hilbertscher Raum und $K\colon H \to H$ 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 $v_j \in H$, $\|v_j\| = 1$, $\lambda_j \neq \lambda_k$ für $j \neq k$, $\lambda_j \in \mathbb{R}$, $\lambda_j \to \lambda \neq 0$ $(j \to \infty)$. Da (v_j) beschränkt ist, ist die Folge Kv_j relativ kompakt, somit auch die Folge $\lambda_j^{-1}Kv_j$ (hierbei wurde $\lambda \neq 0$ benutzt). 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 10.1. Die v_j müssen daher ein abzählbar unendliches Orthonormalsystem bilden, welches nicht relativ kompakt sein kann – siehe Beweis von Lemma 10.2 □

Folgerung aus Lemma 10.3: Der Nullraum 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.

Satz 10.1 *Es sei H ein Hilbertscher Raum und $K\colon H \to H$ kompakt, linear und selbstadjungiert. Dann gibt es ein $u \in H$ mit $\|u\| = 1$ und*

$$|(Ku, u)| = \sup \big\{ |(Kv, v)| \,\big|\, \|v\| = 1 \big\} =: s\,. \tag{10.1}$$

Dieses Element u ist Eigenvektor zum Eigenwert $\lambda = (Ku, u)$, und λ ist der betragsmäßig größte Eigenwert.

Beweis:

(i) Wir zeigen zunächst, daß das Supremum bei (10.1) durch einen Vektor u angenommen wird. Sei (u_m) eine *Maximalfolge*, d.h. $\|u_m\| = 1$ und

$$|(Ku_m, u_m)| \to s \qquad (m \to \infty)\,.$$

Wegen der schwachen Kompaktheit beschränkter Mengen im Hilbertraum und der vorausgesetzten Kompaktheit von K dürfen wir nach Umnumerierung annehmen, daß

$$u_m \rightharpoonup u \quad \text{schwach} \qquad \text{und} \qquad Ku_m \to Ku \quad \text{stark.}$$

Daraus folgt $(Ku_m, u_m) \to (Ku, u)$, denn

$$\begin{aligned} |(Ku_m, u_m) - (Ku, u)| &\le |(Ku_m - Ku, u_m)| + |(Ku, u_m - u)| \\ &\le \|Ku_m - Ku\| + o(1) = o(1)\,. \end{aligned}$$

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, $\|u_m\| = 1$ können wir $\|u\| \le 1$ schließen:

$$\|u\|^2 = o(1) + (u_m, u) \le o(1) + \|u\|\,,$$

also $\|u\|^2 \le \|u\|$ und $\|u\| \le 1$.
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 wäre für $\tilde{u} := \theta^{-1}\, u \quad \|\tilde{u}\| = 1$, 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 (10.1) von u ist gezeigt.

(ii) Wir zeigen, daß Maximal– bzw. Minimalstellen u von (Ku, u), $\|u\| = 1$, Eigenvektoren sind. Sei $v(t) = u\cos t + v\sin t,\ v \in H,\ \|v\| = 1,\ v\perp u$. Wegen der Maximaleigenschaft (o.B.d.A.) von u gilt $(Ku, u) \geq (Kv(t), v(t))$, denn $\|v(t)\| = 1$, und somit

$$\frac{d}{dt}\big(Kv(t), v(t)\big)|_{t=0} = 0\,.$$

Andererseits ist (da K selbstadjungiert)

$$\frac{d}{dt}\big(Kv(t), v(t)\big) = 2\,\big(K\dot{v}(t), v(t)\big) = 2\,(-Ku\sin t + Kv\cos t,\ u\cos t + v\sin t)\,.$$

Damit gilt

$$0 = \frac{d}{dt}\big(Kv(t), v(t)\big)|_{t=0} = 2\,(Kv, u) = 2\,(v, Ku) \quad \text{für alle } v\perp u\,. \tag{10.2}$$

Wegen des Projektionssatzes ist $Ku = \lambda u + v_0,\ v_0\perp u$, und wegen (10.2)

$$0 = (v_0, Ku) = \|v_0\|^2,$$

d.h. $v_0 = 0$ und $Ku = \lambda u$. Ist schließlich $\mu \in \mathbb{R}$ ein weiterer Eigenwert von K zum normierten Eigenvektor u_0, so gilt

$$|\mu| = |(\mu u_0, u_0)| = |(u_0, Ku_0)| \leq \sup_{\|u\|=1} |(u, Ku)|\,.$$

□

Hilfssatz 10.1 *Sei H Hilbertscher Raum und $K\colon H \to H$ 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

$$0 = (Ku, u) + (Kv, v) + 2\,(Ku, v) = 2\,(Ku, v) \quad \text{für alle } v \in H\,.$$

Daraus folgt $Ku = 0$. □

Wir konstruieren uns nun weitere, betragsmäßig monoton fallende Eigenwerte von K mit Hilfe der folgenden Konstruktion:
Sei λ_1 der nach Satz 10.1 existierende betragsmäßig größte Eigenwert. Es gilt

$$\begin{aligned} Ku &= \lambda_1 u \quad \text{für } u \in V_1^+ , \\ Ku &= -\lambda_1 u \quad \text{für } u \in V_1^- . \end{aligned}$$

Die Eigenräume V_1^+, V_1^- sind nach Lemma 10.2 endlich dimensional, nach Lemma 10.1 ist $V_1^+ \perp V_1^-$ und $\dim V_1^+ \oplus V_1^- \geq 1$ nach Satz 10.1. Zur Konstruktion des betragsmäßig *zweitgrößten* Eigenwertes lösen wir die *Optimierungsaufgabe*:

$$\text{Gesucht ist } u \in H \text{ mit } \|u\| = 1 \text{ und } u \perp W_1,\ W_1 := V_1^+ \oplus V_1^- ,\ \text{so daß } |(Ku, u)| = \max . \tag{10.3}$$

(Im Gegensatz zur Konstruktion des betragsmäßig größten Eigenwertes besteht in der Optimierungsaufgabe noch die *zusätzliche Nebenbedingung* $u \perp W_1$.)

Lemma 10.4 *Sei $K\colon H \to H$ linear, selbstadjungiert und kompakt und $W_1 \neq H$. Dann gibt es eine Lösung u der Aufgabe (10.3).*

Beweis: Wegen Hilfssatz 10.1 nehmen wir o.B.d.A. an, daß

$$\sigma := \sup \left\{ (Ku, u) \middle| \, \|u\| = 1 \, , \; u \perp W_1 \right\} > 0 \, .$$

Andernfalls ist $W_1^\perp$ Eigenraum zum Eigenwert Null. Sei (u_m) eine Maximalfolge, d.h. $\|u_m\| = 1$, $u_m \perp W_1$, $(Ku_m, u_m) \to \sigma$. Wegen der schwachen Kompaktheit beschränkter Mengen im Hilbertraum gibt es eine Teilfolge $(u_m)_{m \in \Lambda}$, so daß

$$u_m \rightharpoonup u \quad \text{schwach in } H, \quad (m \to \infty,\ m \in \Lambda) \, .$$

Wegen der Kompaktheit von K gilt $Ku_m \to Ku$ und $(Ku_m, u_m) \to (Ku, u)$. Ferner gilt $u \perp W_1$ und

$$\|u\|^2 = (u, u_m) + o(1) \leq \|u\| \, \|u_m\| + o(1) \leq \|u\| + o(1),$$

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 σ

$$\left(K \tfrac{u}{\|u\|}, \tfrac{u}{\|u\|} \right) \leq \sigma \, ,$$

somit $\sigma \leq \sigma \|u\|^2$, woraus $\|u\| \geq 1$ folgt. Der obige schwache Limes u ist somit Lösung von (10.3). □

Lemma 10.5 *Jede Lösung $u \in H$ von (10.3), $H \neq W_1$, ist Eigenvektor zum Eigenwert $\lambda_2 = (Ku, u)$. Es gilt $|\lambda_1| > |\lambda_2|$, und es gibt keine weiteren Eigenwerte λ von K mit $|\lambda_1| > |\lambda| > |\lambda_2|$.*

Beweis: Sei u Lösung von (10.3). 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"

$$v(t) = u \cos t + v \sin t\,.$$

Sei o.B.d.A.

$$\sigma = \sup\big\{(Kz, z)\big|\|z\| = 1,\ z \perp W_1\big\} > 0.$$

($\sigma = 0$ ist nicht möglich wegen Hilfssatz 10.1.) Wegen der Maximaleigenschaft von u gilt

$$(Ku, u) \geq \big(Kv(t),\ v(t)\big)\,,$$

und wie beim Beweis von Lemma 10.1 schließen wir wieder

$$(v, Ku) = 0 \qquad \text{für alle } v \perp u,\ v \perp W_1\,. \tag{10.4}$$

(Zweimalige) Anwendung des Projektionssatzes ergibt

$$Ku = \lambda_2 u + w + v,\ w \in W_1,\ v \perp u,\ v \perp W_1\,. \tag{10.5}$$

Hieraus und aus (10.4) erhalten wir durch Bildung des Skalarproduktes mit v, daß $\|v\|^2 = 0$ und somit $v = 0$ ist. Skalare Multiplikation von (10.5) 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

$$Ku = \lambda_2 u\,.$$

u ist also Eigenvektor.
Wir zeigen noch, daß zwischen $|\lambda_1|$ und $|\lambda_2|$ kein Betrag eines anderen Eigenwertes liegen kann:
Sei $|\lambda_1| > |\lambda| > |\lambda_2|$ und λ Eigenwert zum Eigenraum V. Da Eigenräume selbstadjungierter Operatoren zu verschiedenen Eigenwerten zueiander orthogonal sind, gilt $V \perp W_1$ und für $v \in V,\ \|v\| = 1$

$$|\lambda| = |(Kv, v)| \leq \sup\big\{\,|(Kz, z)|\,\big|\,z \perp W_1,\ \|z\| = 1\big\} = |\lambda_2|\,.$$

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 $j = 1, \ldots, N-1$. Es gelte

$$|\lambda_1| > |\lambda_2| > \ldots > |\lambda_{N-1}| > 0\,,$$
$$|\lambda_1| = \sup\{|(Ku,u)|\; \|u\| = 1,\; u \in H\},$$

und es existiere kein Eigenwert mit $|\lambda_{j-1}| > |\lambda| > |\lambda_j|$, $j = 1, \ldots, N-1$. Wir konstruieren V_N^+ und V_N^- durch die folgende Optimierungsaufgabe:

Gesucht ist $u \in H,\; \|u\| = 1,\; u \perp W_j = V_j^+ \oplus V_j^-,\; j = 1, \ldots, N-1$, so daß

$$|(Ku,u)| = \sigma_N := \sup\big\{|(Kz,z)|\;\big|\; \|z\| = 1,\; z \perp W_j,\; j = 1, \ldots, N-1\big\}\,. \tag{10.6}$$

Lemma 10.6 *Sei $H \neq \bigoplus_{j=1}^{N-1} W_j$. Unter den Voraussetzungen von Lemma 10.4 besitzt die Aufgabe (10.6) eine Lösung u.*

Der Beweis geschieht ähnlich wie in Lemma 10.4 und wird daher hier nicht abgedruckt.

Lemma 10.7 *Jede Lösung von (10.6) ist Eigenvektor zum Eigenwert λ_N. Es gilt $|\lambda_{N-1}| > |\lambda_N|$, und es gibt keinen Eigenwert λ von K mit $|\lambda_{N-1}| > |\lambda| > |\lambda_N|$.*

Beweis: Das Element Ku von (10.6) hat nach dem Projektionssatz die Darstellung

$$Ku = \pm\lambda_N u + \sum_{j=1}^{N-1} w_j + v\,, \quad w_j \in W_j \tag{10.7}$$

mit $v \perp w_j$, $j = 1, \ldots, N-1$, $v \perp u$. Die Orthogonalität $u \perp W_j$, $j = 1, \ldots, N-1$ gilt nach Konstruktion ohnehin. Wir zeigen, daß die Elemente w_j und v in (10.7) 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

$$0 = (v,Ku) = \|v\|^2\,,$$

wobei die Beziehung $(v,Ku) = 0$ aus der Beziehung

$$\frac{d}{dt}\big(v(t),Kv(t)\big)|_{t=0} = 0\,, \qquad v(t) := u\cos t + \frac{v}{\|v\|}\sin t$$

folgt. Ist schließlich λ 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 10.5 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 $|\lambda_{j-1}| > |\lambda_j|$ erhalten. Ferner ist noch zusätzlich der Eigenwert Null möglich. Es können nicht über abzählbar viele Eigenwerte vorliegen, und die Eigenwerte λ_j müssen wegen Lemma 10.3 gegen Null gehen.

Die Vollständigkeit des Systems der Eigenfunktionen

Satz 10.2 *Sei K eine lineare, vollstetige, selbstadjungierte Abbildung eines Hilbertraumes H in sich und V_j^+, V_j^- die in Lemma 10.6 konstruierten Eigenräume von K zum Eigenwerte $\pm\lambda_j$, $j \in \mathbb{N}$. Dann gilt*

$$H = N(K) \oplus H_0 \,, \qquad H_0 := \bigoplus_{j=1}^{\infty} \left(V_j^+ \oplus V_j^-\right) .$$

Erläuterung: H_0 besteht aus allen konvergenten Reihen $\sum\limits_{j=1}^{\infty}(\mu_j^+ v_j^+ + \mu_j^- v_j^-)$, $\quad v_j^{\pm} \in V_j^{\pm}$, $\mu_j^{\pm} \in \mathbb{R}$ oder $\mathbb{C}$. 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

$$|(Ku,u)| = \max! \,, \quad \|u\| = 1 \,, \quad u \perp N(K) \oplus H_0 \,. \tag{10.8}$$

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

$$w(t) = \cos t \; u + \sin t \; w \,, \ \|w\| = 1 \,, \ w \perp u \,, \ w \perp N(K) \oplus H_0 \,.$$

(Falls es ein solches w nicht gibt, gilt $Ku = \lambda u + v$, $v \perp N(K) \oplus H_0$ und man springe zu (10.11) mit $w = 0$.)
Es gilt

$$\frac{d}{dt}\big(Kw(t), w(t)\big)\Big|_{t=0} = 0 \,, \tag{10.9}$$

da $w(t)$ zur Konkurrenz in (10.8) zugelassen ist.
Aus (10.9) folgt

$$(Ku, w) = 0 \quad \text{für } w \perp u,\ w \perp N(K) \oplus H_0 \,. \tag{10.10}$$

Allgemein gilt wegen des Projektionssatzes

$$Ku = \lambda u + w + v \,, \qquad w \perp u \,,\ v \perp u \,,\ v \in N(K) \oplus H_0 \,,\ w \perp v \,. \tag{10.11}$$

Durch skalare Multiplikation mit w ergibt sich

$$\|w\|^2 = (Ku, w) = 0 \,, \quad \text{also } Ku = \lambda u + v \,,$$

und durch skalare Multiplikation mit v

$$\|v\|^2 = (Ku, v) = (u, Kv)\,.$$

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 $\sum\limits_{j=1}^{\infty} \big(|\zeta_j^+|^2 + |\zeta_j^-|^2\big) < \infty$. Da K insbesondere stetig, folgt $Kz = \sum\limits_{j=1}^{\infty} \big(\zeta_j^+ \lambda_j^+ v_j^+ + \zeta_j^- \lambda_j^- v_j^-\big)$. Beachte, daß $|\lambda_j^\pm| \leq \|K\|$. Die letzte unendliche Summe konvergiert somit in H, und damit gilt $Kz \in H_0$.)

Aus (10.11) folgt damit die Eigenwertgleichung

$$Ku = \lambda u\,, \qquad u \neq 0\,.$$

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 $|\lambda_j| \to 0$. Dies ist aufgrund der Konstruktion des λ_j nicht möglich. □

Anwendungen auf Rand–Eigenwertprobleme für elliptische Differentialoperatoren der Gestalt

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

Sei Ω eine beschränkte offene Punktmenge des $\mathbb{R}^n$, $a_{ik} \in L^\infty(\Omega)$, $a_{ik} = a_{ki}$, $c \in L^\infty(\Omega)$. Es gelte die Elliptizitätsbedingung

$$\sum_{i,k=1}^{n} a_{ik}(x)\, \xi_i\, \xi_k \geq \alpha_0 |\xi|^2\,, \qquad \xi \in \mathbb{R}^n \tag{10.12}$$

mit einer Konstanten α_0.

Wir betrachten das Eigenwertproblem

$$\begin{aligned} -\sum_{i=1}^{n} D_i(a_{ik} D_k u) + c\,u = \lambda u\,, \qquad u \neq 0 \\ u|_{\partial\Omega} = 0 \end{aligned} \tag{10.13}$$

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 brauchte 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 Ω beschränkt ist.)

Wir führen noch die mit der Einbettung von $H^{1,2}$ nach L^2 zusammenhängende Abbildung $J\colon L^2 \to H^{1,2}$ ein:

$$(Ju, v)_{H^{1,2}} := (u, v)_{L^2}$$

Die Restriktion von L^2 auf $H^{1,2}$ wird ebenfalls mit J bezeichnet. Mit diesen Bezeichnungen führt die schwache Formulierung des Eigenwertproblems (10.13) auf die Gleichung

$$Au = \lambda J u\,. \tag{10.14}$$

Hierbei ist $A\colon H_0^{1,2} \to H_0^{1,2}$ eine lineare Abbildung, welche der Gårdingschen Ungleichung (Satz 9.6)

$$(Au, u)_{H^{1,2}} \geq \alpha \|u\|^2_{H^{1,2}} - \lambda_0 \|u\|^2_{L^2}$$

mit Konstanten $\alpha, \lambda_0 > 0$ genügt. Dies schreiben wir in der Form

$$\big((A + \lambda_0 J)u, u\big)_{H^{1,2}} \geq \alpha \|u\|^2_{H^{1,2}}\,. \tag{10.15}$$

Das Eigenwertproblem (10.14) formulieren wir äquivalent um:

$$(A + \lambda_0 J)u = (\lambda + \lambda_0) J u\,,$$

Wir setzen

$$B = A + \lambda_0 J\,, \quad \mu = \lambda + \lambda_0\,.$$

Damit lautet das Eigenwertproblem

$$Bu = \mu J u\,. \tag{10.16}$$

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 $B^{\frac{1}{2}} = C$ wird in einer Übungsaufgabe gezeigt. Im $\mathbb{R}^n$ läßt sich die Existenz der Wurzel einer positiven semidefiniten hermitischen Matrix über die Transformation auf Diagonalform beweisen:
$B = \overline{S}^T DS$, D Diagonalmatrix mit nichtnegativem Diagonalelement, $\overline{S}^T S = I$. Man hat dann $B^{\frac{1}{2}} = \overline{S}^T \sqrt{D} S$, $\sqrt{D} :=$ Diagonalmatrix mit den Elementen $\sqrt{\lambda_j}$.

(10.16) formen wir um mit

$$v = B^{\frac{1}{2}} u \, , \quad B^{\frac{1}{2}} v = \mu J B^{-\frac{1}{2}} v \, , \text{ in}$$

$$\frac{1}{\mu} v = B^{-\frac{1}{2}} J B^{-\frac{1}{2}} v \, . \tag{10.17}$$

Die Abbildung $B^{-\frac{1}{2}}$ ist linear, beschränkt und selbstadjungiert, die Abbildung J ebenfalls; überdies ist J kompakt. Damit ist die Abbildung

$$K = B^{-\frac{1}{2}} J B^{-\frac{1}{2}}$$

linear, beschränkt und selbstadjungiert, und wir können die in Kapitel 9 entwickelte Theorie anwenden. Zur Vereinfachung ist es nützlich zu wissen, daß die Nullräume $N(K)$ und $N(J)$ verschwinden.

Lemma 10.8

$$N(K) = 0 \, , \quad N(J) = 0 \, .$$

Beweis: Es ist $(Ju, v)_{H^{1,2}} = (u, v)_{L^2}$. Gilt also $u \in N(J)$, so folgt

$$(u, v)_{L^2} = 0 \qquad \text{für alle } v \in H_0^{1,2}(\Omega) \, .$$

Da $B^{-\frac{1}{2}}$ invertierbar, folgt auch $N(K) = 0$. □

Wir betrachten somit das Eigenwertproblem

$$Ku = \tfrac{1}{\mu} u \quad \text{ in } H^{1,2} \, ,$$

$N(K) = 0$, $K = K^*$, K kompakt. Aus Lemma 10.8 sowie Satz 10.1 und Satz 10.2 erhalten wir

Satz 10.3 *Das elliptische Rand–Eigenwertproblem (10.13) besitzt – unter den an a_{ik} und c genannten Voraussetzungen – abzählbar–unendlich viele Eigenwerte ν_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 $H^{1,2}(\Omega)$.*

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.

Beweis: Die Zahl ν ist genau dann Eigenwert des Problems (10.13), wenn die Zahl $\frac{1}{2+\lambda_0}$ Eigenwert von $K = B^{-\frac{1}{2}} J B^{\frac{1}{2}}$ ist. Da die Eigenwerte von K gegen Null gehen, d.h. $\frac{1}{\nu_j+\lambda_0} \to 0$, folgt $|\nu_j| \to \infty$, und da $(Ku, u) \geq 0$, gilt $\nu_j \to \infty$ $(j \to \infty)$ für die Eigenwerte ν_j von (10.13). $\square$

Teil II

Allgemeine lineare Funktionalanalysis

1 Grundbegriffe der Funktionalanalysis

Metrische Räume

Definition 1.1 *Ein* metrischer Raum *ist ein Vektorraum V über $\mathbb{R}$ oder $\mathbb{C}$, auf dem eine* Abstandsfunktion $\rho\colon V \times V \to \mathbb{R}$ *gegeben ist, die die folgenden Eigenschaften für alle $x, y \in V$ 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 ρ metrisiert“ und nennt ρ eine *Metrik*.

