Vorlesungen über

Funktionalanalysis

von

Rolf Leis

Vorwort

In den Jahren 1992–1996 habe ich eine Vorlesungsreihe zur Einführung in die Analysis für Studenten der Mathematik und Physik an der Universität Bonn gehalten. Diese Reihe begann mit den Vorlesungen über Infinitesimalrechnung I–IV für das Grundstudium. Daran schlossen sich im Hauptstudium jeweils zweisemestrige Vorlesungen über Funktionalanalysis und Partielle Differentialgleichungen an.

Im Anschluß an diese Vorlesungen habe ich meine Aufzeichnungen den Hörern in Form von Skripten zur Verfügung gestellt. Der vorliegende Text über Funktionalanalysis ist aus den entsprechenden Skripten entstanden. Dabei habe ich den Stoffumfang bewußt auf den einer zweisemestrigen Vorlesung beschränkt und mich bei der Auswahl darum bemüht, Anwendungen aufzuzeigen, die besonders auch für Studenten der Physik mit dem Nebenfach Mathematik interessant sein können. Der Text ist der erste Band aus der Reihe dieser Vorlesungen für das Hauptstudium, im zweiten Band werden Partielle Differentialgleichungen behandelt.

Danken möchte ich vor allem Frau R. Müller für die hervorragende TEX-Niederschrift und vielen Hörern meiner Vorlesungen für ihre Kommentare zu den Skripten und für ihre Hilfe bei deren Erstellung. Herr F. Linke hat den Text der Skripten kritisch durchgesehen und das Sachverzeichnis erstellt.

Bonn, im Herbst 1997

Rolf Leis

Inhaltsverzeichnis

Einführung **1**

1 Grundlagen **5**
1.1 Algebraische Strukturen . . . 5
1.1.1 Das Zornsche Lemma . . . 9
1.1.2 Fortsetzung linearer Abbildungen . . . 10
1.2 Topologische Grundlagen . . . 13
1.2.1 Topologische Räume . . . 13
1.2.2 Metrische Räume . . . 15
1.2.3 Räume von Abbildungen . . . 18
1.2.4 Kompakte Mengen . . . 20
1.2.5 Der Satz von Arzelà-Ascoli . . . 24
1.3 Normierte Räume und stetige lineare Abbildungen . . . 26
1.3.1 Normierte Räume . . . 27
1.3.2 Beispiele . . . 29
1.3.3 Stetige lineare Abbildungen . . . 31
1.3.4 Lineare Homöomorphismen . . . 33
1.3.5 Endlichdimensionale Räume . . . 35
1.4 Hilberträume . . . 37
1.4.1 Skalarprodukte . . . 37
1.4.2 Der Approximationssatz . . . 39
1.4.3 Orthonormalsysteme und separable Räume . . . 40
1.4.4 Projektoren . . . 45
1.5 Dualräume und beschränkt lineare Funktionale . . . 47
1.5.1 Der Satz von Hahn-Banach . . . 48
1.5.2 Konjugierte Abbildungen . . . 50
1.5.3 Der Rieszsche Darstellungssatz . . . 52

2 Fixpunktsätze **54**
2.1 Der Banachsche Fixpunktsatz . . . 54
2.2 Der Brouwersche Fixpunktsatz . . . 54
2.3 Der Schaudersche Fixpunktsatz . . . 56

3 Der Satz von Baire-Hausdorff und Folgerungen **59**
3.1 Das Prinzip von der gleichmäßigen Beschränktheit . . . 60
3.2 Der Satz von Banach-Steinhaus . . . 61
3.3 Die Sätze von der offenen und der inversen Abbildung . . . 62

4 Schwache Konvergenz **64**
4.1 Schwache Kompaktheit . . . 66
4.2 Schwache Konvergenz in speziellen Räumen . . . 68
4.3 Konvexe Mengen . . . 71
4.4 Schwache Topologien . . . 74

5 Abgeschlossene Operatoren **77**
5.1 Der Satz vom abgeschlossenen Graphen . . . 79
5.2 Abschließbare Abbildungen . . . 80
5.3 Adjungierte Abbildungen . . . 83
5.4 Ein Beispiel: Der Δ-Operator . . . 89