Im metrischen Raum läßt sich ein Konvergenzbegriff einführen.

Definition 1.2 *Eine Folge $(u_m)_{m\in\mathbb{N}}$, $u_m \in V$* konvergiert *genau dann* gegen *ein Element u, wenn*

$$\rho(u_m, u) \to 0 \qquad (m \to \infty)\,.$$

Man schreibt $u_m \to u$ in V oder bezüglich ρ $(m \to \infty)$. 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*

$$\rho(u_m, u_k) < \varepsilon \qquad \text{für } m, k \geq n_0\,,$$

so gibt es ein $u \in V$ mit $u_m \to u$ $(m \to \infty)$.

Beispiele metrischer Räume:

a) $\mathbb{R}^n$ mit $\rho(x,y) = |x-y|$ - kennt jeder!

b) Der Folgenraum $l = \big\{ (c_1, c_2, \dots ,) \,\big|\, c_j \in \mathbb{R} \text{ oder } \mathbb{C} \big\}$ mit der Metrik

$$\rho(a,b) = \sum_{j=1}^{\infty} \frac{1}{2^j} \frac{|a_j - b_j|}{1 + |a_j - b_j|} \tag{1.1}$$

Wir zeigen, daß in der Tat ρ die Eigenschaften einer Metrik erfüllt. Die positive Definitheit und die Symmetrie sind offensichtlich. Die Dreiecksungleichung ergibt sich folgendermaßen:

$$\rho(a,b) = \sum_{j=1}^{\infty} \frac{1}{2^j}\left(1 - \frac{1}{1 + |a_j - b_j|}\right) \le \sum_{j=1}^{\infty} \frac{1}{2^j}\left(1 - \frac{1}{1 + |a_j - c_j| + |c_j - b_j|}\right)$$
$$= \sum_{j=1}^{\infty} \frac{1}{2^j}\left(\frac{|a_j - c_j| + |c_j - b_j|}{1 + |a_j - c_j| + |c_j - b_j|}\right) \le \rho(a,c) + \rho(c,b)\,.$$

Man überlegt sich leicht:

Lemma 1.1 *$c^{(k)} \to c$, $c^{(k)}$, $c \in l$, konvergiert genau dann bezüglich der Metrik ρ in (1.1), wenn die Komponenten $c_j^{(k)}$ von $c^{(k)}$ gegen die Komponenten c_j von c konvergieren.*

a) *Der Raum der meßbaren Funktionen*

Sei $\tilde{F} = \{f : \Omega \to \mathbb{R} \text{ oder } \mathbb{C} \,|\, f \text{ meßbar } \}$ mit einer beschränkten meßbaren Menge Ω, und

$$N = \{f \in \tilde{F} \,|\, f = 0 \text{ fast überall}\}\,.$$

Wir betrachten den Restklassenraum

$$F = \tilde{F}/N\,.$$

Etwas unpräzise kann man sagen, daß F aus allen Funktionen besteht, die bis auf einer Menge vom Maß Null definiert sind und zwei Funktionen als gleich angesehen werden, wenn sie sich nur auf einer Menge vom Maß Null unterscheiden.
Als Metrik in F definieren wir

$$\rho(f,g) = \int_{\Omega} \frac{|f-g|}{1+|f-g|}\,dx\,.$$

Man prüft leicht nach, daß ρ eine Metrik in F definiert. Die Dreiecksungleichung beweist man wie in Beispiel (ii).

Satz 1.1 *Die Räume $\mathbb{R}^n, l, F$ sind bezüglich der angegebenen Metriken vollständig.*

Beweis: Für den $\mathbb{R}^n$ weiß dies jeder. Für den Raum l schließt man aus $\rho\big(a^{(m)}, a^{(k)}\big) \to 0$ $(m, k \to \infty)$, daß die Komponenten $a_j^{(m)}$ Cauchyfolgen in $\mathbb{R}$ oder $\mathbb{C}$ bilden. Es gibt daher ein $a \in l$, so daß $a^{(m)} \to a$ *komponentenweise*. Daraus folgt

$$\rho\big(a^{(m)}, a\big) \leq \sum_{j=1}^{N} \frac{1}{2^j} \frac{\big|a_j^{(m)} - a_j\big|}{1 + \big|a_j^{(m)} - a_j\big|} + \sum_{j=N+1}^{\infty} \frac{1}{2^j} = o(1) + \frac{1}{2^N} \to 0 \text{ für } N \to \infty\,,$$

was zu zeigen war.

Für den Raum F betrachten wir eine Folge f_j meßbarer Funktionen (genau genommen, Klassen $[f_j]$) mit $\int\limits_\Omega \frac{|f_j - f_k|}{1+|f_j - f_k|}\,dx \to 0$.

Für eine Teilfolge schließt man dann $f_j - f_k \to 0$ fast überall, somit $f_j \to f$ fast überall mit einer meßbaren Funktion f. Hierbei werden Standardsätze der Maßtheorie benutzt, z.B. daß der punktweise Limes meßbarer Funktionen meßbar ist.

Für $j, k \geq n_0$ gilt

$$\rho(f_j, f_k) = \int\limits_\Omega \frac{|f_j - f_k|}{1 + |f_j - f_k|}\,dx < \frac{\varepsilon}{2}\,.$$

Wir lassen den Index k die konstruierte Teilfolge durchlaufen, $k \to \infty$, und gehen nach dem Satz von Lebesgue 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 Ω mit der obigen Metrik des Raumes der meßbaren Funktionen. Offensichtlich führen Cauchyfolgen aus diesem Raum heraus.

In einem metrischen Raum lassen sich in Analogie zum $\mathbb{R}^n$ die Begriffe „beschränkte Menge“, „offene Menge“, „abgeschlossene Menge“, „Häufungspunkt einer Folge oder Menge“ definieren. Als ε–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 (Beweis siehe Kantorowitsch/Akilov, Funktionalanalysis in normierten Räumen, Kap. I,3.).

Ist der Raum unendlich–dimensional, so sind jedoch beschränkte abgeschlossene Mengen nicht mehr folgenkompakt.

Normierte Räume und Banachräume

Definition 1.3 *Ein* normierter Raum *ist ein Vektorraum über* $\mathbb{R}$ *oder* $\mathbb{C}$ *mit einer Abbildung* $\|.\| : V \to \mathbb{R}$, *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: $\|\lambda u\| = |\lambda| \, \|u\|$, $\lambda \in \mathbb{R}$ *oder* $\mathbb{C}$.

(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

Lemma *Sei* V *ein metrischer Raum mit einer translations– 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

$$u_m \to u \text{ in } V, \; m \to \infty \Leftrightarrow \|u_m - u\| \to 0 \, .$$

Definition 1.4 *Ein* Banachraum *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 $\|a\| = \Big(\sum\limits_{j=1}^{n} |a_j|^2 \Big)^{\frac{1}{2}}$ als Norm (dies weiß jeder);

b) Der Folgenraum $l^p = \{a = (a_1, a_2, a_3, \dots,) \, | \, a_j \in \mathbb{R} \text{ oder } \mathbb{C}, \; \|a\|_p < \infty\}, 1 \leq p < \infty$ mit der Norm

$$\|a\|_p = \Big(\sum_{j=1}^{\infty} |a_j|^p \Big)^{\frac{1}{p}} \, .$$

Die Dreiecksungleichung – zunächst für endliche Summen – folgt zum Beispiel aus der Darstellung

$$\|a\|_p = \sup\left\{ \sum_{j=1}^{N} a_j \overline{b}_j \,\middle|\, \sum_{j=1}^{N} |b_j|^{\frac{p}{p-1}} = 1 \right\}.$$

Dies folgert man seinerseits über die Theorie der Extremwerte mit Nebenbedingungen - Infinitesimalrechnung II.

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 Ω eine meßbare Punktmenge des $\mathbb{R}^n$ und $\tilde{L}^p = \tilde{L}^p(\Omega)$ die Menge aller meßbaren Funktionen $f\colon \Omega \to \mathbb{R}$ oder $\mathbb{C}$ (oder $\mathbb{R}^n$ oder $\mathbb{C}^n$) mit $\|f\|_p = \Big(\int\limits_\Omega |f|^p\,dx\Big)^{\frac{1}{p}} < \infty$ für $p < \infty$ und $\|f\|_\infty = \operatorname{ess\,sup}|f| < \infty$. Hierbei ist das *wesentliche Supremum* von f über Ω definiert durch

$$\operatorname{ess\,sup}|f| = \inf\{c \in \mathbb{R} \,|\, c \geq |f| \text{ fast überall in } \Omega\}\,,$$

d.h. das wesentliche Supremum ist das Infimum der „wesentlichen Schranken c“. Die so definierte Größe $\|f\|_p$ auf L^p erfüllt die Eigenschaften einer Norm mit Ausnahme der Definitheit, d.h. es ist zwar $\|f\|_p \geq 0$, doch es gibt von Null verschiedene Elemente g mit $\|g\|_p = 0$. Dies gilt für die Funktionen g aus der Menge der Funktionen

$$N = \{g \in \tilde{L}^p |\; g = 0 \text{ fast überall in } \Omega\}\,.$$

Um aus $\tilde{L}^p$ einen normierten Raum zu konstruieren, definiert man den Restklassenraum

$$L^p(\Omega) = \tilde{L}^p(\Omega)/N$$

und bezeichnet die Elemente aus $L^p(\Omega)$ ebenfalls – unpräziserweise – als Funktionen, wobei zwei Funktionen als nicht verschieden angesehen werden, wenn sie auf einer Menge vom Maß Null übereinstimmen.

d) Der Raum der stetigen Funktionen

Es sei $M \subset \mathbb{R}^n$ oder $\mathbb{C}^n$ und

$$C\big(\overline{M}\big) = \{f\colon M \to \mathbb{R} \text{ oder } \mathbb{C} \,|\, f \text{ ist gleichmäßig stetig auf } M\}\,.$$

Mit der Norm

$$\|f\|_\infty = \sup\{|f(x)|,\ x \in M\}$$

ist $C(\overline{M})$ ein Banachraum.

Es herrscht eine Verwechslungsmöglichkeit in der Bezeichnung $\|f\|_\infty$, wenn $M \cap B_R$ für eine Kugel B_R das Maß Null hat – das wesentliche Supremum würde diese Menge „vergessen". Ist M offen oder Abschließung einer offenen Menge, so ist für stetige Funktionen $\operatorname{ess\,sup}|f| = \|f\|_\infty$. (Ist M kompakt, wird das Supremum sogar angenommen.) Da $C\big(\overline{M}\big)$ ein Banachraum ist, ist in dem Fall, daß $\operatorname*{ess\,sup}_M |f| = \sup_M |f|$ gilt, der Raum $C\big(\overline{M}\big)$ ein abgeschlossener Teilraum. Ist Ω offen, so bezeichnet man üblicherweise $C(\Omega)$ als den Raum der stetigen Funktionen über Ω. Da für $f \in C(\Omega)$ der Fall eintreten kann, daß $f(x) \to \infty$ für $x \in \Omega,\ x \to x_0 \in \partial\Omega$, ist $C(\Omega)$ kein Banachraum bezüglich $\|.\|_\infty$.

e) Der Raum der Hölderstetigen Funktionen C^α, $0 < \alpha < 1$

Sei $M \subset \mathbb{R}^n$. Die *Hölderseminorm* einer Funktion $f\colon M \to \mathbb{R}$ oder $\mathbb{C}$ ist definiert durch

$$[f]_\alpha = \sup\left\{ \frac{|f(x) - f(y)|}{|x-y|^\alpha} \,\middle|\, x, y \in M,\ x \neq y \right\},$$

der Raum der Hölderstetigen Funktionen auf M ist definiert durch

$$C^\alpha(M) = \big\{ f : M \to \mathbb{R} \,\big|\, [f]_\alpha < \infty \big\}\,.$$

Mit der Norm

$$\|f\|_\alpha = \|f\|_\infty + [f]_\alpha$$

ist C^α 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 Ω eine offene Punktmenge des $\mathbb{R}^n$. Dann bezeichnet $C^m\big(\overline{\Omega}\big)$ bzw. $C^{m,\alpha}\big(\overline{\Omega}\big)$ die Menge aller Funktionen $f\colon \Omega \to \mathbb{R}$ oder C, die in Ω 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.

$$\|f\|_{\infty,m+\alpha} = \|f\|_{\infty,m} + [\nabla^m f]_\alpha$$

ist C^m und $C^{m,\alpha}$ ein Banachraum.

g) Die Sobolevräume $H^{m,p}(\Omega)$, $m \in \mathbb{N}$, $1 \leq p \leq \infty$.

Es sei Ω eine offene Punktmenge. Mit $\tilde{H}^{1,p}(\Omega)$ bezeichnen wir den linearen Raum aller Funktionen („Elemente“) aus $L^p(\Omega)$, die erste verallgemeinerte Ableitungen in $L^p(\Omega)$ haben, d.h. es existieren Funktionen $u^{(i)} \in L^p$ mit

$$\int\limits_\Omega u \frac{\partial \varphi}{\partial x_i}\, dx = - \int\limits_\Omega u^{(i)}\, \varphi\, dx$$

für alle $\varphi \in C_0^\infty(\Omega)$.

Viel benötigt werden die Sobolevschen Ungleichungen:

Satz 1.2 *Sei $\Omega \subset \mathbb{R}^n$, Ω beschränkt offen und $\partial\Omega$ Lipschitz–stetig. Dann existiert eine Konstante $K = K(n, \partial\Omega)$ mit*

$$\|u\|_q \leq K\, \|u\|_{1,p}, \qquad q = \frac{np}{n-p}$$

sofern $p < n$.

Ist $u \in H_0^{1,p}(\Omega)$, Ω offen $\subset \mathbb{R}^n$, so gilt ohne weitere Einschränkung an Ω

$$\|u\|_q \leq K\, \|\nabla u\|_p, \qquad q = \frac{np}{n-p}$$

mit einer universellen, nicht von Ω und n abhängigen Konstanten.

Im Grenzfall $p = n$ gelten zahlreiche Besonderheiten. Für den Fall $p > n$ gilt

Satz 1.3 *Sei $p > n$. Unter den Voraussetzungen an Ω wie in Satz 1.2 gilt*

$$H^{1,p}(\Omega) \subset C^\alpha(\overline{\Omega})$$

mit $\alpha = 1 - \frac{n}{p}$. Für die Normen gilt entsprechend

$$\|u\|_{C^\alpha} \leq K \|u\|_{H^{1,p}}\,.$$

Man schreibt $D_i u = u^{(i)}$ und auch $\nabla = (D_1, D_2, \dots, D_n)$. Mit der Norm

$$\|u\|_{1,p} = \|u\|_p + \|\nabla u\|_p$$

wird $H^{1,p}$ ein Banachraum. Analog werden die Räume $H^{m,p}$ definiert.

h) Die schwachen L^p–Räume („Marčinkiewič–Räume“)

Zur Motivation der folgenden Definition überlegen wir uns das

Lemma *Sei $u \in L^p(\Omega)$. Dann gilt für alle $L \geq 1$*

$$\mu\big(x \mid |u(x)| \geq L\big) \leq \frac{\|u\|^p}{L^p}\,. \tag{1.2}$$

Beweis:

$$\|u\|^p \geq \int_{(|u| \geq L)} |u|^p \, dx \geq L^p \, \mu(|u| \geq L)\,.$$

□

Eigenschaft (1.2) gibt Anlaß zur folgenden Definition:

Definition 1.5 *Eine meßbare Funktion u liegt* schwach in L^p*, wenn eine Konstante K existiert, so daß für alle $L \geq 1$*

$$\mu\big(x \mid |u(x)| \geq L\big) \leq \frac{K}{L^p}\,. \tag{1.3}$$

Man schreibt dann $u \in L^p_w(\Omega)$ (w wie „weakly“). Das Infimum der Konstanten in (1.3) ist die Norm in L^p_w. Daß dieses Infimum die Eigenschaft einer Norm hat, ist schwer zu beweisen. Man kann zeigen, daß für beschränktes Ω für $\varepsilon > 0$, $p - \varepsilon \geq 1$

$$L^p(\Omega) \subset L^p_w(\Omega) \subset L^{p-\varepsilon}(\Omega)$$

mit *echten* Inklusionen $\subset$.

Die schwachen L^p–Räume werden in der sogenannten Interpolationstheorie benötigt. Ein anderes Beispiel ist der Satz

Satz 1.4 *Sei Ω 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 $\nabla^2 u \in L^1_w$.

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 $1 < p < \infty$.

i) Die Morrey–Räume

Für die vertiefende Theorie der partiellen Differentialgleichungen sind die Morrey–Bedingung und die dazugehörenden Räume sehr wichtig.

Definition 1.6 *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*

$$\int_{B_R \cap \Omega} |u|^p \, dx \leq K R^{n-\lambda}\,. \tag{1.4}$$

gilt.

Der Morrey–Raum $L^{p,\lambda}(\Omega)$ ist derjenige Teilraum von $L^p(\Omega)$, dessen Elemente die Bedingung (1.4) erfüllen. Sehr ähnlich zu $L^{p,\lambda}(\Omega)$ ist der gewichtete L^p–Raum, der aus allen L^p–Funktionen mit der Eigenschaft

$$\int\limits_{\Omega} |u|^p \frac{1}{|x-x_0|^{n-\lambda}}\, dx \le K \tag{1.5}$$

besteht.

Die *Norm* in $L^{p,\lambda}$ wird definiert durch das Infimum der zulässigen Konstanten K in (1.4) oder, äquivalent

$$\|u\|_{L^{p,\lambda}} = \sup\left\{ R^{\lambda-n} \int\limits_{B_R\cap\Omega} |u|^p\, dx \;\middle|\; 0 < R < R_0,\ B_R(x) \subset \mathbb{R}^n,\ x_0 \in \mathbb{R}^n \right\}$$

mit $B_R(x_0) := \big\{x \,\big|\, |x-x_0| \le R\big\}$.

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:

Satz 1.5 *Sei $\Omega \subset \mathbb{R}^n$ beschränkt und habe Lipschitz–stetigen Rand. Die $H^{1,p}(\Omega)$–Funktion u genüge der Morrey–Bedingung*

$$\int\limits_{B_R\cap\Omega} |\nabla u|^p\, dx \le K R^{n-p+\alpha p} \tag{1.6}$$

mit einer Konstanten K für alle Kugeln $B_R \subset R^n$, $0 < R \le R_0$ 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 1.5 ist ein äußerst wichtiger Stetigkeitstest.

Satz 1.6 (Marčinkiewič–Stampacchia) *Sei $\Omega \subset \mathbb{R}^n$ beschränkt und habe Lipschitz–Rand $L^q(\Omega) \subset H^{1,p;\lambda}$ mit*

$$\frac{1}{q} = \frac{1}{p} - \frac{1}{\lambda},$$

sofern $\lambda > 0$. Die entsprechenden Normen werden entsprechend abgeschätzt:

$$\|u\|_q \le K\, \|u\|_{1,p;\lambda}.$$

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

Definition 1.7 *Eine Funktion $f\colon [a,b] \to \mathbb{R}$ hat* beschränkte Variation, *wenn für alle $N \in \mathbb{N}$ und alle Pärchen t_j, t'_j, $j = 1, \ldots, N$ mit $t_j < t'_j < t_{j+1}$ die sogenannte Totalvariation*

$$V(f) = \sup\left\{ \sum_{j=1}^{N} |f(t_j) - f(t'_j)| \,\middle|\, N \in \mathbb{N},\ t_j < t'_j < t_{j+1},\ j = 1, \ldots, N,\ t_j, t'_j \in [a,b] \right\} \tag{1.7}$$

endlich ist.

Die Menge der Funktionen mit beschränkter Variation bildet einen Banachraum $BV[a,b]$ bezüglich der Norm

$$\|f\|_{BV} = \|f\|_{L^1} + V(f)\,.$$

Im Fall von n Dimensionen betrachtet man additive, reellwertige Mengenfunktionen, die etwa auf den Lebesgue–meßbaren Teilmengen einer beschränkten meßbaren Menge Ω definiert sind und definiert als Totalvariation einer Mengenfunktion Φ

$$V(\Phi) = \sup\left\{ \sum_{j=1}^{N} \Phi(e_j) \,\middle|\, e_j \subset \Omega, e_j \text{ meßbar },\ e_j \cap e_k = \emptyset,\ j \neq k,\ j,k = 1, \ldots, N,\ N \in \mathbb{N} \right\}.$$

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.

Definition 1.8 *Ein* topologischer linearer Raum *ist ein Vektorraum V über $\mathbb{R}$ oder $\mathbb{C}$, 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 gegeben, wenn es zu jedem Punkt $x \in V$ ein System $\mathfrak{U}_x$ von Umgebungen $U \subset V$ gibt, so daß

1. $x \in U$ für $U \in \mathfrak{U}_x$,
2. $U \cap U' \in \mathfrak{U}_x$, falls $U, U' \in \mathfrak{U}_x$,
3. $W \in \mathfrak{U}_x$, falls ein $U \in \mathfrak{U}_x$ mit $U \subset W$ existiert.
4. Ist U eine Umgebung von x und $y \in U$, dann existiert eine Umgebung W von y mit $W \subset U$

Läßt man den dritten Punkt weg, so spricht man von einer Umgebungsbasis. Durch die Topologie wird ein Konvergenzbegriff und Stetigkeitsbegriff erklärt.
Die Verträglichkeit der algebraischen Operationen in V mit der Topologie bedeutet, daß diese stetige Funktionen bezüglich der Topologie sind.

Um eine Topologie auf V zu definieren, genügt es, eine Basis von Nullumgebungen anzugeben, da man die Umgebungsbasis eines allgemeinen Punktes durch Translation erhält.

Eine größere Bedeutung für Anwendungen in der Analysis haben die lokalkonvexen Räume.

Definition 1.9 *Ein* lokalkonvexer (topologischer linearer) Raum *ist ein topologischer linearer Raum mit einer aus* konvexen *Mengen bestehenden Null-Umgebungsbasis.*

Eine wichtige Klasse von lokalkonvexen Räumen sind Räume mit *Multi-Norm.*

Definition 1.10 *Sei V ein linearer Raum. Eine* Multi–Norm *ist eine Familie von Semi–Normen $\|.\|_\iota$, auf V, $\iota \in I$, die in ihrer Gesamtheit definit sind, d.h. aus $\|u\|_\iota = 0$ für alle $\iota \in I$ folgt $u = 0$.*

Erinnerung: Eine Semi–Norm erfüllt die Dreiecksungleichung und die positive Homogenität, braucht aber nicht definit zu sein.

Anmerkung: 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:**

1. Die Räume $C(\Omega), C^m(\Omega)$.

   Sei $\Omega \subset \mathbb{R}$ offen und $\Omega_j \subset \Omega$ offen, $\bigcup\limits_{j=1}^{\infty} \Omega_j = \Omega$. Der Raum $C(\Omega)$ bzw. $C^m(\Omega)$ aller auf Ω stetigen bzw. m–mal stetig differenzierbaren Funktionen wird mit Hilfe der Multi–Norm

   $$\|u\|_j = \sup_{\Omega_j} |u|$$

   bzw.

   $$\|u\|_j = \sup_{\Omega_j} \left(|u| + \sum_{k=1}^{m} |\nabla^k u| \right)$$

   ein lokalkonvexer Raum.

2. Die Räume $L^p_{\text{loc}}(\mathbb{R}^n)$, $H^{m,p}_{\text{loc}}(\mathbb{R}^n)$.

   Es handelt sich hierbei um die linearen Räume von Funktionen, deren Restriktion auf beschränkte Mengen in L^p bzw. $H^{m,p}$ liegt. Eine Multi–Norm ist

   $$\|u\|_j = \|u\|_{L^p(\Omega_j)}$$

   mit $\Omega_j \subset \mathbb{R}^n$, beschränkt und meßbar, $\bigcup\limits_{j=1}^{\infty} \Omega_j = \mathbb{R}^n$. Der Raum $H^{m,p}_{\text{loc}}$ wird analog definiert.
   Entsprechend bildet man $L^p_{\text{loc}}(\Omega)$–Räume für $\Omega \neq \mathbb{R}^n$.

3. Der Raum $\mathcal{D} = C_0^\infty(\Omega)$ (Schwartzscher Raum)

   Es sei $\Omega \subset \mathbb{R}^n$ offen. $C_0^\infty(\Omega)$ besteht aus allen unendlich *oft differenzierbaren* Funktionen f, deren Träger $\operatorname{supp} f = \overline{\{x \in \Omega \,|\, f(x) \neq 0\}}$ beschränkt ist. Sei $\Omega_j \subset \Omega$ offen und beschränkt, $\bigcup\limits_{j=1}^{\infty} \Omega_j = \Omega$. Eine Multi–Norm in $\mathcal{D}$ ist gegeben durch

   $$\|f\|_{j,m} := \sup_{\Omega_j} |\nabla^m f|, \qquad j = 1, 2 \ldots, \; m = 0, 1 \ldots, .$$

4. 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

   $$\|f\|_{j,m} = \sup_{\mathbb{R}^n} |\nabla^m f| \left(|x| + 1\right)^j .$$

Auch in lokalkonvexen linearen Räumen läßt sich definieren, wann eine Menge beschränkt ist.

Definition 1.11 *Eine Menge M in einem lokalkonvexen linearen Raum E heißt* beschränkt*, wenn*

$\sup\{\langle\varphi, x\rangle | x \in M\} < \infty$ *für jedes $\varphi \in E^*$ = Raum der stetigen linearen Funktionale.*

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 $\mathcal{D}$ und $\mathcal{S}$ gilt die Inklusion „Beschränktheit $\Rightarrow$ relative Kompaktheit“, obwohl sie unendlich-dimensional sind.

2 Beschränkte lineare Funktionale und beschränkte lineare Operatoren

Definition 2.1 *Ein* lineares Funktional *auf einem normierten Raum V ist eine lineare Abbildung von V in die reellen oder komplexen Zahlen. Das lineare Funktional heißt* beschränkt, *wenn eine Konstante K existiert mit*

$$|f(u)| \leq K \,\|u\| \quad \textit{für alle } u \in V\,.$$

Die Norm eines beschränkten linearen Funktionals f wird durch

$$\|f\| = \sup_{\|u\|=1} |f(u)| \tag{2.1}$$

definiert.

Lemma 2.1 *Ein lineares Funktional ist genau dann beschränkt, wenn es stetig ist.*

Der Beweis verläuft wie im Hilbertraum.

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 $\langle f, u\rangle = f(u)$ für $f \in V^*,\ u \in V$ üblich.

Lemma 2.2 *V^* ist bezüglich der in (2.1) definierten Norm ein Banachraum.*

Der Beweis geschieht ähnlich wie im Hilbertraum.

Mit Hilfe des Dualraums läßt sich die *schwache Konvergenz* in einem normierten Raum definieren.

Definition 2.2 *Eine Folge $(u_m)_{m\in\mathbb{N}}$ in einem normierten Raum V* konvergiert schwach gegen *ein Element $u \in V$, in Zeichen:*

$$u_m \rightharpoonup u \quad \textit{(schwach) in } V \ (m \to \infty)\,,$$

wenn für jedes $f \in V^*$

$$\langle f, u_m\rangle \to \langle f, u\rangle\,.$$

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*:*

$$f_m \overset{*}{\rightharpoonup} f \qquad \textit{im Sinne der $*$-schwachen Konvergenz}$$

bedeutet, daß für alle $u \in V$

$$\langle f_m, u\rangle \to \langle f, u\rangle \qquad (m \to \infty)\,.$$

Satz 2.1 ($*$–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 $*$–schwache 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 Teilfolge $\Lambda \subset \mathbb{N}$, so daß

$$\langle f_m, u_j\rangle \quad \text{konvergiert für alle } j,\ m \in \Lambda, m \to \infty\,.$$

Wir setzen

$$\langle f, u_j\rangle = \lim \langle f_m, u_j\rangle \quad m \in \Lambda,\ m \to \infty\,.$$

Da 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 $\langle f_m, u\rangle \to \langle f, u\rangle$ auf alle $u \in V$ durch Abschließung ausdehnen. Die Beschränktheit von f sieht man aus der Gleichung bzw. Ungleichung

$$|\langle f, u\rangle| = o(1) + |\langle f_m, u\rangle| \le o(1) + \|f_m\| \cdot \|u\| \le o(1) + K\|u\|\,.$$

Der $*$–schwache Limes der f_m und die Teilfolge sind damit konstruiert. □

Dualräume von konkreten Funktionenräumen

Satz 2.2 *Sei Ω 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

$$\langle f, u\rangle = \int_\Omega g \cdot u\, dx\,. \tag{2.2}$$

Umgekehrt beschreibt jedes $g \in L^q(\Omega)$ über Gleichung (2.2) ein Element $f \in (L^q(\Omega))^*$.

Aufgrund dieses Satzes sagt man, daß L^q der Dualraum zu L^p 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 Ω 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:

$$\langle f, u \rangle = \int_\Omega g\, u\, dx \qquad u \in L^2(\Omega)\,. \tag{2.3}$$

Wir zeigen, daß sogar $g \in L^{\frac{p}{p-1}}(\Omega)$ gilt. Durch Abschließung in L^p, da L^2 dort dicht ist, folgt dann aus (2.3) die behauptete Darstellung. Hierzu definieren wir:

$$g_L = \begin{cases} g & \text{falls} \quad |g| \leq L \\ 0 & \text{falls} \quad |g| > L \end{cases}$$

und wählen in (2.3) $u = |g_L|^s \operatorname{sign} g$. Da g_L beschränkt, ist $u \in L^2$ und damit zulässig in (2.3). Es gilt

$$\begin{aligned} \int_\Omega |g_L|^{s+1}\, dx = \int_\Omega g \cdot u\, dx &= \langle f, u \rangle \leq K \cdot \|u\|_p \\ &= K \Big(\int_\Omega |g_L|^{sp}\, dx \Big)^{\frac{1}{p}}. \end{aligned}$$

Wir wählen s so, daß $s + 1 = sp$, d.h. $s = \frac{1}{p-1}$. Wegen $s + 1 = \frac{p}{p-1}$ folgt dann

$$\int_\Omega |g_L|^{\frac{p}{p-1}}\, dx \leq K \Big(\int_\Omega |g_L|^{\frac{p}{p-1}} \Big)^{\frac{1}{p}},$$

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 $L \to \infty$ folgt aus der Definition des Lebesgue-Integrals für unbeschränkte Funktionen, daß $g \in L^{\frac{p}{p-1}}$. $\square$

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*

$$\sum_{k=1}^{N} f(\xi_k)\big(g(x_{k+1}) - g(x_k)\big)$$

bezüglich einer Zerlegung $a = x_0 < x_1 < \ldots x_N = b$ von $[a, b]$ und einer Belegung $\xi_k \in [x_k x_{k+1}]$. Das Riemann–Stieltjes–Integral $\int\limits_a^b f\, dg$ wird dann ähnlich wie das klassische Riemann–Integral als Grenzwert der Riemann–Stieltjes–Summen bei gegen Null gehender Feinheit
$\max\limits_{k=1,\ldots,N} |x_k - x_{k+1}|$ definiert.
Es gilt die Abschätzung

$$\left| \int\limits_a^b f\, dg \right| \leq \|f\|_\infty V(g)\,,$$

so daß $\int\limits_a^b f\, dg$ ein beschränktes lineares Funktional auf $C[a, b]$ definiert. $V(G)$ ist hierbei die Totalvariation von g, siehe Def. (1.7). Umgekehrt läßt sich zeigen, daß jedes $\varphi \in C^*[a, b]$ als Riemann–Stieltjes–Integral dargestellt werden kann. 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 δ_k mit

$$\langle \delta, f \rangle = f(0)\,.$$

Offensichtlich ist $\delta \in C^*[-1, 1]$. Die Darstellung als Stieltjes–Integral lautet:

$$\langle \delta, f \rangle = \int\limits_{-1}^{1} f\, dg$$

mit

$$g(t) = \begin{cases} 1 & t \geq 0 \\ 0 & t < 0 \end{cases}$$

Der Dualraum zu $L^\infty(\Omega)$.

Sei Ω meßbar und beschränkt, $\Omega \subset \mathbb{R}^n$, sowie ϕ eine additive reelle Mengenfunktion, welche auf den Lebesgue–meßbaren Teilmengen von Ω definiert ist. Additiv bedeutet:

$$\phi(e_1 \cup e_2) = \phi(e_1) + \phi(e_2)\,,\ e_1 \cap e_2 = \emptyset\,.$$

Man kann sich z.B. $\phi(e) = \int\limits_e w\,dx$ vorstellen, $w \in L^1$.

Definition 2.3 *ϕ ist von beschränkter Variation, wenn*

$$\sup\left\{\left|\phi\left(\bigcup_{j=1}^N e_k\right)\right| \,\middle|\, e_k \subset \Omega,\ e_k \text{ meßbar},\ \ e_j \cap e_k = \emptyset,\ j \neq k\right\} < \infty\,.$$

Für jedes ϕ von beschränkter Variation und $f \in C^\infty(\Omega)$ läßt sich das Lebesgue–Stieltjes–Integral über die entsprechenden Lebesgue–Stieltjes–Summen definieren:

$$\sum_{j=1}^N l_j\,\phi(l_{j-1} < f \leq l_j)\,,\quad l_0 < -\|f\|_\infty\,,\ l_N > \|f\|_\infty\,.$$

Die entsprechenden Limites werden als *Radon–Integral* oder *Lebesgue–Stieltjes–Integral* $\int\limits_\Omega f\,d\phi$ bezeichnet. Man kann zeigen, daß sich jedes beschränkte lineare Funktional auf $L^\infty(\Omega)$ in der Form

$$\int\limits_\Omega f\,d\phi$$

mit einer additiven, auf den Lebesgue-meßbaren Teilmengen von Ω definierten Mengenfunktion ϕ darstellen läßt. Hierbei hat $\phi(e)$ *nicht* notwendig die Gestalt $\int\limits_e w\,dx$.

Literatur: Kantorowitsch-Akilov, XI.4.

3 Das Prinzip der gleichmäßigen Beschränktheit

Dieses Kapitel behandelt einen grundlegenden Satz über Folgen beschränkter linearer Operatoren in Banachräumen.

Satz 3.1 *Sei (A_n) eine Folge linearer besschränkter Operatoren, die einen Banachraum B in einen normierten Raum V abbilden. Die Folge 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:

$$\sup_n \|\!|A_n|\!\| < \infty\,.$$

Anmerkung: Für *nichtlineare* Abbildungen ist die entsprechende Aussage bereits im $\mathbb{R}^1$ falsch:

$$f_n(x) = \begin{cases} |x| & \text{für } |x| \leq n \\ n & \text{für } |x| \geq n \end{cases}$$

Eine alternative Formulierung ist der Satz von der Festlegung der Singularität:

Lemma *Gilt $\sup_n \|\!|A_n|\!\| = \infty$, so gibt es ein x mit $\sup_n \|A_n x\| = \infty$.*

Beweis: Wir zeigen zunächst, daß für eine lineare Abbildung A aus der Ungleichung

$$\|Ax\| \leq C \qquad \text{für } x \in B_\delta(x_0) = \big\{y \in B \big|\ \|y - x_0\| \leq \delta\big\}$$

die Ungleichung

$$\|\!|A|\!\| \leq \frac{2C}{\delta} \tag{3.1}$$

folgt, d.h. ist A punktweise beschränkt auf einer kleinen Kugel mit Mittelpunkt x_0 – egal, wo x_0 liegt – so erfüllt die Operatornorm eine konkrete Abschätzung. Dies gilt entsprechend dann für Folgen von Operatoren.

Nachweis von (3.1): 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

$$\delta\|Ay\| \le C + \|Ax_0\| \le 2C\,.$$

Übergang zum Supremum bezüglich $\|y\| = 1$ liefert (3.1).

Es folgt nun der Hauptteil des Beweises: Angenommen, daß $\|A_n\| \to \infty$ $(n \to \infty)$. 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 (3.1), daß doch nicht $\|A_n\| \to \infty$. Da die Abbildung p auf jeder noch so kleinen Kugel unbeschränkt ist, liegt die Menge $E_k = \{x \in B \,|\, p(x) > k\}$ *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 $x \in U_\varepsilon(x_0)$. Dies wieder impliziert $p(x) > k$ für $x \in U_\varepsilon(x_0)$.

In *vollständigen* Räumen gilt jedoch, daß der Schnitt von abzählbar vielen offenen Mengen, die jeweils dicht in B liegen, *nicht leer* ist. Man beachte, daß z.B. schon im $\mathbb{R}^1$ der Schnitt abzählbar vieler offener Mengen leer sein kann. Wenn die offenen Mengen aber dicht liegen, geht dies nicht.

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.

Es ergibt sich somit die Existenz eines Elementes $x_0 \in \bigcap_{k=1}^{\infty} E_k$. Für dieses gilt $\sup_n \|A_n x_0\| \ge k$ für alle k, also $\sup_n \|A_n x_0\| = \infty$, was der Voraussetzung widerspricht. □

Zum Beweis von Satz 3.1 benötigen wir das folgende Lemma, welches eine duale Formulierung des sogenannten *Baireschen Kategoriensatzes* ist.

Lemma 3.1 *Es sei $\mathcal{O}_i$ eine absteigende Folge von offenen Mengen in einem Banachraum B. Jede der Mengen $\mathcal{O}_i$ sei dicht in B. Dann ist*

$$\bigcap_{i=1}^{\infty} \mathcal{O}_i \ne \emptyset\,.$$

Anmerkung: Läßt man die Forderung der Dichtheit von $\mathcal{O}_i$ fallen, ist die Aussage bereits im $\mathbb{R}^n$ falsch. Beispiel: $\mathcal{O}_i =]0, \frac{1}{i}[$. Es gilt $\mathcal{O}_i \supseteq \mathcal{O}_{i+1}$ und $\frac{1}{j} \notin \mathcal{O}$, $i > j$, sowie $0 \notin \mathcal{O}$. Der Schnitt der $\mathcal{O}_i$ muß daher leer sein.

Beweis: Wir konstruieren eine Folge von ineinander geschachtelten Kugeln $B_{\varepsilon_i}(x_i)$, $\varepsilon_i \to 0$, die in allen $\mathcal{O}_j$ liegen, so daß $x_i \to x^*$. Von diesem Punkt x^* zeigen wir, daß $x^* \in \mathcal{O}_i$

für alle $i \in \mathbb{N}$.
Start: 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 $\mathcal{O}_2 \cap B_{\varepsilon_1/2}(x_1)$ 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 ε_2 mit $B_{\varepsilon_2}(x_2) \subset B_{\varepsilon_1/2}(x_1) \cap \mathcal{O}_2$.
Es gilt $\varepsilon_2 \leq \frac{\varepsilon_1}{2}$.
Diese Konstruktion wird 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 $B_{\varepsilon_{n-1}}(x_{n-1}) \subset \bigcap_{j=1}^{n-1} \mathcal{O}_j$. Die Mittelpunkte x_n konvergieren. Zunächst gilt $\varepsilon_j \leq \varepsilon_1 \cdot 2^{-j+1}$, außerdem

$$x_{j+k} \in B_{\varepsilon_j/2}(x_j)\,, \qquad \text{d.h. } \|x_{j+k} - x_j\| \leq \varepsilon_j/2 \to 0 \quad (j \to \infty)\,.$$

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 $k \to \infty$ übergehen, es ergibt sich $x^* \in B_{\varepsilon_j/2}(x_j) \subset \mathcal{O}_j$. □

Eine äquivalente Folgerung aus Lemma 3.1 ergibt sich durch Komplementbildung: Die Komplemente der Mengen $\mathcal{O}_j$ in Lemma 3.1 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. Dies ist der sogenannte *Bairesche Kategoriensatz.*

Satz 3.2 *Die Vereinigung von abzählbar vielen abgeschlossenen, nirgends dichten Mengen eines Banachraumes kann nicht den ganzen Raum ergeben.*

Anmerkung: Satz und Beweis gelten analog in vollständigen metrischen Räumen.

Unter Benutzung schöner Vokabeln läßt sich Satz 3.2 folgendermaßen formulieren:

Satz 3.2' *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.

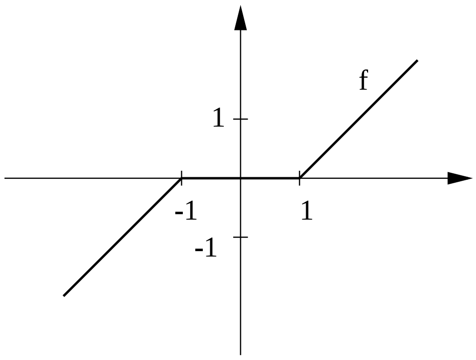
Ebenfalls eine Folge des Baireschen Kategoriensatzes ist der Satz von der offenen Abbildung.

Definition 3.1 *Eine Abbildung eines topologischen Raumes in einen andern heißt* offen, *wenn sie offene Mengen in offene überführt.*

Satz 3.3 (Satz von der offenen Abbildung) *Jede stetige surjektive lineare Abbildung $A\colon E \to F$ eines Banachraumes E in einen Banachraum F ist offen.*

Erläuterung: Betrachten wir den Fall $E = F = \mathbb{R}^n$. Stetige, surjektive nichtlineare Abbildungen sind selbst hier nicht offen. Beispiel:

$$f(x) = \begin{cases} 0 & \text{für} \quad |x| \le 1 \\ x-1 & \text{für} \quad x > 1 \\ x+1 & \text{für} \quad x < -1 \end{cases}$$



Offensichtlich ist $f\big(U_\varepsilon(1)\big)$ nicht notwendig offen.
Der Fall $E = F = \mathbb{R}^n$ mit einer linearen, surjektiven, stetigen 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 $f\colon \mathbb{R}^n \to \mathbb{R}^n$ offen ist.

In der Theorie der komplexen Funktionen gibt es ebenfalls einen Satz über die Gebietsinvarianz: „Sei $f\colon \mathbb{C} \to \mathbb{C}$ holomorph und nicht konstant. Dann ist f offen.“

Beweis des Satzes von der offenen Abbildung:

1. Schritt: Wir zeigen, daß eine Konstante $c > 0$ existiert mit

$$\overline{AB_1} \supset B_{2c} \tag{3.2}$$

Hierbei ist $B_r(z) = \big\{x \in E \text{ bzw. } F \;\big|\; \|x - z\| < r\big\}$, $B_r = B_r(0)$ und AB_1 das Bild von B_1 unter A, der Querstrich bedeutet die Abschließung.

Beweis: Setze $X_n = n\overline{AB_1}$. Da A surjektiv ist, gilt $\bigcup\limits_{n=1}^{\infty} X_n = F$, und nach dem Baireschen Kategoriensatz muß eines der X_n eine Kugel enthalten, d.h. $\operatorname{int} X_{n_0} \neq \emptyset$. Daraus folgt, daß auch $\operatorname{int} X_1 = \operatorname{int} \overline{AB_1} \neq \emptyset$, 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

$$-y_0 + B_{4c}(y_0) = B_{4c}(0) \subset \overline{AB_1} + \overline{AB_1} \subset 2\overline{AB_1}\,.$$

Daraus ergibt sich (3.2).