6 Spezielle Funktionenräume **93**
6.1 Klassische Funktionenräume . . . 93
6.2 Lebesguesche Räume . . . 95
6.3 Einige Abschätzungen . . . 96
6.4 Spezielle Techniken . . . 98
6.4.1 Faltungen . . . 98
6.4.2 Der Friedrichssche Glättungsoperator . . . 100
6.4.3 Dirac-Folgen . . . 101

6.4.4 Die Fouriertransformation . . . 103
6.5 Distributionen . . . 106
6.6 Sobolevräume . . . 110
6.6.1 Stark gleich Schwach . . . 115
6.6.2 Stetige Einbettungen . . . 118
6.6.3 Kompakte Einbettungen . . . 121
6.6.4 Ein Fortsetzungssatz . . . 123
6.7 Anwendungen . . . 123
6.7.1 Das Hausdorffsche Momentenproblem . . . 123
6.7.2 Ein Ergodensatz . . . 126
6.7.3 Minimalprobleme der Variationsrechnung . . . 128

7 Über das Spektrum linearer, insbesondere kompakter linearer Abbildungen 131
7.1 Resolvente und Spektrum . . . 131
7.2 Kompakte lineare Abbildungen . . . 143
7.3 Die Riesz-Schauder Theorie . . . 147
7.4 Die Integralgleichungsmethode in der Potentialtheorie . . . 157
7.5 Das Spektrum kompakter linearer Abbildungen . . . 160
7.6 Das Spektrum kompakter normaler Abbildungen . . . 163
7.7 Klassifikation kompakter linearer Abbildungen . . . 172

8 Selbstadjungierte Fortsetzungen symmetrischer Abbildungen 175
8.1 Die Cayleytransformation . . . 175
8.2 Halbbeschränkte Abbildungen . . . 182

9 Der Spektralsatz für selbstadjungierte Abbildungen 185
9.1 Spektralscharen . . . 185
9.2 Der Spektralsatz . . . 191
9.3 Folgerungen . . . 199

10 Ergänzungen 207
10.1 Halbgruppen . . . 207
10.1.1 Gleichmäßig stetige Halbgruppen . . . 208
10.1.2 Stark stetige Halbgruppen . . . 210
10.1.3 Stark stetige Gruppen . . . 218
10.2 Ein Darstellungssatz . . . 219
10.3 Spektraldarstellung unitärer Abbildungen . . . 220

Literaturverzeichnis 221

Bezeichnungen 223

Namen- und Sachverzeichnis 225

Einführung

Diese Vorlesung über *Funktionalanalysis* soll einen Kurs über Infinitesimalrechnung fortsetzen. In ihr wird die Analysis der *Funktionale*, das heißt der Abbildungen $f : X \to \mathbb{K}$ behandelt. Dabei ist X ein Vektorraum über $\mathbb{K}$. Die Bezeichnung Funktional wurde wohl 1903 von Jacques Hadamard, 1865–1963, eingeführt.

Die Funktionalanalysis ist in diesem Jahrhundert entstanden. Noch während meines Studiums wurde eine solche Vorlesung nicht angeboten. Damals erschien gerade das heute klassische Buch von *Riesz & Sz.-Nagy*, und wir haben es mit Begeisterung gelesen.

Die Funktionalanalysis ist ein sehr reizvolles Gebiet. Sie lernen in ihr eine Menge moderner Mathematik. So werden verallgemeinerte Ableitungen und Räume solcher differenzierbarer Abbildungen behandelt. Überhaupt ist die Tendenz zum Abstrakten stark; und auf diese Weise bringt man wieder etwas Ordnung in die Vielfalt der Einzelergebnisse. Andererseits sollte man dabei aber auch nicht zu weit gehen und die Funktionalanalysis nicht zur Strukturmathematik verkümmern lassen. Ich werde mich deshalb immer bemühen, den Anwendungsbezug erkennen zu lassen, wobei ich vor allem an Anwendungen aus der Mathematischen Physik denke.