2. Schritt: Wir zeigen, daß mit obigem c

$$B_c \subset AB_1 \tag{3.3}$$

Beweis: Um (3.3) zu beweisen, müssen wir zu $y \in B_c$ ein $x \in B_1$ mit $Ax = y$ finden. Wegen (3.2) existiert zu jedem $\varepsilon > 0$ ein $z \in E$ mit $\|z\| < \frac{1}{2}$ und $\|y - Az\| < \varepsilon$, denn

$$\overline{AB_{1/2}} \supset B_c \ni y\,.$$

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 $\varepsilon = \frac{c}{4}$. Da $\overline{AB_{1/4}} \supset B_{c/2} \ni y - Az_1$, 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

$$\|z_n\| < \frac{1}{2^n} \quad \text{und} \quad \|y - A(z_1 + z_2 + \ldots z_n)\| < \frac{c}{2^n}\,.$$

Die Elemente $x_n = z_1 + z_2 + \ldots z_n$ bilden daher eine Cauchyfolge mit Limes x. Es gilt $\|x_n\| \leq \|z_1\| + \sum\limits_{j=2}^{\infty} \frac{1}{2^j}$ und somit

$$\|x\| \leq \|z_1\| + \sum_{i=2}^{\infty} \frac{1}{2^i} < \frac{1}{2} + \sum_{i=2}^{\infty} \frac{1}{2^i} = 1\,.$$

Ferner gilt $\|y - Ax_n\| \to 0$, und schließlich $y - Ax = 0$. Das gewünschte x mit $\|x\| < 1$ ist damit konstruiert. □

Aus dem Satz über die offene Abbildung folgt der Satz von der stetigen Inversen, da die Stetigkeit einer Abbildung äquivalent ist, daß Urbilder offener Mengen offen sind. Ist nun V Urbild einer offenen Menge der Gestalt $A^{-1}U$, U offen, so ist $A(A^{-1}U)$ offen, wenn A offen ist, d.h. V ist offen. Wir haben somit bewiesen:

Satz 3.4 (Satz von der stetigen Inversen) *Es seien E und F Banachräume und $A\colon E \to F$ linear, stetig und bijektiv. Dann ist A^{-1} stetig.*

Eine weitere elegante Folge ist

Satz 3.5 (Satz vom abgeschlossenen Graphen) *Seien E und F Banachräume, $A\colon E \to F$ 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 3.4 den Raum $G(A)$ mit der Graphennorm, als Bildraum den Raum E; als Abbildung $A_0\colon G(A) \to E$ definieren wir

$$A_0(x, Ax) = x_A\,.$$

Offensichtlich ist A_0 bijektiv. Außerdem ist A_0 beschränkt, da

$$\|A_0(x, Ax)\| = \|x\| \leq \|x\| + \|Ax\| = \|x\|_{G(A)}\,.$$

Nach Satz 3.4 ist damit A_0^{-1} stetig, d.h. für $x_j \to x$ folgt $(x_j, Ax_j) \to (x, Ax)$ und somit $Ax_j \to Ax$. □

Folgerung (Äquivalenz von Normen):

Satz 3.6 *Der Banachraum B sei bezüglich der Normen $\|.\|_0$ und $\|.\|_1$ 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 3.4 an mit $A = I$ als identischer Abbildung. Der Raum E bzw. F sei der mit $\|.\|_0$ bzw. $\|.\|_1$ 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 3.4 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:*

Satz 3.7 *Sei H ein Hilbertscher Raum und $A\colon H \to H$ 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! Ist nämlich $\|x_j - x_k\|_A \to 0$, so folgt $\|x_j - x_k\| \to 0$ und $\|Ax_j - Ax_k\| < \varepsilon$ für $j, k \geq n_0(\varepsilon)$. Hieraus schließen wir $|(Ax_j - Ax_k, \varphi_j)| < \varepsilon$ für alle φ_j mit $\|\varphi_j\| \leq 1$, j, k wie oben.

Da $(Ax_k, \varphi_j) = (x_k, A\varphi_j) \to (x, A\varphi_j) = (Ax, \varphi_j)$ für $k \to \infty$, folgt

$$|(Ax_j - Ax, \varphi_j)| < \varepsilon \qquad \|\varphi_j\| \leq 1\,,$$

und somit

$$\|Ax_j - Ax\| \leq \varepsilon \qquad j \geq n_0\,.$$

Da $\|x\| \leq \|x\|_A$ und H bezüglich beider Normen vollständig ist, folgt aus Satz 3.6 die Ungleichung

$$\|x\|_A \leq K\,\|x\|$$

und somit

$$\|Ax\| \leq (K-1)\,\|x\|\,.$$

□

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

$$Au := u' + cu = f \quad \text{ in } [0,T]\,, \qquad u(0) = 0 \in \mathbb{R}^n \tag{3.4}$$

mit gegebenen Funktionen $f \in L^p(0,T)$, $c \in L^\infty(0,T)$. Die Vektorfunktion $u \in H^{1,p}(0,T;\mathbb{R}^n)$ ist gesucht. Der in (3.4) definierte Operator A bildet den Raum $H^{1,p}_{0)} = \{v \in H^{1,p} |\ v(0) = 0\}$ nach $L^p(0,T)$ab. A ist stetig als Abbildung von $H^{1,p}_{0)}$ nach L^p, außerdem ist A surjektiv *und* injektiv, da sich (3.4) 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 (3.4) stetig von der rechten Seite f abhängt, d.h. konvergieren die Funktionen f in L^p, so konvergiert die Lösung 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 $A\colon C^1[a,b] \cap H_0^{1,p} \to C[a,b]$ auffaßt.

Sonstiges:

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.

Definition 3.2 *Ein lokalkonvexer Raum V heißt* tonneliert, *wenn jede abgeschlossene, konvexe, absorbierende Menge eine Nullumgebung ist.*
Eine Menge M heißt absorbierend, *wenn zu jedem $u \in V$ ein $\lambda \in \mathbb{R}$ existiert mit $\frac{u}{\lambda} \in M$.*

Wir betrachten lokalkonvexe Räume E und F und eine Menge $U = \{A_\iota | \ \iota \in J\}$ von linearen stetigen Operatoren $A_\iota \colon E \to F$.

Definition 3.3 *Die Menge U heißt* punktweise beschränkt, *wenn für jedes x die Menge $U(x) = \{A_\iota(x) | \ \iota \in J\}$ beschränkt ist (siehe Definition (1.11) zur Beschränktheit in lokalkonvexen Räumen)*

Definition 3.4 *Die Menge U heißt* gleichgradig stetig, *wenn man zu jeder Nullumgebung W im Raum F eine Nullumgebung V im Raum E finden kann, so daß*

$$U_\iota(V) \subset W \qquad \text{für alle } \iota \in J$$

gilt.

Der Begriff „gleichgradig stetig“ ersetzt die gleichmäßige Beschränktheit der Operatornormen.

Satz 3.8 *Sei E tonneliert. Dann ist jede punktweise beschränkte Menge von stetigen linearen Operatoren gleichgradig stetig.*

Beweis: siehe Kantorowitsch/Akilov, XI,3.

4 Der Satz von Hahn–Banach

Der Satz von Hahn–Banach befaßt sich mit der Fortsetzung von linearen Funktionalen, 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.

Satz 4.1 (Satz von Hahn–Banach, analytische Form, reeller Fall) *Sei V ein Vektorraum über $\mathbb{R}$ und $p\colon V \to \mathbb{R}$ subadditiv und positiv homogen, d.h. $p(\lambda x) = \lambda p(x)$ für alle $x \in E$ 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 $\varphi\colon W \to \mathbb{R}$ ein lineares Funktional mit*

$$\varphi(x) \leq p(x) \qquad \text{für alle } x \in W\,.$$

Dann gibt es eine Fortsetzung $f\colon V \to \mathbb{R}$ von φ, d.h. $f|_W = \varphi$, so daß

$$f(x) \leq p(x) \qquad \text{für alle } x \in V\,.$$

Anmerkung: Besonders wichtig ist der Fall, daß $p(x) = \|x\|$, wenn V normierter Raum mit Norm $\|.\|$ ist.

Eine Anwendung für Ungeduldige: Nicht jedes beschränkte lineare Funktional φ auf $L^\infty(a,b)$ läßt sich in der Form

$$\varphi(z) = \int_a^b g\, z\, dx\,, \qquad z \in L^\infty(a,b)$$

mit einer Funktion $g \in L^1$ darstellen. In anderen Worten:

Satz *$L^1(a,b)$ ist ein echter Teilraum von $(L^\infty)^*$.*

Da $(L^1)^* = L^\infty$, folgt der

Satz

$L^1(a,b)$ ist nicht reflexiv.

Beweis (der Aussage $(L^\infty)^* \neq L^1$): Der Raum $C(a,b)$ ist ein echter Teilraum von L^∞. Wir betrachten das Dirac–Funktional $\varphi = \delta$ mit

$$\delta(z) = z(0)\,, \qquad z \in C(a,b)$$

((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^∞ mit $f(z) \le \|z\|$ für alle $z \in L^\infty$. Die letzte Ungleichung gilt auch für $-z$, somit ist $|f(z)| \le \|z\|$ und $f \in (L^\infty)^*$. Das Dirac–Funktional läßt sich aber nicht in der Form $\delta(z) = \int\limits_a^b z\,dx$ darstellen, somit auch nicht das Funktional f.
□

Zum Beweis von Satz 4.1 zeigen wir zunächst, daß das Funktional φ in Satz 4.1 auf einen größeren Raum $W \oplus \langle x_0 \rangle$ unter Erhaltung der Ungleichung $\varphi(x) \le p(x)$ fortgesetzt werden kann.

Lemma 4.1 (Elementarerweiterung) *Unter den Voraussetzungen von Satz 4.1 gibt es zu jedem $x_0 \in V - W$ eine lineare Fortsetzung $f_1 : W \oplus \langle x_0 \rangle \to \mathbb{R}$ von $\varphi\colon W \to \mathbb{R}$, so daß*

$$f_1(y) \le p(y) \qquad \textit{für alle } y \in W \oplus \langle x_0 \rangle\,. \tag{4.1}$$

Beweis: Wir setzen

$$f_1(x + tx_0) = \varphi(x) + t\alpha\,,$$

wobei α noch definiert werden muß. Jedenfalls ist f_1 linear auf $W \oplus \langle x_0 \rangle$. Um ein geeignetes α zu finden, welches die Ungleichung (4.1) erfüllt, beachten wir, daß wegen der positiven Homogenität von p nur sichergestellt werden muß, daß für alle $x \in W$

$$\begin{aligned} \varphi(x) + \alpha &\le p(x + x_0) \\ \varphi(x) - \alpha &\le p(x - x_0) \end{aligned} \tag{4.2}$$

gilt. Ersetzt man in (4.2) x durch x/t, $t > 0$, und multipliziert mit t, so ergibt sich (4.1) für $y = x \pm tx_0$. Um (4.2) zu gewährleisten, muß man offensichtlich

$$\sup_{x\in W}\{\varphi(x) - p(x - x_0)\} \le \alpha \le \inf_{x \in W}\{p(x + x_0) - \varphi(x)\}$$

sicherstellen. Eine solche Wahl von α ist aber möglich, da

$$\varphi(\xi) - p(\xi - x_0) \le p(x + x_0) - \varphi(x)\,, \qquad x \in W\,. \tag{4.3}$$

Es gilt nämlich

$$\varphi(x) + \varphi(\xi) = \varphi(x + \xi) \leq p(x + \xi) \leq p(x + x_0) + p(\xi - x_0)\,,$$

woraus (4.3) 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 in separablen normierten Räumen, $p(x) = K\|x\|$**:** Es sei $\{x_j \,|\, j \in \mathbb{N}\}$ eine Menge von linear unabhängigen Vektoren aus V, deren lineare Hülle in V dicht ist. Sei
$\{x_j \,|\, j \in \Lambda \subset \mathbb{N}\}$ 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 φ auf $W \oplus \langle x_j; j \in \Lambda \rangle$ unter Beibehaltung der Ungleichung $\varphi(x) \leq K\|x\|$. Durch Abschließung erhalten wir eine Fortsetzung auf den ganzen Raum V, d.h. wir definieren

$$f(x) = \lim_{y_k \to x} f_1(x)\,, \qquad y_k \in W \oplus \langle x_j; j \in \Lambda \rangle\,.$$

□

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

$$a \prec b \qquad \text{oder} \qquad b \prec a\,.$$

Wenn einer der beiden Fälle zutrifft, sagt man dann „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

$$a \leq b \Leftrightarrow a_i \leq b_i \qquad \text{für die Komponenten } a_i \text{ und } b_i\,,\ i = 1, \ldots, n\,.$$

(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 der Halbordnung $\leq$ ist eine Kette.

Eine obere *Schranke* für eine Teilmenge S einer Menge M mit Halbordnung $\prec$ ist ein Element a mit

$$s \prec a \qquad \text{für alle } s \in S\,.$$

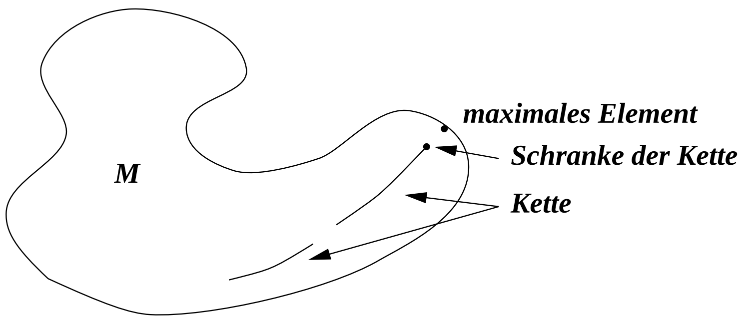
Ein Element $m \in M$ heißt *maximal* (auf M), wenn es von keinem anderen Element $x \in M$ bezüglich $\prec$ übertroffen werden kann, d.h. aus

$$m \prec x \qquad \text{folgt } x = m\,.$$

(Es ist aber durchaus möglich, daß m maximal ist und mit dem Element y nicht vergleichbar ist.)

Lemma 4.2 (Lemma von 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 1 versehen. Man überlegt sich – etwa mit Hilfe komponentenweiser Supremumsbildung – daß jede Kette eine 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 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 des 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. α 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) Irgendein $x_0 \in M$ gehöre zu G.

(ii) Ist für $\beta < \gamma,\ \beta, \gamma \in I$, entschieden, welche x_β zu G gehören, so gehöre x_γ 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_α 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: Wir betrachten die Menge

$$M = \big\{h\colon D(h) \underset{\text{lin}}{\longrightarrow} \mathbb{R} \big| \, D(h) \text{ lin. Teilraum von } V,\ W \subset D(h),$$
$$h|_W = \varphi,\ h(x) \le p(x)\ \forall x \in D(h)\big\}\,.$$

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, $K = \{h_i \in M | \ i \in I\}$, 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 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 4.1 durch. f wäre dann nicht maximal. $\square$

Im *komplexen Fall* sind die Ungleichungen in der reellen Formulierung des Satzes von Hahn–Banach – Satz 4.1 – folgendermaßen abzuändern: Für das lineare Funktional $\varphi\colon W \to \mathbb{C}$ setzt man

$$|\varphi(x)| \le p(x)$$

voraus und erhält eine Fortsetzung $f: V \to C$ von φ mit $|f(x)| \le p(x)$.

Geometrische Formulierung des Satzes von Hahn-Banach als Trennung konvexer Mengen durch Hyperebenen:

Wie im endlich-dimensionalen Fall hat man die

Definition 4.1 *Eine* Hyperebene *in einem Vektorraum V über $\mathbb{R}$ oder $\mathbb{C}$ ist eine Menge der Gestalt*

$$H = \{x \in V | \ \varphi(x) = \alpha\} =: (\varphi = \alpha)$$

mit einer linearen Abbildung $\varphi : V \to \mathbb{R}$ bzw. $\mathbb{C}$, $\varphi \neq 0$, und einer Zahl $\alpha \in \mathbb{R}$ oder $\mathbb{C}$.

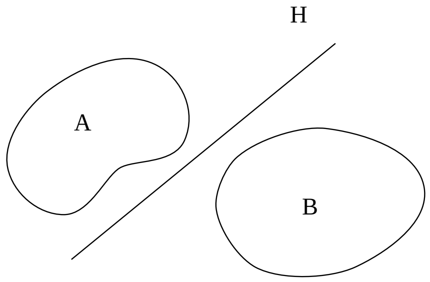
Lemma 4.3 *Sei V ein normierter Vektorraum. Eine Hyperebene $(\varphi = \alpha)$ ist genau dann abgeschlossen, wenn φ stetig ist.*

(Beweis: Übungsaufgabe)

Definition 4.2 *Seien A, B Teilmengen des normierten Vektorraumes V. Man sagt, „die Hyperebene $(\varphi = \alpha)$* trenne *A und B“, wenn*

$$\varphi(x) \leq \alpha \leq \varphi(y) \quad \text{für alle } x \in A \text{ und } y \in B\,. \tag{4.4}$$

(Auf die Reihenfolge von A und B kommt es nicht an, d.h. (4.4) bedeutet auch, daß $(\varphi = \alpha)$ die Mengen B und A trennt.)



Man spricht von „strikter Trennung“, wenn ein $\varepsilon > 0$ existiert mit

$$\varphi(x) + \varepsilon \leq \alpha \leq \varphi(y) - \varepsilon\,, \qquad x \in A,\ y \in B\,.$$

Definition 4.3 *Eine Menge $A \subset V$ heißt* konvex *genau dann, wenn für $x, y \in A$ und $t \in [0,1]$*

$$tx + (1-t)y \in A\,.$$

Satz 4.2 (Satz von Hahn–Banach, geometrische Form) *Sei V ein normierter Raum und A, B zwei nicht–leere, 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 $C \subset V$, $0 \in C$

$$p_C(x) := \inf\{\alpha > 0 \,|\, \alpha^{-1}x \in C\}\,.$$

Lemma 4.4 *Sei C offene konvexe Teilmenge des normierten Vektorraumes V. Es sei $0 \in C$. Dann gilt*

$$0 \leq p_C(x) \leq M\,\|x\| \qquad \text{mit einer Konstanten } M \tag{4.5}$$

und

$$C = \{x \in V \,|\, p_C(x) < 1\}\,. \tag{4.6}$$

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 $p_C(x) \le \alpha_0 = \frac{\|x\|}{r}$ nach Definition von p_C. Daraus folgt (4.5).

- Die positive Homogenität von p_C folgt aus

$$\begin{aligned} p_C(\lambda x) &= \inf\{\alpha > 0 \,|\, \alpha^{-1}\lambda x \in C\} = \inf\{\beta\lambda \,|\, \beta^{-1}x \in C\} = \\ &= \lambda \inf\{\beta \,|\, \beta^{-1}x \in C\} = \lambda\, p_C(x)\,, \qquad \lambda > 0\,. \end{aligned}$$

- Wir beweisen (4.6): Ist $x \in C$, so gilt auch $(1+\varepsilon)\,x \in C$ für genügend kleines ε wegen der Offenheit von C. Daher $p_C(x) \le \frac{1}{1+\varepsilon} < 1$, d.h. $C \subset \{x \in V \,|\, p_C(x) < 1\}$. 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: Sei $x, y \in V,\ \varepsilon > 0$. Da

$$p_C\left(\frac{x}{p_C(x)+\varepsilon}\right) = \frac{1}{p_C(x)+\varepsilon}\, p_C(x) < 1\,,$$

folgt

$$\frac{x}{p_C(x)+\varepsilon} \in C \quad \text{und entsprechend } \frac{y}{p_C(y)+\varepsilon} \in C\,.$$

Aus Konvexitätsgründen ist

$$\frac{tx}{p_C(x)+\varepsilon} + \frac{(1-t)\,y}{p_C(y)+\varepsilon} \in C\,, \qquad 0 \le t \le 1\,.$$

Setzt man

$$t = \frac{p_C(x)+\varepsilon}{p_C(x)+p_C(y)+2\varepsilon}\,,$$

folgt

$$\frac{x+y}{p_C(x)+p_C(y)+2\varepsilon} \in C\,.$$

Nach Definition von C folgt daraus

$$p_C(x+y) \le p_C(x) + p_C(y) + 2\varepsilon\,.$$

Grenzübergang $\varepsilon \to 0$ ergibt die Behauptung.

$\square$

Lemma 4.5 (Trennung einer konvexen Menge und eines Punktes) *Sei V 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*

$$f(x) < f(x_0) \qquad \text{für alle } x \in C\,.$$

Beweis: O.B.d.A. sei $0 \in C$. Auf dem eindimensionalen Raum $\langle x_0 \rangle$ definieren wir das lineare Funktional φ durch

$$\varphi(tx_0) = t\,.$$

Es gilt $\varphi(x) \leq p_C(x)$, da andernfalls

$$p_C(x) < \varphi(x)\,, \quad \text{d.h. } p_C(tx_0) < t \text{ und } p_C(x_0) < 1\,.$$

Dies würde nach Lemma 4.4 $x_0 \in C$ bedeuten. Nach dem Satz von Hahn–Banach in der analytischen Form läßt sich φ zu einem linearen Funktional $f\colon V \to \mathbb{R}$ fortsetzen mit

$$f(x) \leq p_C(x)\,.$$

Insbesondere ist $f(x_0) = 1$ und andererseits $f(x) < 1$ für $x \in C$ wegen (4.6). □

Wir schreiten nun zum Beweis von Satz 4.2:

Beweis: Die Menge $C = A \ominus B = \{a - b \mid a \in A,\ b \in B\}$ 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 4.5 gibt es $f \in V^*$ (d.h. ist stetig und linear) mit

$$\begin{aligned} f(z) < f(0) = 0\,, &\qquad z \in C\,,\\ f(x) < f(y)\,, &\qquad x \in A,\ y \in B\,. \end{aligned}$$

Wir wählen α so, daß

$$\sup_{x \in A} f(x) \leq \alpha \leq \inf_{y \in B} f(y)\,.$$

Die Hyperebene $(f = \alpha)$ ist wegen Lemma 4.3 abgeschlossen und trennt A und B.
□

Im Unendlich-dimensionalen lassen sich beliebige konvexe Mengen nicht notwendig trennen.

Beispiel: $V = L^2(\Omega)$ mit L^2–Norm, Ω offen, $\subset \mathbb{R}^n$.

$$A = \left\{ z \in C(\Omega) \cap L^2 \middle|\ \|z\|_2 < 1 \right\}, \qquad B = \{y_0\}$$

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

$$f(z) < f(y_0), \qquad z \in C(\Omega) \cap L^2, \quad \|z\|_{L^2} < 1,$$

gilt, so folgt

$$(u_0, z) < (u_0, y_0), \qquad \|z\|_{L^2} < 1,$$

und, da $C(\Omega) \cap B_1$ dicht in $B_1 = \left\{ w \in L^2 \middle|\ \|w\|_2 \leq 1 \right\}$, ist

$$\|u_0\| < (u_0, y_0) \leq \tfrac{1}{2}\|u_0\|.$$

q.e.d.

Wir notieren ohne Beweis eine Variante der geometrischen Form des Satzes von Hahn–Banach:

Satz 4.3 *Sei V ein normierter Vektorraum und $A, B \subset V$ konvexe, disjunkte nichtleere Teilmengen von V. Ferner sei A abgeschlossen und B kompakt. Dann gibt es eine abgeschlossene Hyperebene, die A und B strikt trennt.*

Beweis siehe Brezis, Analyse fonctionnelle, Th.1.7. (Man approximiert A durch die offene Menge $A + B_\varepsilon$.)

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: In einer Spezialisierung werden disjunkte konvexe Mengen, in denen die eine einen *„algebraisch inneren“* Punkt enthält, durch Hyperebenen getrennt.

Eine Reihe von Sätzen über Hilberträume – z.B. über die normale Lösbarkeit linearer Gleichungen – läßt sich – abgeändert – auf Banachräume B übertragen, wenn man mit dem Raumpaar 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 $\|f\|_{V^*} = \sup\limits_{\substack{\|u\| \leq 1 \\ u \in V}} \langle f, u\rangle$

Lemma 4.6 *Sei V normierter Raum und $W \subset V$ 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. □

Lemma 4.7 *Sei V normierter Vektorraum. Zu jedem $u \in V$ existiert ein $f \in V^*$ mit $\|f\|_{V^*} = \|u\|_V$ und $\langle f, u\rangle = \|u\|^2$.*

Beweis: Man wendet Lemma 4.6 mit $W = \langle u\rangle$ und $\varphi(tu) = t\,\|u\|^2$ an. Es gilt dann

$$\|\varphi\| = \sup\left\{|t|\,\|u\|^2 \,\middle|\, |t|\|u\|| \leq 1\right\} = \|u\|\,.$$

Nach dem Satz von Hahn–Banach gibt es eine Fortsetzung $f \in V^*$ von φ 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

$$|\langle f, v\rangle| \leq \|u\| \cdot \|v\|\,.$$

□

Folgerung:

Lemma 4.8 *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 4.7 konstruierte f und erhält $0 = \langle f, u\rangle = \|u\|^2$. □

Anmerkung: Das Element f in Lemma 4.7 muß nicht eindeutig sein – dies gilt nur, wenn die Norm in V^* *strikt* konvex ist, d.h. für jedes Pärchen $f_1, f_2 \in V^*$ gilt

$$\|tf_1 + (1-t)\,f_2\| < t\,\|f_1\| + (1-t)\,\|f_2\|\,, \qquad 0 < t < 1\,.$$

Die Abbildung, die jedem $u \in V$ die Menge der Elemente f zuordnet, so daß die Aussagen von Lemma 4.7 gelten, nennt man die *Dualitätsabbildung.*

Nach Definition von $\|f\|_{V^*}$, $f \in V^*$, gilt

$$\|f\|_{V^*} := \sup\left\{|\langle f, u\rangle| \,\middle|\, \|u\| \leq 1, u \in V\right\}, \tag{4.7}$$

nach Lemma 4.7 hingegen gilt

$$\|u\| = \max\{|\langle f, u\rangle| \mid \|f\|_{V^*} \leq 1\,, \quad f \in V^*\}\,, \tag{4.8}$$

denn es gibt ein f_0 mit $\langle f_0, u\rangle = \|u\|^2\,,\ \|f_0\|_{V^*}\rangle = \|u\|$. Das Element $\|u\|^{-1} f_0$ realisiert das Maximum in (4.8).
Man kann beweisen, daß auch das „sup“ in (4.7) genau dann angenommen wird, wenn der Raum reflexiv ist.

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

$$g\colon V \to \mathbb{R} \cup \{\infty\}$$

heißt *konvex*, wenn

$$g\big(tx + (1-t)\,y\big) \leq tg(x) + (1-t)\,g(y) \qquad \text{für alle } x, y \in V, 0 \leq t \leq 1.$$

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.

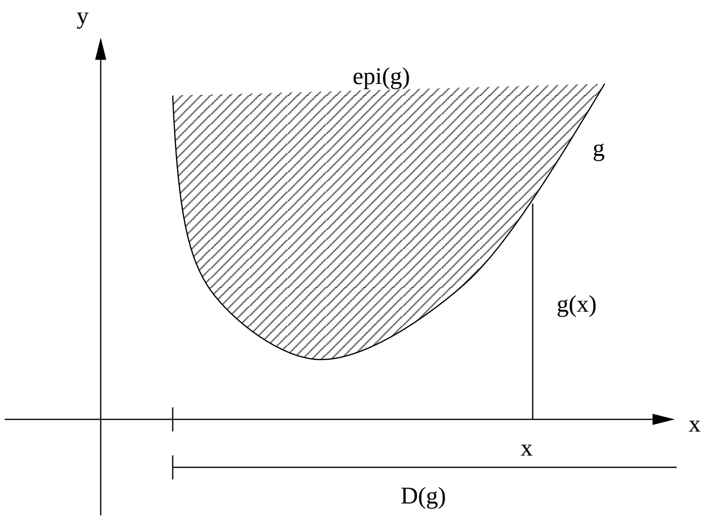
Definition 5.1 *$g\colon D(g) \to \mathbb{R} \cup \{\infty\}$ heißt* unterhalbstetig, *wenn für alle konvergenten Folgen (x_n) aus V mit $x_n \to x$ die Ungleichung*

$$g(x) \leq \liminf_{n\to\infty} g(x_n)$$

gilt.

Definition 5.2 *Der* Epigraph *einer konvexen Funktion $g\colon V \to \mathbb{R} \cup \{\infty\}$ ist die Menge*

$$\operatorname{epi} g = \Big\{(x, \lambda) \in V \times \mathbb{R} \Big|\; g(x) \leq \lambda\Big\}\,.$$



Lemma 5.1 *g ist genau dann unterhalbstetig, wenn der Epigraph von g abgeschlossen ist.*

Beweis: Übungsaufgabe.

Wenn g konvex ist, ist auch $\operatorname{epi} g$ konvex; die Umkehrung ist ebenfalls richtig.

Definition 5.3 *Sei $g\colon V \to \mathbb{R} \cup \{\infty\}$ und $g \not\equiv \infty$, d.h. $D(g) \neq \emptyset$. Die* duale Funktion $g^*\colon V^* \to \mathbb{R} \cup \{\infty\}$ *ist definiert durch*

$$g^*(f) = \sup_{u \in V} \big\{ \langle f, u \rangle - g(u) \big\}\,, \qquad f \in V^*\,.$$

Es ist klar, daß g^* konvex *ist, selbst wenn g nicht konvex sein sollte.*

Beispiel: $V = \mathbb{R}$, $\quad g(x) = \frac{|x|^p}{p}$

$$g^*(\xi) = \sup_{x \in \mathbb{R}} \left\{ \xi \cdot x - \frac{|x|^p}{p} \right\}.$$

Das Supremum x^* wird angenommen; es gilt dann $\xi - |x^*|^{p-1} \operatorname{sign} x^* = 0$ und somit

$$g^*(\xi) = |x^*|^{p-1} (\operatorname{sign} x^*) \cdot x^* - \frac{|x^*|^p}{p} = \left(1 - \frac{1}{p}\right) |x^*|^p\,.$$

Wegen $|x^*| = |\xi|^{\frac{1}{p-1}}$ folgt

$$g^*(\xi) = \frac{|\xi|^q}{q} \qquad \text{mit } q = \frac{p}{p-1}\,.$$

Lemma 5.2 *Sei V ein normierter Raum und $g\colon V \to \mathbb{R} \cup \{\infty\}$ konvex, unterhalbstetig, und es sei $g \not\equiv \infty$. 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 Variante „Trennung einer konvexen abgeschlossenen Menge und einer konvexen und kompakten Menge“ an. Als Grundraum wählen wir $V \times \mathbb{R}$, als abgeschlossene konvexe Menge $A = \operatorname{epi} g$ und $B = (x_0, \lambda_0)$. Nach der Definition von $\operatorname{epi} g$ ist $A \cap B = \emptyset$. A ist konvex und abgeschlossen und B kompakt, da letztere nur aus einem Punkt besteht. Es gibt daher eine trennende abgeschlossene Hyperebene H, die A und B strikt trennt. Die Hyperebene hat die Gestalt

$$H = \big\{ (x, \lambda) \in V \times \mathbb{R} \,|\, \langle f, x \rangle + k\lambda = \alpha \big\}$$

mit einem $f \in V^*$, $k \in \mathbb{R}$, $\alpha \in \mathbb{R}$. Es gilt hierbei

$$\langle f, x_0 \rangle + k\lambda > \alpha \qquad \forall (x,\lambda) \in \operatorname{epi} g\,. \tag{5.1}$$

$$\langle f, x \rangle + k\lambda_0 < \alpha\,. \tag{5.2}$$

Da $\big(x_0, g(x_0)\big) \in \operatorname{epi} g$, folgt

$$\langle f, x_0 \rangle + kg(x_0) > \alpha > \langle f, x_0 \rangle + k\lambda_0\,.$$

Daraus folgt $k > 0$. Aus (5.1) folgt mit $\lambda = g(x)$

$$\big\langle -\tfrac{1}{k} f, x \big\rangle - g(x) < -\tfrac{\alpha}{k}\,, \qquad x \in D(g)\,,$$

und damit

$$g^*\big(-\tfrac{1}{k} f\big) \leq \tfrac{\alpha}{k} < \infty\,.$$

□

Die *biduale konvexe Funktion* zu g hat die Gestalt $g^{**}(x) = \sup\limits_{f \in V^*} \big\{\langle f, x\rangle - g^*(f)\big\}$.

Satz 5.1 *Sei V normierter Raum und $g\colon V \to \mathbb{R} \cup \{\infty\}$ konvex und unterhalbstetig sowie $g \not\equiv \infty$. Dann gilt $g^{**} = g$.*

Beweis: Sei zunächst $g \geq 0$. 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

$$g(x) \geq g^{**}(x)\,.$$

(Dies gilt auch ohne die Voraussetzung $g \geq 0$.)
Angenommen, es würde $g^{**} \neq g$ gelten. Dann gäbe es ein x_0 mit

$$g^{**}(x_0) < g(x_0)\,. \tag{5.3}$$

($g(x_0) = \infty$ ist zugelassen, aber dann beinhaltet die Widerspruchsannahme $g^{**}(x_0) < \infty$.)
Wir wenden den Satz von Hahn–Banach an und trennen die Menge $\operatorname{epi} g$ und den Punkt $(x_0, g^{**}(x_0))$, denn wegen (5.3) liegt $(x_0, g^{**}(x_0))$ nicht in $\operatorname{epi} g$. 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

$$\langle f, x \rangle + k\lambda > \alpha \qquad \forall (x,\lambda) \in \operatorname{epi} g\,, \tag{5.4}$$

$$\langle f, x_0 \rangle + kg^{**}(x_0) < \alpha\,. \tag{5.5}$$

Läßt man in (5.4) λ gegen ∞ gehen, folgt $k \geq 0$. Aus (5.4) schließt man mit $\lambda = g(x) \geq 0$

$$\langle f, x\rangle + (k+\varepsilon)\, g(x) > \alpha\,, \qquad\qquad x \in D(g)\,,$$

und damit

$$\Big\langle -f\,\tfrac{1}{k+\varepsilon},\ x\Big\rangle - g(x) < -\tfrac{\alpha}{k+\varepsilon}\,,$$

und durch Übergang zum Supremum in x

$$g^*\big(-f\,\tfrac{1}{k+\varepsilon}\big) \leq -\tfrac{\alpha}{k+\varepsilon}\,.$$

Weiterhin folgt nach Definition von $g^{**}(x_0)$

$$\begin{aligned} g^{**}(x_0) &\geq \Big\langle -f\,\tfrac{1}{k+\varepsilon},\ x_0\Big\rangle - g^*\big(-f\,\tfrac{1}{k+\varepsilon}\big) \geq \\ &\geq \Big\langle -f\,\tfrac{1}{k+\varepsilon},\ x_0\Big\rangle + \tfrac{\alpha}{k+\varepsilon}\,. \end{aligned}$$

Damit folgt

$$\langle f, x_0\rangle + (k+\varepsilon)\, g^{**}(x_0) \geq \alpha\,.$$

Der Grenzübergang $\varepsilon \to 0$ führt zu einem Widerspruch mit (5.5).

Dies war der Fall $g \geq 0$.
Den allgemeinen Fall erhält man durch Abändern von g mit einem $f_0 \subset D(g^*)$

$$\overline{g}(x) = g(x) - \langle f_0, x\rangle + g^*(f_0) \geq 0\,.$$

Für $\overline{g}$ weiß man, daß $\overline{g}^{**} = \overline{g}$. Es gilt

$$\begin{aligned} \overline{g}^*(f) &= \sup_{x\in V} \big\{\langle f, x\rangle - \overline{g}(x)\big\} = \\ &= \sup_{x\in V} \big\{\langle f + f_0, x\rangle - g(x) - g^*(f_0)\big\} = \\ &= g^*(f+f_0) - g^*(f_0) \\ \overline{g}^{**}(x) &= \sup_{f\in V^*} \{\langle f+f_0, x\rangle - g^*(f+f_0)\} + g^*(f_0) \\ &= \sup_{\tilde{f}\in V^*} \{\langle \tilde{f}, x\rangle - g^*(\tilde{f})\} - \langle f_0, x\rangle + g^*(f_0) \\ &= g^{**}(x) - \langle f_0, x\rangle + g^*(f_0) \end{aligned}$$

Daraus folgt $g(x) = g^{**}(x)$. $\square$

Orlicz–Räume

Definition 5.4 *Es sei $g\colon \mathbb{R} \to \mathbb{R}$ eine nichtnegative konvexe Funktion und $\Omega \subset \mathbb{R}^n$ meßbar. Mit $\hat{L}_g$ („*Orliczklasse*“) bezeichnen wir die Menge aller meßbaren Funktionen $f\colon \Omega \to \mathbb{R}$ mit*

$$G(f) := \int_\Omega g(f)\,dx < \infty\,.$$

Der Orlicz–Raum *L_g ist die Menge aller Äquivalenzklassen $[f]$ von Funktionen f, für die*

$$\|f\|_g = \sup\left\{ \int_\Omega \varphi \cdot f\,dx \,|\, G^*(\varphi) \le 1 \right\} < \infty$$

ausfällt. Hierbei ist $[f]$ die Klasse aller Funktionen, die sich von f nur auf einer Menge vom Maß Null unterscheiden, und G^* ist die zu G duale konvexe Funktion. Die Größe $\|f\|_g$ nennt man die Orlicz–Norm *von f zur Funktion g.*

Der Kuhn–Tucker–Satz der konvexen Optimierung

Die Funktionen $F\colon \mathbb{R}^n \to \mathbb{R}$ und $g\colon \mathbb{R}^n \to \mathbb{R}^m$ seien stetig und konvex. Wir betrachten die Optimierungsaufgabe. Gesucht ist $x^* \in \mathbb{R}^n$, so daß

$$F(x^*) \le F(x) \qquad \text{für alle} \quad x \in \mathbb{R}^n \text{ mit } g(x) \le 0\,.$$

Man schreibt auch

$$F(x) = \min!\,, \qquad g(x) \le 0\,, \quad x \in \mathbb{R}^n\,. \tag{5.6}$$

Die Menge $M = \{x \in \mathbb{R}^n \,|\, g(x) \le 0\}$ 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 (5.6) lautet

$$L(x,\lambda) = F(x) + \lambda \cdot g(x)\,, \qquad x \in \mathbb{R}^n\,, \quad \lambda \in \mathbb{R}^m\,.$$

Für $\lambda \ge 0$ (d.h. die Komponenten λ_i von λ sind ≥ 0) ist $L(x,\lambda)$ konvex in x.

Definition 5.5 *Ein* Sattelpunkt *der Lagrange–Funktion L zu (5.6) ist ein Paar $(x^*, \lambda^*) \in \mathbb{R}^n \times \mathbb{R}^m$ mit $\lambda^* \geq 0$ und*

$$L(x^*, \lambda) \leq L(x^*, \lambda^*) \leq L(x, \lambda^*) \tag{5.7}$$

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 (5.6) ergibt sich durch den folgenden Satz:

Satz 5.2 *Ist (x^*, λ^*) Sattelpunkt von L, so ist x^* eine Lösung der Aufgabe (5.6). Sind überdies F und g differenzierbar, so gilt*

$$\nabla_x F(x^*) + \sum_{j=1}^{m} \lambda_j^* \, \nabla g_j(x^*) = 0\,. \tag{5.8}$$

Anmerkung: Offensichtlich entspricht (5.8) der aus der Analysis unter anderen Voraussetzungen bekannten Aussage zur Existenz von Lagrange–Multiplikatoren.

Beweis: (Der Satz von Hahn–Banach wird hier noch nicht verwendet.) (5.7) bedeutet

$$F(x^*) + \lambda \cdot g(x^*) \leq F(x^*) + \lambda^* \cdot g(x^*) \leq F(x) + \lambda^* \cdot g(x) \tag{5.9}$$

für $\lambda \in \mathbb{R}^n$, $\lambda \geq 0$ und $x \in \mathbb{R}^n$.
Aus der ersten Ungleichung folgt

$$(\lambda - \lambda^*) \cdot g(x^*) \leq 0\,, \qquad \lambda \geq 0\,, \quad \lambda \in \mathbb{R}^m\,.$$

Läßt man die Komponenten von λ gegen $+\infty$ sausen, 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 $\lambda = 0$, folgt

$$F(x^*) \leq F(x^*) + \lambda^* \cdot g(x^*)\,,$$

also $\lambda^* \cdot g(x^*) \geq 0$. Damit folgt

$$\lambda^* \cdot g(x^*) = 0\,. \tag{5.10}$$

Aus der zweiten Ungleichung in (5.9) folgt dann

$$F(x^*) \leq F(x) + \lambda^* \cdot g(x) \leq F(x)\,,$$

da $\lambda^* \geq 0,\ g(x) \leq 0$. x^* ist also tatsächlich Minimalstelle. □

Anmerkung: Die Beziehung (5.10) wird als besonderes Lemma festgehalten:

Lemma 5.3 („Gleichgewichtssatz“) *Sei (x^*, λ^*) Sattelpunkt der Lagrange–Funktion L von (5.6). Dann gilt*

$$\lambda^* \cdot g(x^*) = 0\,.$$

Die Bedingung(5.8) heißt auch „dritte Kuhn–Tucker–Bedingung“, die Bedingung (5.10) „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 (5.6) ein $\lambda^* \in \mathbb{R}^m,\ \lambda^* \geq 0$, zu konstruieren, so daß (x^*, λ^*) Sattelpunkt ist, benötigt man eine Zusatzbedingung, die manchmal als *Slater–Bedingung* in den Optimierungsbüchern herumgeistert:

$$\textit{Es gibt einen inneren Punkt der Menge } M\,. \tag{5.11}$$

Analytisch heißt dies: Es existiert $x_0 \in \mathbb{R}^n$ mit $g_j(x_0) < 0,\ j = 1, \ldots, m$.

Satz 5.3 (Kuhn–Tucker–Satz der konvexen Optimierung) *Sei $F\colon \mathbb{R}^n \to \mathbb{R}$ und $g\colon \mathbb{R}^n \to \mathbb{R}^m$ konvex und stetig. Die Menge $M = (g \leq 0)$ besitze einen inneren Punkt. Dann ist ein Punkt $x^* \in \mathbb{R}^n$ genau dann Lösung der Optimierungsaufgabe (5.6), wenn es ein $\lambda^* \in \mathbb{R}^m,\ \lambda^* \geq 0$ gibt, so daß (x^*, λ^*) Sattelpunkt der Lagrange–Funktion L zu (5.6) ist.*

Beweis: Die eine Richtung des Satzes ist bereits in Satz 5.2 festgehalten. Für die Umkehrung müssen wir den Multiplikator λ^* konstruieren. Hierzu betrachten wir die Mengen

$$A = \Big\{ y = (y_0, \ldots, y_m) \in \mathbb{R}^{m+1} \,\Big|\, y_0 \geq F(x),\ y_j \geq g_j(x),\ j = 1, \ldots, m \\ \text{für mindestens ein } x \Big\}$$

sowie

$$B = \Big\{ y = (y_0, \ldots, y_m) \in \mathbb{R}^{m+1} \,\Big|\, y_0 < F(x^*),\ y_j < 0,\ j = 1, \ldots, m \Big\}\,.$$

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 gäbe mit $F(x) \le y_0 < F(x^*)$, $g_j(x) \le y_j < 0$ für ein $x \in \mathbb{R}^n$, so daß x^* nicht minimal wäre. B ist offen (benötigt man beim Hahn–Banach im $\mathbb{R}^n$ nicht), A und B sind offensichtlich nicht leer). Damit läßt sich der Satz von Hahn–Banach über die Trennung konvexer Mengen anwenden und es gibt $\mu \in \mathbb{R}^m$, $\mu \ne 0$, so daß

$$\mu \cdot y > \mu \cdot z \qquad \text{für alle } y \in A,\ z \in B\,. \tag{5.12}$$

Läßt man die Komponenten von z nach $-\infty$ sausen (beachte die Definition von B), ergibt sich aus (5.12)

$$\mu \ge 0\,.$$

Aus (5.12) folgt für $y \in A, z \in \overline{B}$ (Abschließung von B)

$$\mu \cdot y \ge \mu \cdot z\,. \tag{5.13}$$

Aus (5.13) ergibt sich mit

$$z = \big(F(x^*), 0, \dots, 0\big) \in \overline{B}$$

und

$$y = \big(F(x),\ g(x)\big) \in A\,.$$

$$\mu_0\, F(x) + \sum_{j=1}^{m} \mu_j\, g_j(x) \ge \mu_0\, F(x^*), \qquad x \in \mathbb{R}^n\,. \tag{5.14}$$

Daraus entnimmt man $\mu_0 > 0$. Andernfalls wäre $\sum_{j=1}^{m} \mu_j\, g_j(x) \ge 0$, $x \in \mathbb{R}^n$, und insbesondere

$$\sum_{j=1}^{m} \mu_j\, g_j(x_0) \ge 0$$

mit dem Punkt x_0 aus der Slater–Bedingung $g(x_0) < 0$. Da $\mu_j \ge 0$ – wie schon bewiesen –, ergäbe sich ein Widerspruch. Wir setzen

$$\lambda^* = \frac{1}{\mu_0}\,(\mu_1, \dots, \mu_m)\,.$$

Es gilt $\lambda^* \ge 0$, und wegen (5.14)

$$F(x) + \lambda^* \cdot g(x) \ge F(x^*) \tag{5.15}$$

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

$$\lambda^* \cdot g(x^*) = 0\,. \tag{5.16}$$

Aus (5.15) und (5.16) ergibt sich die eine der beiden Ungleichungen der Sattelpunktsbedingung

$$F(x) + \lambda^* \cdot g(x) \geq F(x^*) + \lambda^* \cdot g(x^*)\,.$$

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. □

Eine weitere Anwendung des Satzes von Hahn-Banach sind

Alternativsätze für lineare Ungleichungssysteme

Satz 5.4 *Sei A eine $m \times n$–Matrix und $b \in \mathbb{R}^m$. Entweder besitzt*

$$Ax = b\,, \qquad x \geq 0 \tag{5.17}$$

eine Lösung $x \in \mathbb{R}^n$, oder

$$A^T y \geq 0\,, \qquad b \cdot y < 0 \tag{5.18}$$

besitzt eine Lösung $y \in \mathbb{R}^m$.

Beweis:

(i) (5.17) und (5.18) können nicht gleichzeitig lösbar sein, denn andernfalls existierte $x \in \mathbb{R}^n$ und $y \in R^m$ mit

$$0 > b \cdot y = (Ax) \cdot y = x^T A^T y \geq 0\,.$$

(ii) Ist (5.17) *unlösbar*, so liegt b *nicht* in dem von den Spaltenvektoren $a^{(1)}, \ldots, a^{(n)} \in \mathbb{R}^m$ aufgespannten Kegel

$$K = \left\{ \sum_{j=1}^{n} \xi_j a^{(j)} \right\}.$$

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

$$K_b = \left\{\alpha y \middle|\, y \in B_\varepsilon(b)\,, \quad \alpha \in \mathbb{R}^+\right\}$$

gilt

$$\operatorname{int} K_b \cap K = \emptyset\,.$$

Da K konvex und abgeschlossen, $\operatorname{int} K_b$ konvex und offen, läßt sich der Satz von Hahn–Banach anwenden. Es gibt daher eine Hyperebene, welche K und $\operatorname{int} K_b$ trennt, d.h. es existiert $a \in \mathbb{R}^m$, $a \neq 0$ mit

$$a \cdot y < a \cdot z\,, \qquad y \in K\,, \quad z \in \operatorname{int} K_b\,. \tag{5.19}$$

Da $y = 0 \in K$ und $b \in \operatorname{int} K_b$, folgt

$$a \cdot b > 0\,. \tag{5.20}$$

Da $0 \in \overline{\operatorname{int} K_b} = K_b$, folgt

$$a \cdot y \leq 0\,, \qquad y \in K\,,$$

insbesondere

$$a \cdot a^{(j)} \leq 0\,, \qquad j = 1, \dots, m\,.$$

Dies bedeutet

$$A^T a \leq 0\,. \tag{5.21}$$

Mit $-a = y$ ergibt sich aus (5.20) und (5.21)

$$A^T y \geq 0\,, \qquad y \cdot b < 0\,,$$

d.h. eine Lösung von (5.18).

□

6 Adjungierte lineare Operatoren in Banachräumen

Es seien E und F Banachräume. Wir betrachten nicht notwendig beschränkte lineare Operatoren $A\colon D(A) \to F$ 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 *Nullraum* ist definiert durch:

$$N(A) := \{x \in E |\ Ax = 0\}$$

Der Vektorraum der beschränkten linearen Funktionale auf E bzw. F wird wieder mit E^* bzw. F^* bezeichnet.

Definition 6.1 (des adjungierten Operators A^*) *Sei $A\colon D(A) \subset E \to F$. 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 Abschließung auf ganz E zu einem Element $f \in E^*$ fortgesetzt werden kann:

$$\langle f, u\rangle = \varphi(u) = \langle v, Au\rangle \qquad u \in D(A)\,.$$

Man definiert*:*

$$A^* v = f\ .$$

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.
Der Operator $A^*\colon D(A^*) \subset F^* \to E^*$ heißt der zu A *adjungierte Operator*.

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 Definition:

Definition 6.2 *Sei B ein Banachraum. Das* Orthogonalkomplement *einer Menge $M \subset B$ ist die Menge*

$$M^{\perp} := \{f \in B^* | \langle f, m \rangle = 0 \quad \textit{für alle } m \in M\} .$$

Das Biorthogonalkomplement *$M^{\perp\perp}$ ist die Menge*

$$M^{\perp\perp} = \{u \in B | \langle f, u \rangle = 0 \textit{ für alle } f \in M^{\perp}\}.$$

Satz 6.1 *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 $f_j \to f,\ f_j \in M^{\perp}$, folgt $0 = \langle f_j, m \rangle \to \langle f, m \rangle$, 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 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

$$\langle f, x \rangle < \alpha < \langle f, x_0 \rangle , \qquad x \in \overline{W} .$$

Da $\overline{W}$ ein linearer Teilraum ist, folgt $\langle f, x \rangle = 0$ für alle $x \in \overline{W}$, d.h. $f \in W^{\perp}$. 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. □

Lemma 6.1 *Sei $A \colon D(A) \subset E \to F$ linear, $\overline{D(A)} = E$. Dann ist A^* abgeschlossen, d.h. der Graph $G(A) := \big\{(v, A^*v) \big| \ v \in D(A^*)\big\}$ ist abgeschlossen in $F^* \times E^*$.*

Beweis: Sei $v_n \to v,\ A^*v_n \to f,\ v_n \in D(A^*)$, in F^* bzw. E^*. Wir müssen zeigen:

(i) $v \in D(A^*)$ und

(ii) $A^*v = f$.

Es gilt:

$$\langle v_n, Au \rangle = \langle A^*v_n, u \rangle , \quad u \in D(A) ,$$

nach Grenzübergang

$$\langle v, Au \rangle = \langle f, u \rangle \,,$$

woraus man $v \in D(A^*)$ und $A^* v = f$ abliest. □

Satz 6.2 *Sei $A\colon D(A) \subset E \to F$ 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) $N(A^*)^\perp \supset R(A)$

Anmerkung: „=“ hat man noch nicht sofort, es könnten ja weiter Elemente zu $R(A^*)$ orthogonal sein.

Beweis: Ein genialer Trick, den man in dem Buch „Analyse fonctionnelle“ von Brezis findet, macht den Beweis einfach:

(i) Wir betrachten den Graphen

$$G = G(A) = \big\{(u, Au) \in E \times F \big| u \in D(A)\big\}$$

und den Raum $L = E \times \{0\} \subset E \times F$. Die Beweisidee besteht darin, $(G+L)^\perp$ auszurechnen:

$$\begin{aligned}(G+L)^\perp &= \{(u+v, Au) \in E \times F \mid u \in D(A),\ v \in E\big\}^\perp = \\ &= \big\{(w, Au) \in E \times F \bigm| u \in D(A),\ w \in E\big\}^\perp = \\ &= \big\{(0, f^*) \in E^* \times F^* \big| \ f^* \perp Au\big\} = \\ &= \{0\} \times R(A)^\perp \,.\end{aligned}$$

Andererseits gilt ganz allgemein für die Summe zweier Räume

$$(G+L)^\perp = G^\perp \cap L^\perp,$$

denn

$$(G+L)^\perp = \big\{z + y \big| z \in G\,,\ y \in L\big\}^\perp = \big\{\varphi^* \big| \langle \varphi^*, z\rangle + \langle \varphi^*, y \rangle = 0\,, \quad z \in G,\ y \in L\big\}\,.$$

Setzt man z bzw. $y = 0$, erhält man, daß das linear beschränkte Funktional φ^* 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

$$\begin{aligned}\{0\} \times R(A)^\perp &= G^\perp \cap L^\perp = \\ &= \Big\{(e^*, f^*) \in E^* \times F^* \Big| \begin{array}{l}\langle e^*, u\rangle + \langle f^*, Au\rangle = 0 \text{ und} \\ \langle e^*, v\rangle + \langle f^*, 0\rangle = 0\,, \quad u \in D(A),\ v \in E\end{array}\Big\}\end{aligned}$$

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

$$\Big\{(0, f^*) \in E^* \times F^* \Big| \langle f^*, Au\rangle = 0\,, \quad u \in D(A)\Big\}$$

und damit gleich

$$\Big\{(0, f^*) \Big| A^* f^* = 0\Big\} = \{0\} \times N(A^*)\,.$$

Damit erhalten wir

$$R(A)^\perp = N(A^*)\,. \tag{6.1}$$

(i) = (6.1) ist damit bewiesen.

Anmerkung: Die folgende simple Überlegung reicht nicht zum Nachweis von (6.1). 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 $\big(G^\perp + L^\perp\big)^\perp$, G und L wie im Beweis für (i): Einerseits ist

$$\begin{aligned}\big(G^\perp + L^\perp\big)^\perp &= G^{\perp\perp} \cap L^{\perp\perp} = G \cap L = \\ &= \Big\{(u, Au) \Big| u \in D(A)\Big\} \cap \Big\{(v, 0) \Big| v \in E\Big\} = \\ &= \Big\{(u, Au) \Big| Au = 0,\ u \in E\Big\} = \\ &= N(A) \times \{0\},\end{aligned} \tag{6.2}$$

da G und L abgeschlossen sind.

Andererseits ist $L^\perp = \big\{(0,h^*) \in E^* \times F^* \big|\ h^* \in F^*\big\}$.

$G^\perp + L^\perp = \big\{(e^*, f^*) + (0,h^*) \in E^* \times F^* \big| \langle e^*, u\rangle + \langle f^*, Au\rangle = 0, u \in E, h^* \in F^*\big\}$

Da alle $h^* \in F^*$ zugelassen sind, folgt

$G^\perp + L^\perp = \big\{(e^*, g^*) \big|\ g^* \in F^*\,,\ \text{zu } e^* \text{ existiert } f^*, \text{so daß } \langle e^*, u\rangle + \langle f^*, Au\rangle = 0\big\}$.

Wenn für ein Paar (e^*, f^*) die Gleichung $\langle e^*, u\rangle + \langle f^*, Au\rangle = 0$ für alle u gilt, so ist $f^* \in D(A^*)$ und

$$\langle e^* + A^* f^*, u\rangle = 0\,, \qquad u \in D(A)\,.$$

Daraus folgt $e^* = -A^* f^*$. Umgekehrt impliziert die letzte Gleichung, daß

$$\langle e^*, u\rangle + \langle f^*, Au\rangle = 0\,.$$

Daraus folgt

$$\begin{aligned} G^\perp + L^\perp &= \big\{(-A^* f^*, g^*) \big|\ f^* \in D(A)\,,\ g^* \in F^*\big\} \\ &= R(A^*) \times F^*. \end{aligned}$$

Anwendung der $\perp$-Operation ergibt

$$\big(G^\perp + L^\perp\big)^\perp = R(A^*)^\perp \times \{0\}\,.$$

Zusammen mit (6.2) folgt

$$N(A) \times \{0\} = R(A^*)^\perp \times \{0\}$$

und damit $N(A) = R(A^*)^\perp$.

(iii) Aus (i) folgt

$$N(A^*)^\perp = R(A)^{\perp\perp}\,,$$

allgemein gilt jedoch für einen linearen Teilraum V, daß

$$V^{\perp\perp} = \overline{V}\,, \tag{6.3}$$

woraus (iii) folgt, wenn man $V = R(A)$ wählt. In der Tat, Orthogonalkomplemente sind immer abgeschlossen; auf der anderen Seite ist

$$(V^\perp)^\perp = \{v \in E |\ \langle f^*, v\rangle = 0\,, f^* \in V^\perp\}\,.$$

Wenn $f^* \in V^\perp$, ist $\langle f^*, u\rangle = 0$ für alle $u \in V$, d.h. die $(V^\perp)^\perp$ definierende Menge enthält V, und wegen der Abgeschlossenheit von $V^{\perp\perp}$ ergibt sich (6.3). Ähnlich schließt man für (iv).

$\square$

Wie im Fall des Hilbertraumes läßt sich auch die Aussage des vorigen Satzes verschärfen, wenn man die Abgeschlossenheit des Bildbereiches $R(A)$ fordert:

Satz 6.3 *Sei A: $D(A) \subset E \to F$ 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) $R(A^*) = N(A)^\perp$.

Anmerkung: (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 6.2, (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. $\square$

Beispiel: Sei $I = [\alpha, \beta]$ ein beschränktes Intervall des $\mathbb{R}^n$ und $E = L^p(I)$, $F = L^p(I)$, mit $p \in (1, \infty)$. Sei $D(A) = H^{2,p}(I) \cap H_0^{1,p}(I)$. Hierbei ist $H^{2,p}$ der Sobolev-Raum der L^p-Funktionen mit zweiten verallgemeinerten Ableitungen in L^p, oder, äquivalent, die Menge aller C^1-Funktionen, deren erste Ableitung absolutstetig in I ist und deren zweite - fast überall existierende Ableitung - in L^p liegt. $H_0^{1,p}(I)$ besteht aus den $H^{1,p}$-Funktionen, die auf dem Rand von I, als on α und β verschwinden. (In einer Dimension ist die punktweise Restriktion einer $H^{1,p}$-Funktion erlaubt.)
Offensichtlich ist $D(A)$ dicht in L^p (bezüglich der L^p-Norm). Der Operator A : $D(A) \to L^p$ sei definiert durch

$$Au = -u'' + cu$$

mit einer vorgegebenen L^∞-Funktion c. Da für $u \in H^{2,p}$, $u \in C(I)$, 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 $Au_m \to f$ in L^p, $u_m \to u$ 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 $Au_m \to Au$. Wir können damit aussagen, daß der Nullraum 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 Nullraum von A^* bestimmen. Wir wissen, daß

$$A^* : D(A^*) \subset L^q(I) \to L^q(I)\,, \qquad q = \frac{p}{p-1}\,,$$

abbildet und daß $N(A^*)$ aus allen Elementen $g \in L^q$ besteht, für die

$$\int\limits_\alpha^\beta g(-u'' + cu)\,dt = 0 \qquad \text{für alle } u \in D(A)\,.$$

Ähnlich wie im Hilbert-Raum-Fall schließt man damit zunächst, daß g zweite verallgemeinerte Ableitungen hat und daß $g'' \in L^q,\ q = p(p-1)$. Es folgt dann weiter

$$-g'' + cg = 0 \qquad \text{für in } I \tag{6.4}$$

und

$$g(\alpha) = g(\beta) = 0\,. \tag{6.5}$$

Der Nullraum von A^* besteht also aus den Lösungen des Randwertproblems (6.4), (6.5), und ist das Randwertproblem

$$-u'' + cu = f \in L^p\,, \qquad u(\alpha) = \alpha(\beta) = 0 \tag{6.6}$$

lösbar, so muß f die Orthogonalitätsrelation („Paarung von (L^q, L^p)“)

$$\int\limits_\alpha^\beta fg\,dx = 0$$

erfüllen. Daß dies auch notwendig für die Lösbarkeit von (6.6) ist, folgt aus Satz 6.3,(iii), wofür man die Abgeschlossenheit von A beweisen muß. Darüberhinaus läßt sich zeigen, daß $R(A)$ abgeschlossen ist. Der Nachweis geschieht ähnlich wie im Hilbertraum, indem A als Summe eines koerzitiven und eines kompakten Operators geschrieben wird.

7 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 arithmethische Mittelung) mit dem gleichen Limes bilden kann. U.a. 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, eine Folge (u_m), $u_m \in B$ heißt *schwach konvergent*, $u_m \rightharpoonup u$, genau dann, wenn

$$\langle f, u_m \rangle \to \langle f, u \rangle$$

für alle $f \in B^*$.

Satz 7.1 *Eine konvexe Menge C eines Banachraumes B ist genau dann bezüglich der starken Topologie abgeschlossen, wenn sie bezüglich der schwachen Topologie abgeschlossen ist.*

Beweis: Die Richtung

$$\text{„}C \text{ ist schwach abgeschlossen“} \Rightarrow \text{„}C \text{ ist stark abgeschlossen“}$$

ist trivial, da die schwache Topologie gröber ist als die starke. Sei also C stark abgeschlossen. Wir müssen zeigen, daß das Komplement 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

$$\langle f, x_0 \rangle < \alpha < \langle f, y \rangle , \qquad y \in C .$$

Die Menge

$$V = \{x \in B | \ \langle f, x \rangle < \alpha\}$$

ist aber eine bezüglich der schwachen Topologie offene Menge, somit eine offene, mit C disjunkte Umgebung von x_0. Damit ist C offen bezüglich der schwachen Topologie.
□

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), $v_j \in B$, von konvexen Linearkombinationen*

$$v_j = \sum_{k=1}^{N_j} c_k^j u_{m_{kj}}\,,\ c_k^j \geq 0\,,\ \sum_{k=1}^{N_j} c_k^j = 1\,,\ m_{kj} \in \mathbb{N},\ N_j \in \mathbb{N}, \tag{7.1}$$

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 die Abschließung $\overline{C}$ bezüglich der starken Topologie. Ist nun $u_m \rightharpoonup u$ schwach, so muß wegen Satz 7.1 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 (7.1).
□

Beispiel: *Jedes konvexe, koerzitive, unterhalbstetige Funktional f auf einer abgeschlossenen konvexen Menge C eines reflexiven Banachraumes besitzt einen Minimierer.* Die Begriffe „abgeschlossen“ und „unterhalbstetig“ sind bezüglich der starken Topologie gemeint.

Beweis: Sei (u_m), $u_m \in C$ Minimalfolge von f, d.h.

$$f(u_m) \to \inf_C f \qquad (m \to \infty)\,.$$

Dann gibt es eine schwach konvergente Teilfolge $(u_m)_{m\in\Lambda}$, $u_m \rightharpoonup u$, $m \in \Lambda$, und wegen Satz 7.1 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

$$f(v_j) \leq \sum_{k=1}^{N_j} c_k^j\, f(u_{m_{kj}}) \leq \sum_{k=1}^{N_j} c_k^j \left(\inf_C f + \varepsilon\right) = \inf_C f + \varepsilon\,.$$

(Die m_{kj} können so gewählt werden, daß $f(u_{m_{kj}}) = \inf_C f + e$, $|e| < \varepsilon$, indem man m_{1j} genügend groß wählt.) Da f unterhalbstetig ist und $v_j \to u$ stark, gilt

$$f(u) \leq \lim\inf_{j\to\infty} f(v_j) \leq \inf_C f + \varepsilon\,.$$

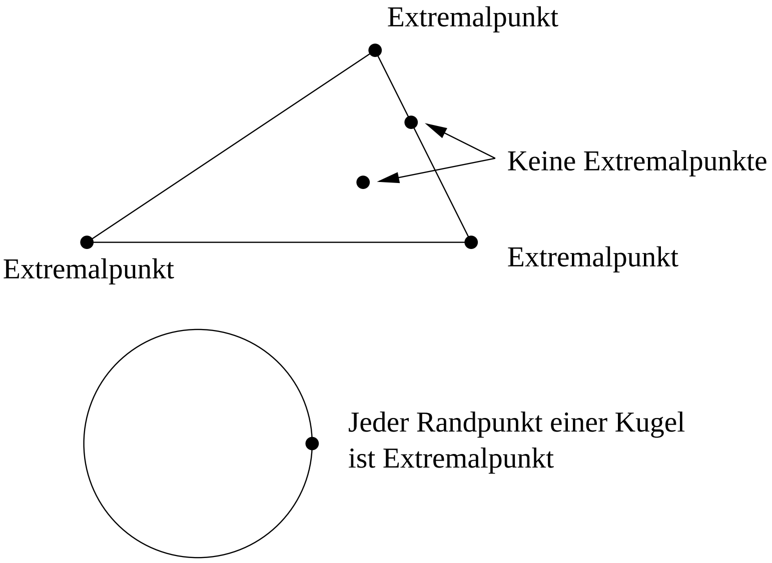
Grenzübergang $\varepsilon \to 0$ ergibt $f(u) = \inf_C f$, d.h. das Minimum wird angenommen. □

8 Der Satz von Krein–Milman und der Satz von Ljapunov

Der Satz von Krein–Milman stellt die Existenz von Extremalpunkten kompakter (auch schwach kompakter) konvexer Mengen sicher. Eine berühmte Anwendung ist der Satz von Ljapunov über die Konvexität des Wertebereiches von Vektormaßen.

Definition 8.1 *Es sei K eine Teilmenge eines normierten Raumes B. Ein Punkt $x \in K$ heißt* Extremalpunkt von K, *wenn aus der Darstellung $x = \alpha y + (1-\alpha)z,\ y, z \in K,\ 0 < \alpha < 1$ die Gleichheit $x = y = z$ folgt.*

Mit anderen Worten: Ist x ein Extremalpunkt, so kann x nicht echte konvexe Linearkombination verschiedener Elemente aus K sein.



Es gibt noch den Begriff des *exponierten Punktes*: Ein Punkt $x \in K$ heißt *exponiert*, wenn x auf einer Stützhyperebene S zu K liegt und wenn x der einzige Punkt von K auf S ist. Jeder Extremalpunkt ist exponierter Punkt, aber nicht umgekehrt.

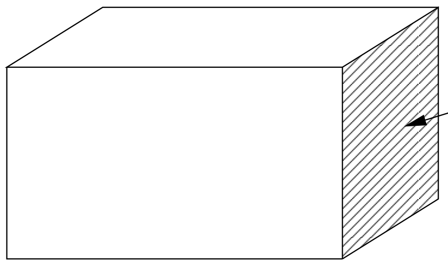
Der Satz von Krein–Milman in reflexiven Banachräumen:

Satz 8.1 *Jede konvexe, abgeschlossene und beschränkte Teilmenge M eines reflexiven Banachraumes B besitzt mindestens einen Extremalpunkt.*

Beweis: Wir nennen eine nichtleere Teilmenge A von M eine *extremale Teilmenge* von M, wenn die Inklusion

$$\alpha x + (1-\alpha)y \in A\,, \quad x, y \in M,\ \alpha \in]0,1[\text{ fest},$$

die Inklusion $x, y \in A$ impliziert.
Extremale Teilmengen werden auch als *„Wände“* bezeichnet.



M selbst ist extremale Teilmenge von M, einpunktige extremale Teilmengen von M sind Extremalpunkte.

Es sei $\mathcal{F}$ die (nichtleere!) Familie aller abgeschlossenen nichtleeren extremalen Teilmengen von M, die durch die mengentheoretische Inklusion mit einer Halbordnung versehen ist. Wir werden zeigen, daß $\mathcal{F}$ bezüglich dieser Halbordnung mindestens ein minimales Element M_0 besitzt, welches aus genau einem Punkt, einem Extremalpunkt, besteht. Hierzu benutzen wir das *Lemma von Zorn*, welches die Existenz eines minimalen Elementes bezüglich einer Halbordnung sicherstellt, sofern man weiß, daß jede maximale *total geordnete* Unterfamilie der zugrundeliegenden Familie (in unserem Falle $\mathcal{F}$) ein minimales Element hat. Eine Unterfamilie U von $\mathcal{F}$ heißt *„linear“* oder *„total geordnet“*, wenn für $M_1, M_2 \in U$ entweder $M_1 \subseteq M_2$ oder $M_2 \subseteq M_1$ gilt. Sie heißt *maximal*, wenn es keine linear geordnete, U echt enthaltende Unterfamilie $\tilde{U} \subset \mathcal{F}$ gibt.

Daß total geordnete Unterfamilien von $\mathcal{F}$ in unserem Fall ein minimales Element haben, folgt aus der schwachen Kompaktheit der Menge M: Ist U eine linear geordnete Unterfamilie von $\mathcal{F}$ und gilt

$$\bigcap \{M_\iota | M_\iota \in U\} = \emptyset\,,$$

so gilt dies wegen der schwachen Kompaktheit von M auch für endlich viele M_i, d.h.

$$\bigcap \{M_\iota |\ \iota \in \Lambda\} = \emptyset, \qquad |\Lambda| < \infty\,.$$

Dies ist ein Widerspruch, da die M_ι linear geordnet und nichtleer sind. (Wir haben früher bewiesen, daß beschränkte, abgeschlossene, konvexe Mengen in reflexiven Räumen schwach folgenkompakt sind. Man kann noch folgern, s. etwa Köthe, Topologische Lineare Räume, daß sie auch überdeckungskompakt sind, und wir haben soeben eine hierzu äquivalente Formulierung benutzt.)

Es gilt also $N_0 = \bigcap \{M_\iota |\ M_\iota \in U\} \neq \emptyset$, und man überlegt sich sofort, daß N_0 selbst abgeschlossene, nichtleere, extremale Teilmenge von M ist. Wegen der Maximalität von U ist $N_0 \in U$ und N_0 ist das minimale Element von U. Nach dem Lemma von Zorn gibt es daher mindestens ein minimales Element A_0 in $\mathcal{F}$. Angenommen, A_0 enthielte zwei verschiedene Punkte x_1, x_2. Dann gibt es ein $f \in B^*$ mit $\langle f, x_1\rangle = 0$, $\langle f, x_2\rangle \neq 0$. Setze $\beta = \inf\big\{\langle f, x\rangle \big|\ x \in A_0\big\}$. Dann ist die Menge $A_1 = \big\{x \in A_0 \big|\ \langle f, x\rangle = \beta\big\}$ eine nichtleere abgeschlossene konvexe eigentliche Teilmenge von A_0. Wir zeigen, daß A_1 *extremale* Teilmenge von M ist: Es seien $x, y \in M$, $0 < \alpha < 1$, so daß $\alpha x + (1-\alpha)\, y \in A_1$. Da $A_1 \subset A_0$ und A_0 extremale Teilmenge von M ist, folgt zunächst $x, y \in A_0$. Nach Definition von β gilt $\langle f, x\rangle \geq \beta$, $\langle f, y\rangle \geq \beta$. Angenommen, für mindestens eines der Elemente x, y gälte $\langle f, x\rangle > \beta$ bzw. $\langle f, y\rangle > \beta$, so folgte

$$\big\langle f, \alpha x + (1-\alpha) y\big\rangle = \alpha\langle f, x\rangle + (1-\alpha)\langle f, y\rangle > \beta\,,$$

d.h. $\alpha x + (1-\alpha)y$ wäre nicht in A_1 enthalten. Dies ist ein Widerspruch, und es muß gelten $\langle f, x\rangle = \beta$, $\langle f, y\rangle = \beta$, d.h. $x, y \in A_1$ und A_1 muß extremal sein. Dies widerspricht der Minimalität von A_0, und die Annahme, A_0 bestehe aus mehr als einem Punkt, ist damit zum Widerspruch geführt. □

Eine bekannte Anwendung des Satzes von Krein–Milman ist der Satz von Ljapunov über den Wertebereich von Vektormaßen, den wir in einer einfachen Situation beweisen werden: Es sei $[a, b]$ ein Intervall des $\mathbb{R}^n$ und $f\colon [a, b] \to \mathbb{R}^n$ meßbar und beschränkt. Wir setzen

$$\begin{aligned} M &= \left\{ \int_a^b f u\, dt \,\middle|\, 0 \leq u \leq 1\,, \quad u \in L^\infty[a, b] \right\} \qquad \text{und} \\ M_0 &= \left\{ \int_E f\, dt \,\middle|\, E \subset [a, b] \text{ meßbar } \right\}. \end{aligned}$$

Die Ungleichung $0 \leq u \leq 1$ für $u \in L^\infty$ soll bis auf eine Ausnahmemenge vom Maß Null im Intervall $[a, b]$ gelten. Offenbar sind M und M_0 Teilmengen des $\mathbb{R}^1$ und es gilt $M_0 \subseteq M$, da jedes Element $\int_E f\, dt$ von M_0 in der Form $\int_a^b f u\, dt$, $u = \begin{cases} 1 & \text{auf } E \\ 0 & \text{auf } [a, b]\backslash E \end{cases}$, geschrieben werden kann. Solche Mengen treten beim Studium der erreichbaren Mengen von Steuerungssystemen auf.

Lemma 8.1 *Die Menge M ist eine abgeschlossene und konvexe Teilmenge des $\mathbb{R}^n$.*

Dies folgt aus den Ergebnissen von Kapitel 7.

Mit Hilfe des Satzes von Krein–Milman werden wir zeigen:

Satz 8.2 $M = M_0$.

Da M abgeschlossen und konvex ist, folgt als Korollar der

Satz 8.3 (Satz von Ljapunov) *Der Wertebereich des Vektormaßes* $J(E) = \int\limits_E f\,dt$, $E \subset [a,b]$ *meßbar, ist abgeschlossen und konvex.*

Unter einem Vektormaß versteht man eine σ-additive Funktion eines Maßraumes in den $\mathbb{R}^n$.

Satz 8.3 ist die spezielle Form des folgenden allgemeinen Satzes von Ljapunov über Werte von nicht–atomaren Vektormaßen in endlichen Maßräumen, der auch in der Wahrscheinlichkeitstheorie/Statistik und anderen Gebieten (z.B. Theorie der konvexen Körper) eine Rolle spielt. Wir geben kurz eine Erläuterung der Begriffe:
Eine Familie $\sum$ von Teilmengen E einer Menge S heißt eine *Algebra von Teilmengen*, wenn

(i) $S \in \sum$,

(ii) Mit $E \in \sum$ ist auch $S - E \in \sum$,

(iii) Ist $E_1, E_2 \in \sum$, so ist auch $E_1 \cap E_2$ und $E_1 \cup E_2 \in \sum$.

Eine Algebra $\sum$ von Teilmengen heißt *σ–Algebra*, wenn für jede abzählbare Familie $\{E_i \in \sum \mid i = 1, 2, \dots\}$ die Inklusionen $\bigcap\limits_{i=1}^{\infty} E_i \in \sum$ und $\bigcup\limits_{i=1}^{\infty} E_i \in \sum$ folgen.

Eine Abbildung $\mu\colon \sum \to \mathbb{R}^n$ heißt *additiv*, wenn für jedes Paar $E_1, E_2 \in \sum$ mit $E_1 \cap E_2 = \emptyset$ die Gleichung

$$\mu(E_1 \cup E_2) = \mu(E_1) + \mu(E_2)$$

gilt.

Eine Abbildung $\mu\colon \sum \to \mathbb{R}^n$ heißt *σ–additiv*, wenn für jede abzählbare Familie $\{E_i \in \sum \mid i = 1, 2, \dots\}$ mit $E_i \cap E_k = \emptyset$, $i \neq k$, die Gleichung

$$\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mu(E_i)$$

gilt. μ wird dann auch „Vektormaß" genannt.
Das Tripel $(\mu, S, \sum)$ nennt man auch einen „Maßraum".

Ein Vektormaß auf einer σ–Algebra heißt nicht–atomar, wenn zu jedem $E \in \sum$ mit $\mu(E) \neq 0$ ein $E' \subset E, E' \in \sum$, existiert mit $0 \neq |\mu(E')| < |\mu(E)|$. Mit diesen Bezeichnungen lautet der allgemeine Satz von Ljapunov

Satz 8.4 *Der Wertebereich eines nicht–atomaren Vektormaßes auf einem Maßraum ist abgeschlossen und konvex.*

Eine Beweisskizze findet sich in dem Buch von Lee–Markus. Wir werden hier nur die Version mit Lebesgue–meßbaren Mengen beweisen.

Beweis von Satz 8.2 Wir präsentieren einen „einfachen“ Beweis, dessen Hauptidee auf Lindenstrauß zurückgeht, vgl. das Buch von Hermes–Lasalle, Time optimal control. Für jedes $u \in \mathcal{K} := \big\{v \in L^\infty[a,b] \,\big|\, 0 \le v \le 1 \text{ fast überall}\big\}$ sei $T\colon \mathcal{K} \to \mathbb{R}^n$ definiert durch

$$Tu := \int_a^b f u \, dt\,.$$

Es sei nun $z \in M$, d.h. es gibt ein $u \in \mathcal{K}$ mit $Tu = z$. Wir müssen zeigen, daß es eine meßbare Menge $E \subset [a,b]$ gibt, so daß $T\chi_E = \int_a^b f\chi_E \, dt = \int_E f \, dt = z$ gilt. Hierbei ist χ_E die sogenannte charakteristische Funktion von E, d.h. $\chi_E = \begin{cases} 1 & \text{auf } E \\ 0 & \text{sonst} \end{cases}$. Es sei $C = \{v \in \mathcal{K} | \; Tv = z\}$. Die Menge C ist nicht leer, da $u \in C$, und konvex wegen der Linearität von T. Fassen wir C als eine Teilmenge von $L^2[a,b]$ auf, so ist die Menge C abgeschlossen in der schwachen Topologie von L^2, s. Satz (7.1). Nach dem Satz von Krein–Milman (Satz 8.1) hat die Menge C einen Extremalpunkt. Daraus folgt die Existenz einer Funktion $u^* \in C$ mit der Eigenschaft „u^* liegt nicht im Inneren irgendeines ganz in C liegenden Intervalls“, d.h. aus $u^* = \alpha u_1 + (1-\alpha)u_2, 0 < \alpha < 1, u_1, u_2 \in C$ folgt $u^* = u_1$ oder $u^* = u_2$.

Wir zeigen, daß der obige Extremalpunkt u^* von C die charakteristische Funktion einer Menge E ist. Angenommen, dies wäre nicht der Fall. Dann gibt es eine meßbare Menge $G \subset [a,b]$ mit $|G| \neq 0$, so daß $0 < u^* < 1$ auf G gilt. Daraus folgt die Existenz einer Zahl $\varepsilon > 0$ und einer meßbaren Menge G_0 mit $|G_0| \neq 0$, so daß $\varepsilon < u^* < 1 - \varepsilon$ auf G_0 gilt. Zum Beweis sei $H_j = \big\{t \in [a,b] \big| \, \frac{1}{j} < u^*(t) < 1 - \frac{1}{j}\big\}$. Es gilt dann $\Big|\bigcup_{j=1}^\infty H_j\Big| \neq 0$, und damit existiert ein j_0 mit $|H_{j_0}| \neq 0$. Setze $G_0 = H_{j_0}$, $\varepsilon = \frac{1}{j_0}$. Da $|G_0| \neq 0$, enthält G_0 $(n+1)$ disjunkte meßbare Teilmengen G_i, $i = 1, \dots, n+1$, deren Maß nicht verschwindet. (Hier wird ausgenutzt, daß das Lebesgue–Maß nicht–atomar ist.)

Da $(n+1)$ Vektoren im $\mathbb{R}^n$ stets linear abhängig sind, gibt es Zahlen λ_i mit $\sum\limits_{i=1}^{n+1} |\lambda_i| \neq 0$ und

$$\sum_{i=1}^{n+1} \lambda_i \int\limits_{G_i} f \, dt = 0 \, .$$

Wir dürfen zusätzlich annehmen, daß $|\lambda_i| < \varepsilon, \; i = 1, \dots, n+1$. Setzt man

$$\chi_i = \begin{cases} 1 & \text{auf } G_i \\ 0 & \text{auf } [a,b] - G_i \end{cases} , \qquad i = 1, \dots, n+1,$$

so gilt für $\alpha \in \mathbb{R}$

$$T\Big(u^* + \alpha \sum_{i=1}^{n+1} \lambda_i \chi_i\Big) = z + \alpha \sum_{i=1}^{n+1} \lambda_i \int\limits_a^b f \chi_i \, dt = z \, ,$$

und für $\alpha \in [-1, 1]$

$$0 \leq u^* + \alpha \sum_{i=1}^{n+1} \lambda_i \chi_i \leq 1 \text{ auf } [a,b] \, ,$$

denn $\Big|\sum\limits_{i=1}^{n+1} \lambda_i \chi_i\Big| \leq \varepsilon$ auf $\bigcup\limits_{i=1}^{n+1} G_i = G_0$, d.h. auf der Menge, wo $\varepsilon < u^* < 1 - \varepsilon$. Damit ist gezeigt, daß $u^* + \alpha \sum\limits_{i=1}^{n+1} \lambda_i \chi_i \in C$ gilt, d.h. u^* ist innerer Punkt eines in C liegenden Intervalls. u^* kann daher kein Extremalpunkt sein und die Annahme, daß u^* verschieden von 0 oder 1 auf einer Menge nicht verschwindenden Maßes ist, ist zum Widerspruch geführt. u^* ist daher die charakteristische Funktion einer meßbaren Menge E, und der Satz ist bewiesen. □

Teil III

Anhang

A Fixpunktsätze

Die nichtlineare Funktionalanalysis beschäftigt sich mit nichtlinearen Abbildungen zwischen unendlich–dimensionalen topologischen linearen Räumen sowie zugehörigen Gleichungen. U.a. werden Aussagen zur Lösbarkeit nichtlinearer Gleichungen und Eigenschaften ihrer Lösungen gegeben. Von besonderem Interesse sind Abbildungen, die aus Differential– und Integralgleichungen und zugehörigen Funktionenräumen herrühren.

Säulen der nichtlinearen Funktionalanalysis sind *Fixpunktsätze*; besonders wichtig und bekannt sind:

(i) der Banachsche Fixpunktsatz

(ii) der Fixpunktsatz von Schauder

(iii) die Theorie der monotonen Operatoren

(iv) die Leray–Schauder–Theorie

Diese vier Prinzipien wollen wir kurz vorstellen:

1. Der Banachsche Fixpunktsatz

Diesen Satz lernt man im Fall des $\mathbb{R}^n$ bereits in den Anfängervorlesungen kennen.

Satz A.1 *Eine kontrahierende Abbildung eines Banachraumes B in sich besitzt einen Fixpunkt. Eine Abbildung $A\colon B \to B$ heißt* kontrahierend, *wenn*

$$\|Ax - Ay\| \leq q\|x - y\| , \qquad x, y \in T \tag{A.1}$$

mit einer Konstanten $q < 1$.

Satz A.1 gilt analog in vollständigen Räumen, eine Variante besteht darin, daß T eine Teilmenge C – zumeist eine Kugel oder konvexe Menge – in sich selbst abbildet und die Kontraktionsbedingung dafür nur für die Elemente x, $y \in C$ gefordert wird.

Der Beweis von Satz A.1 beruht darauf, indem man sich den gesuchten Fixpunkt als Limes der Elemente (=iterierten Vektoren) $A^m x_0, x_0$ beliebige „Startwerte“, konstruiert.

Eine klassische Anwendung des Banachschen Fixpunktsatzes sind bekanntlich Anfangswertprobleme für gewöhnliche Differentialgleichungen

$$x = f(.,x)\,, \quad x(t_0) = x_0\,,$$

mit einer Lipschitz–Bedingung für die stetigen Funktion f bezüglich des zweiten Arguments.

Man wählt als Grundraum den Raum $B = C([t_0, t_0 + \delta], \mathbb{R}^n)$. Die Abbildung A ist definiert durch

$$(Ax)(t) = x_0 + \int\limits_{t_0}^{t} f\big(s, x(s)\big)\, ds \tag{A.2}$$

Durch die Kleinheit von δ wird die Kontraktionsbedingung (A.1) erreicht.

Erfüllt f eine globale Lipschitzbedingung

$$|f(t,\xi) - f(t,\eta)| \le L|\xi - \eta|\,,$$

so läßt sich die Existenz einer Lösung des Anfangswertproblems auf einem *großen* Intervall $[t_0, T]$ durch geschickte Wahl der Norm erreichen.

Man definiert die Abbildung A, auf die der Banachsche Fixpunktsatz angewandt werden soll, wie in (A.2). Als Norm wählt man

$$\|x\| = \sup_{t\in[t_0,T]} \left|e^{-\lambda t} x(t)\right|$$

mit einem genügend großen Parameter λ. Die Kontraktionsbedingung wird folgendermaßen bewiesen:

$$
\begin{aligned}
\|Ax - Ay\| &= \sup_{t\in[t_0,T]} \left| e^{-\lambda,t} \int\limits_{t_0}^{t} (f(s,x(s) - f(s,y(s)))\, ds \right| = \\
&= \sup_{t\in[t_0,T]} \left| e^{-\lambda t} \int\limits_{t_0}^{t} e^{+\lambda s} e^{-\lambda s} \big(f(s,x(s)) - f(s,y(s)) \big)\, ds \right| \\
&\leq \sup_{t\in[t_0,T]} \left| e^{-\lambda t} \int\limits_{t_0}^{t} e^{+\lambda s} \|f(.,x) - f(.,y)\|\, ds \right| \\
&\leq \sup_{t\in[t_0,T]} \left| e^{-\lambda t} \int\limits_{t_0}^{t} e^{+\lambda s}\, ds \right| L\|x-y\| \\
&= \sup_{t\in[t_0,T]} \left| e^{-\lambda t} \frac{e^{+\lambda t} - e^{+\lambda t_0}}{\lambda} \right| L\|x-y\| \\
&= \sup_{t\in[t_0,T]} \frac{|1 - e^{-\lambda(t-t_0)}|}{\lambda} L\|x-y\| \\
&= \frac{1}{\lambda}\big(1 - e^{-\lambda(T-t_0)}\big) L\|x-y\|
\end{aligned}
$$

Wählt man etwa $\lambda > 2L$, erhält man die Kontraktionsbedingung $q = \frac{1}{\lambda}\big(1 - e^{-\lambda(T-t_0)}\big)L < 1$.

Die Ausführung zeigt, wie man die Norm durch geschickte Wahl dem Problem anpassen kann.

2. Der Schaudersche Fixpunktsatz

Satz A.2 *Es sei A eine stetige Abbildung einer kompakten konvexen Menge C eines Banachraumes B in sich. Dann besitzt A einen Fixpunkt.*

Alternativ läßt sich voraussetzen, daß C eine abgeschlossene konvexe – nicht notwendig kompakte – Teilmenge von B ist und C stetig und „kompakt" ist, d.h. beschränkte Folgen in relativ kompakte überführt.

Eine bekannte Anwendung ist wiederum die in (A.2) definierte Abbildung A, welche die lokale Lösbarkeit von Anfangswertproblemen für gewöhnliche Differentialgleichungen bei stetigem f – ohne Lipschitzbedingung – sichert.

Eine andere Standardanwendung ist das Randwertproblem

$$\begin{aligned} -\Delta u &= f(u) \quad && \text{a.e. in } \Omega \subset \mathbb{R}^n \\ u &= 0 && \text{auf } \partial\Omega \end{aligned} \tag{A.3}$$

Hierbei ist Ω eine beschränkte Menge des $\mathbb{R}^n$ und f eine vorgegebene, auf ganz $\mathbb{R}$ beschränkte stetige Funktion. Das Problem (A.3) läßt sich in zahlreichen Funktionenräumen definieren. Wir verwenden hier den Raum $L^2(\Omega)$ und benutzen, daß für Gebiete mit glattem Rand die Abbildung $-\Delta$ eine umkehrbar-eindeutige stetige Abbildung von $H^{2,2}(\Omega) \cap H_0^{1,2}(\Omega)$ nach $L^2(\Omega)$ ist.

Problem (A.3) läßt sich damit als Fixpunktgleichung

$$u = (-\Delta)^{-1} f(u)$$

schreiben. Da f beschränkt ist, gilt $\|f(u)\|_{L^2} \leq K$ und $\| - \Delta^{-1} f(u)\|_{H^{2,2}} \leq K$ mit einer Konstanten K für alle $u \in L^2$. Die Abbildung $A = -\Delta^{-1} f(u)$ bildet daher die Kugel B_K aus L^2 in sich ab und führt beschränkte Folgen aus L^2 in $H^{2,2}$-beschränkte Folgen über. Da in $H^{2,2}$ beschränkte Folgen wegen des Rellichschen Einbettungssatzes relativ kompakt in L^2 sind, führt A beschränkte Folgen in relativ kompakte über. Ferner ist A stetig. Nach dem Schauderschen Fixpunktsatz gibt es daher einen Fixpunkt von A und damit eine (sogenannte starke) Lösung von (A.3).

Der Satz von Schauder ist das unendlich–dimensionale Analogon des Brouwerschen Fixpunktsatzes. Der Beweis wird auf letzteren zurückgeführt, indem die Abbildung A mit Hilfe von Projektionen P_j durch „endlich dimensionale" Abbildungen $A_j = P_j A P_j$ approximiert wird.

Das Analogon des Schauderschen Satzes für lokalkonvexe Räume ist als Satz von Tychonoff bekannt. Auch für die Räume L^p, $0 < p < 1$, gibt es einen „Schauderschen Fixpunktsatz".

3. Die Theorie der Monotonen Operatoren

Ein Standardsatz ist der folgende:

Satz A.3 *Es sei A eine stetige Abbildung eines Hilberraumes in sich. A führe beschränkte Mengen in beschränkte über. Ferner sei A monoton, d.h. es gelte für alle $x, y \in H$*

$$(Ax - Ay, x - y) \geq 0\,,$$

und koerzitiv, d.h.

$$(Au, u) \to \infty \qquad \textit{für } \|u\| \to \infty \,.$$

Dann ist die Gleichung $Au = f$ für alle $f \in H$ lösbar.

Im endlich–dimensionalen Fall benötigt man die Monotoniebedingung *nicht*, sie wird im unendlich–dimensionalen Fall zur Rechtfertigung des Grenzüberganges bei der Approximation durch endlich–dimensionale Probleme benötigt.

Satz A.3 läßt sich auf Abbildungen $A\colon B \to B^*$, B Banachraum, B^* Dualraum, verallgemeinern – die Monotoniebedingung lautet dann:

$$\langle Ax - Ay,\ x - y\rangle \geq 0\,.$$

Typische Anwendungen liegen im Bereich elliptischer partieller Differentialgleichungen, etwa

$$-\sum_{i=1}^{n} D_i\big(|\nabla u|^{p-2} D_i u\big) = f \text{ in } \Omega, \quad u \in H_0^{1,p}\,.$$

Die Abbildung A wird in diesem Falle definiert durch

$$\langle Au, \varphi\rangle = \sum_{i=1}^{n} \int\limits_{\Omega} |\nabla u|^{p-2}\, D_i u\, D_i\varphi\, dx\,.$$

Die Theorie der monotonen Operatoren erlaubt zahlreiche Verallgemeinerungen, z.B. Operatoren der Gestalt $Au = S(u, u) \in B^*$, wobei S in einem Argument monoton, im anderen vollstetig ist, oder, noch allgemeiner, sogenannte pseudo–monotone Operatoren, was wir hier nicht definieren wollen.

4. Die Leray–Schauder–Theorie

Diese Theorie ist ein starkes Hilfsmittel zur Verwertung von a priori–Abschätzungen im Hinblick auf Existenznachweise für Randwertprobleme bei partiellen Differentialgleichungen.
Eine einfache – black–box–version – verläuft folgendermaßen: Es sei S eine vollstetige Abbildung eines Banachraumes in sich. Es existiere eine Kugel B_R mit der Eigenschaft: „Für *etwaige* Lösungen u_0 der Gleichung

$$u + \theta S u = f$$

mit $\theta \in [0,1]$ gelte $u_0 \notin \partial B_R$." Dann besitzt die Gleichung $u + S = f$ eine Lösung in B_R.

Mit anderen Worten: Man führt einen „Monotonie–Parameter" $\theta \in [0,1]$ ein. Von der Gleichung $u + \theta S u = f$ wird die Lösbarkeit *nicht* verlangt, sondern nur, daß keine Lösung, sofern sie existiert, im Rand ∂B_R liegt. Im einfachsten Fall sichert man dies durch eine „a priori–Abschätzung" $\|u_0\| < R$.

Hintergrund für die Leray–Schauder–Theorie ist die Deformationsinvarianz des Abbildungsgrades, einer ganzen Zahl, die man dem Tripel $(u + \theta S u, f, B)$ zuordnen kann. Ist $f \in B_R$, $\theta = 0$, so ist der Abbildungsgrad $= 1$, und er bleibt bei der Deformation von u über $u + \theta S u$ nach $u + S u$ konstant, wenn etwaige Lösungen u_0 von $u + \theta S u = f$ nicht in B_R liegen.

Wenn der Abbildungsgrad des Tripels $(u + S u, f, B_R)$ 1 ist, so ist die Gleichung $u + S u = f$ lösbar.

Empfehlenswert ist für Anfänger ohne größere Topologiekentnisse, die Arbeiten von Nagumo über die Definition des Abbildungsgrades mit Hilfe von Funktionaldeterminanten zu lesen.

B Erläuterungen zum Literaturverzeichnis

Literatur zu Teil A

zu §0 Einführung:

Einen ersten Überblick bzw. Eindruck kann man sich anhand der Tafeln und Texte Seite 364–371 in Band 2 des "dtv–Atlas zur Mathematik" (Analysis und angewandte Mathematik) [54] verschaffen. Ausführlichere Informationen erhält man in 17 Artikeln zu Einzelthemen der Funktionalanalysis (z. B. in den Artikeln 199 Hilbert Spaces, 39 Banach Spaces, 407 Topological Linear Spaces, 251 Linear Operators, 72 Compact Operators, 135 Eigenvalue Problems; eine Liste sämtlicher Artikel findet man auf Seite 1531) in "Encyclopedic Dictionary of Mathematics" [30]. Dort findet man auch schon einige historische Hintergrundinformationen.

Den modernsten, ausführlich gehaltenen Überblick, der "angenehm" zu lesen ist, liefert E. Zeidler in Kapitel 11 (Lineare Funktionalanalysis und ihre Anwendungen), 12 (Nichtlineare Funktionalanalysis und ihre Anwendungen) Seite 338–480 des "Teubner–Taschenbuch der Mathematik Teil II" [22]. Neben viel Theorie findet man dort auch viele relevante Beispiele und ein modernes Literaturverzeichnis, welches wir als Ergänzung zu unserem beigefügten Literaturverzeichnis empfehlen.

Älter, kürzer, aber wesentliche Themen behandelnd ist der Bericht zur Funktionalanalysis in [8] Kapitel 11, Seite 422–484, der "Grundzüge der Mathematik, Band III Analysis".

Für historische Bezüge verweisen wir auf:

- die "Historical notes", Seite TVS V.80 - 92, in N. Bourbaki's "Topological Vector Spaces, Chapters 1–5" in der Reihe Elements of Mathematics, [10];
- die "Notes and Remarks" am Ende der einzelnen Kapitel in Dunford–Schwartz: "Linear Operators I, II, III", [15];
- Kapitel XIX "Ein Blick aud die werdende Funktionalanalysis", Seite 599–663, in H. Heuser's "Funktionalanalysis", [27]; sowie schließlich auf
- J. Dieudonné: "History of Functional Analysis", North–Holland, Amsterdam 1981.

Etwas aus der Reihe empfehlen wir auch einen Blick in

- R. Courant und D. Hilbert: “Methoden der Mathematischen Physik I, II”, [13], sowie
- S. Michlin: “Variationsmethoden der mathematischen Physik”, [45],

zu werfen. In diesen beiden Werken werden zahlreiche physikalisch motivierte und relevante Beispiele diskutiert. Der “Courant–Hilbert” ist außerdem aus historischen Gründen empfehlenswert: In einem gewissen Sinne ist es das erste Lehrbuch, welches über Methoden der Funktinalanalysis berichtet. Insbesondere enthält es in Kapitel 7 Band II “Lösung der Rand– und Eigenwertprobleme auf Grund der Variationsrechnung.” eine Methode zur Behandlung von Randwertproblemen für gewöhnliche und partielle Differentialgleichungen zweiter Ordnung, die man als Vorläufer der heutzutage verwendeten funktionalanalytischen Methode (Stichwort: “schwache Theorie”) ansehen muß.

Literatur

[1] N. I. Achieser und I. M. Glasmann: Theorie der linearen Operatoren im Hilbertraum, Akademie–Verlag, Berlin 1975 (sechste, unveränderte Auflage).

[2] P. Alexandroff und H. Hopf: Topologie, Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen Band 45, Springer–Verlag, Berlin 1935 (Berichtigter Reprint 1974).

[3] H. W. Alt: Lineare Funktionalanalysis, Springer–Verlag, Berlin 1992 (zweite, verbesserte Auflage).

[4] A. Ambrosetti and G. Prodi: A Primer of Nonlinear Analysis, Cambridge Studies in Advanced Mathematics 34, Cambridge University Press, Cambridge 1993.

[5] J. P. Aubin: Applied Functional Analysis, J. Wiley, New York 1979.

[6] S. Axler, P. Bourdon and W. Ramey: Harmonic Function Theory, Graduate Texts in Mathematics vol. 137, Springer–Verlag, New York 1992.

[7] G. Bachman and L. Narici: Functional Analysis, Academic Press, New York 1966.

[8] H. Behnke, F. Bachmann, K. Fladt und W. Süss: Grundzüge der Mathematik Band III Analysis, Vandenhoeck & Ruprecht, Göttingen 1968.

[9] L. Bers, F. John and M. Schechter: Partial Differential Equations, Lectures in Applied Mathematics vol. 3, Interscience Publishers (a division of John Wiley & Sons, Inc.), New York 1964.

[10] N. Bourbaki: Topological Vector Spaces, Chapters 1-5 (in Elements of Mathematics), Springer-Verlag, Berlin 1987.

[11] H. Brezis: Analyse Fonctionnelle et Applications, Masson, Paris 1983 (oder in italienisch: Analisi funzionale: Teoria e applicazioni, Liguori Editore, S. r. l., Napoli 1986; englische Übersetzung soll bei Springer in Vorbereitung sein!).

[12] R. F. Brown: A Topological Introduction To Nonlinear Analysis, Birkhäuser, Boston - Basel - Berlin 1993.

[13] R. Courant und D. Hilbert: Methoden der Mathematischen Physik I,II, Springer–Verlag, Berlin 1968 (dritte bzw. zweite Auflage; Heidelberger Taschenbücher Band 30/31).

[14] J. Dieudonné: History of Functional Analysis, North–Holland, Amsterdam 1981.

[15] N. Dunford and J. T. Schwartz: Linear Operators I,II,III; Interscience Publishers, Inc. (a division of John Wiley, New York), New York 1966 (Third Printing).

[16] R. E. Edwards: Functional Analysis, Holt, Rinehart and Winston, New York 1965.

[17] L. C. Evans and R. Gariepy: Measure Theory and Fine Properties of Functions, Studies in Advanced Mathematics, CRC Press, Boca Raton, Florida 1992.

[18] G. Fichera: Linear elliptic differential systems and eigenvalue problems, Lecture Notes in Mathematics Band 8, Springer–Verlag, Berlin 1965.

[19] D. G. Figueiredo: Lectures on The Ekeland Variational Principle with Applications and Detours, Tata Institute of Fundamental Research Bombay vol. 81, Springer–Verlag, Berlin 1989.

[20] G. B. Folland: Introduction to Partial Differential Equations, Princeton University Press, Princeton New Jersey 1995 (Second Edition).

[21] D. Gilbarg and N. S. Trudinger: Elliptic Partial Differential Equations of Second Order, Grundlehren der mathematischen Wissenschaften Band 224, Springer-Verlag, Berlin 1983 (Second edition).

[22] G. Grosche, E. Zeidler, D. Ziegler und V. Ziegler: Teubner–Taschenbuch der Mathematik, Teil II, B. G. Teubner Verlag Stuttgart–Leipzig 1995 (7. Auflage, vollständig überarbeitete und wesentlich erweiterte Neufassung der 6. Auflage der "Ergänzenden Kapitel zum Taschenbuch der Mathematikvon I. N. Bronstein und K. A. Semendjajew").

[23] P. Habala, P. Hájek and V. Zizler: Introduction to Banach Spaces I,II; MATFYZPRESS, vydavatelství Matematicko-fyzikálni fakulty Univerzity Karlovy, Praha 1996.

[24] P. Halmos: A Hilbert Space Problem Book, Springer–Verlag 1974.

[25] E. Harzheim und H. Ratschek: Einführung in die allgemeine Topologie, Wissenschaftliche Buchgesellschaft Darmstadt 1975.

[26] E. Harzheim: Einführung in die Kombinatorische Topologie, Wissenschaftliche Buchgesellschaft, Darmstadt 1978.

[27] H. Heuser: Funktionalanalysis, B. G. Teubner Verlag, Stuttgart 1986 (zweite, neubearbeitete und verbesserte Auflage).

[28] E. Hewitt and K. Stromberg: Real and Abstract Analysis, Springer–Verlag, New York 1965 (Graduate Texts in Mathematics 25).

[29] F. Hirzebruch und W. Scharlau: Einführung in die Funktionalanalysis, Bibibliographisches Institut, Mannheim 1971.

[30] S. Iyanaga and Y. Kawada (ed.): Encyclopedic Dictionary of Mathematics I,II (by the Mathematical Society of Japan), The MIT Press, Cambridge, Massachusetts, and London. England 1980 (First paperback edition).

[31] H. Jeggle: Nichtlineare Funktionalanalysis, B. G. Teubner-Verlag, Stuttgart 1979.

[32] F. John: Partial Differential Equations, Applied Mathematical Sciences vol. 1, Springer–Verlag, New York 1986 (Fourth Editon, second printing).

[33] J. Jost: Postmodern Analysis, Springer–Verlag, Berlin 1998.

[34] J. Jost: Partielle Differentialgleichungen, Springer–Verlag, Berlin 1998.

[35] L. W. Kantorowitsch und G. P. Akilow: Funtionalanalysis in normierten Räumen, Verlag Harri Deutsch, Thun-Frankfurt/Main 1978.

[36] T. Kato: Perturbation Theory for Linear Operators, Springer–Verlag, Berlin 1995 (Corrected Printing of the Second Edition 1980 in der Reihe Classics in Mathematics).

[37] G. Köthe: Topologische Lineare Räume I,II; Springer–Verlag, Berlin 1966/1979 (Grundlehren der mathematischen Wissenschaften Band 107).

[38] A. N. Kolmogorov and S. V. Fomin: Introductory Real Analysis, Prentice-Hall, Inc., Englewood Cliffs, N. J. 1970.

[39] S. Lang: Real and Functional Analysis, Springer–Verlag, New York 1993 (Third Edition; Graduate Texts in Mathematics vol. 142).

[40] R. Leis: Funktionalanalysis I,II; Skript zur Vorlesung im Wintersemester 1994/95 bzw. im Sommersemester 1995, Universität Bonn.

[41] J. Lindenstrauss and L. Tzafriri: Classical Banach Spaces I and II, Springer–Verlag, Berlin 1996 (Originally published as Vol. 92 and Vol. 97 of the Ergebnisse der Mathematik und ihrer Grenzgebiete, Springer–Verlag, Berlin 1977).

[42] H. v. Mangoldt und K. Knopp (bearb. von F. Lösch): Höhere Mathematik Band IV, S. Hirzel Wissenschaftliche Verlagsgesellschaft Stuttgart 1990 (vierte, durchgesehene Auflage).

[43] K. Maurin: Methods Of Hilbert Spaces, Polska Akademia Nauk Monografie Matematyczne Tom 45, PWN–Polish Scientific Publishers, Warszawa 1967.

[44] R. Meise und D. Vogt: Einführung in die Funktionalanlysis, Vieweg–Verlag, Braunschweig/Wiesbaden 1992 (Vieweg Studium;62: Aufbaukurs Mathematik).

[45] S. G. Michlin: Variationsmethoden der Mathematischen Physik, Akademie–Verlag, Berlin 1962.

[46] S. G. Michlin: The Problem of the Minimum of a Quadratic Functional, Holden–Day, San Francisco - London - Amsterdam 1965.

[47] C. Miranda: Problemi Di Esistenza In Analisi Funzionale, Pubblicazione della Classe di Scienze della Scuola Normale Superiore, Pisa 1975 (Ristampa Anno Accademico 1948/49).

[48] C. Miranda: Istituzioni Di Analisi Funzionale Lineare (vol. I,II), Unione Matematica Italiana 1978 (I), 1979 (II), Tipografia Oderisi Editrice - Gubbio 1978/9.

[49] C. Miranda: Partial Differential Equations of Elliptic Type, Ergebnisse der Mathematik Band 2, Springer-Verlag, New York 1970 (Second Revised Edition).

[50] G. Pedersen: Analysis Now, Springer–Verlag, Berlin 1989.

[51] A. Pietsch: Eigenvalues and s-numbers, Cambridge University Press, Cambridge 1987 (Cambridge Studies in Advanced Mathematics 13).

[52] B. v. Querenburg: Mengentheoretische Topologie, Springer–Verlag, Berlin 1979 (zweite, neubearbeitete und erweiterte Auflage).

[53] M. Reed and B. Simon: Methods of Modern Mathematical Physics I; Functional Analysis, Academic Press 1980.

[54] F. Reinhardt und H. Soeder: dtv–Atlas zur Mathematik Band 2, Deutscher Taschenbuch Verlag, München 1980 (dritte Auflage).

[55] F. Riesz und B. Sz.- Nagy: Vorlesungen über Funktionalanalysis, Verlag Harri Deutsch, Thun-Frankfurt 1982 (vierte, durchgesehene Auflage).

[56] R. T. Rockafellar: Convex Analysis, Princeton Landmarks in Mathematics and Physics, Princeton University Press, Princeton, New Jersey 1997 (tenth printing and first paperback).

[57] H. L. Royden: Real Analysis, Macmillan Publishing Company, New York 1988 (Third Edition)

[58] W. Rudin: Functional Analysis, Tata McGraw-Hill Publ. Comp., New Delhi 1985 (9th Reprint).

[59] W. Rudin: Real and Complex Analysis, McGraw-Hill Publ. Comp., New York 1987.

[60] M. Schechter: Principles of Functional Analysis, Academic Press, New York - San Franzisco - London 1971.

[61] A. Taylor and D. Lay: Introduction to Functional Analysis, John-Wiley 1980 (second edition).

[62] H. Triebel: Höhere Analysis, Verlag Harri Deutsch, Thun-Frankfurt 1980 (zweite, verbesserte Auflage).

[63] W. Velte: Direkte Methoden der Variationsrechnung, Leitfäden der angewandten Mathematik und Mechanik Band 26, B. G. Teubner–Verlag, Stuttgart 1976.

[64] W. Walter: Gewöhnliche Differentialgleichungen, Springer-Verlag, Berlin 1996 (Sechste, überarbeitete und erweiterte Auflage).

[65] J. Weidmann: Lineare Operatoren in Hilberträumen, B. G. Teubner Verlag, Stuttgart 1976.

[66] D. Werner: Funktionalanalysis, Springer–Verlag, Berlin 1995.

[67] J. Wloka: Funktionalanalysis und ihre Anwendungen, Walter de Gruyter, Berlin 1971.

[68] K. Yosida: Functional Analysis, Springer–Verlag, New York 1988 (5th edition).

[69] P. Zahn: Ein konstruktiver Weg zur Masstheorie und Funktionalanalysis, Wissenschaftliche Buchgesellschaft, Darmstadt 1978.

[70] E. Zeidler: Applied Functional Analysis, Springer–Verlag, New York 1995 (Applied Mathematical Sciences 108/109).