All das möchte ich Ihnen an einem Beispiel deutlicher machen: Es seien $G \subset \mathbb{R}^n$ ein beschränktes Gebiet und $f \in \mathcal{C}(\partial G)$ am Rande vorgegeben. Eine wichtige Aufgabe in der Physik ist es, ein $u \in \mathcal{K} := \mathcal{C}_2(G) \cap \mathcal{C}(\bar{G})$ zu bestimmen mit

$$\Delta u = 0 \quad \text{und} \quad u|\partial G = f.$$

Die Funktion u beschreibt zum Beispiel das Potential einer elektrischen Ladungsverteilung im Vakuum. Es handelt sich also um eine Randwertaufgabe für u, man spricht von der *Dirichletschen Randbedingung* und von der *Dirichletschen Randwertaufgabe*. Natürlich sind auch andere Randbedingungen möglich. In $\mathbb{R}^2$ ist die Dirichletsche Randwertaufgabe zur Aussage im Riemannschen Abbildungssatz äquivalent.

Es gibt viele Methoden, die Dirichletsche Randwertaufgabe zu lösen. Dieses Problem hat eine große Bedeutung für die weitere Entwicklung der Mathematik gehabt, und es läßt sich spannend darüber berichten. Ein klassischer Beweis, der auch von Riemann selbst verwandt worden ist, verwendet das sogenannte *Dirichletsche Prinzip*. Dieses geht auf Gauß und Kelvin zurück und hängt eng mit den Extremalprinzipien der klassischen Mechanik zusammen. Denken Sie an das *Maupertiussche Prinzip von der kleinsten Wirkung* oder an das *Hamiltonsche Prinzip*.

Man wähle

$$\mathcal{D} := \{ v \in \mathcal{K} \mid D(v) < \infty \text{ und } v|\partial G = f \}$$

mit

$$D(v) := \int_G |\nabla v|^2.$$

Das Integral $D(v)$ nennt man *Dirichletsches Integral* von v. Das Dirichletsche Prinzip besagt dann, daß man die gesuchte Lösung u der Dirichletschen Randwertaufgabe durch Minimieren von $D(v)$ für $v \in \mathcal{D}$ erhält.

Das Minimum von $D(v)$ werde von $u \in \mathcal{D}$ angenommen. Dann folgt für alle Testfunktionen $\varphi \in \mathring{\mathcal{C}}_\infty(G)$ und alle $\varepsilon \in \mathbb{R}$

$$D(u + \varepsilon\varphi) \geq D(u)$$

oder

$$2\varepsilon \int_G \nabla u \cdot \nabla\varphi + \varepsilon^2 \int_G |\nabla\varphi|^2 \geq 0.$$

Daraus erhält man

$$\forall\, \varphi \in \mathring{\mathcal{C}}_\infty(G) \qquad \int_G \nabla u \cdot \nabla\varphi = 0.$$

Aus dieser Beziehung folgt $\Delta u = 0$. Das Minimum u löst also die Dirichletsche Randwertaufgabe.

Die andere Richtung, nämlich den Nachweis der Existenz eines Minimums, hielt man wegen $D(v) \geq 0$ und der physikalischen Evidenz zunächst für trivial und kümmerte sich nicht weiter darum. Riemann verwendete das Prinzip in dieser Form. Der Einspruch und ein Gegenbeispiel wurden erst 1870 von Weierstraß veröffentlicht. Wohl existiert das Infimum von $D(v)$ für $v \in \mathcal{D}$. Es ist aber nicht klar, ob es auch in $\mathcal{D}$ angenommen wird.

Es hat dann etwa dreißig Jahre gedauert, bis nach vielen Versuchen die Dirichletsche Randwertaufgabe befriedigend gelöst werden konnte. Durch dieses intensive Bemühen sind viele Erkenntnisse gewonnen worden. Der Gegenstand unserer Vorlesung ist entstanden und letztlich das, was wir heute „moderne Mathematik“ nennen.

Zwei Lösungsmethoden zur Dirichletschen Randwertaufgabe möchte ich kurz skizzieren: Da ist einmal die *Integralgleichungsmethode*. Man macht hier einen Ansatz

$$u(x) = \frac{1}{2\pi} \int_{\partial G} \mu(y) \, \frac{\partial}{\partial n_y} \frac{1}{|x-y|} \, dy$$

mit einer stetigen *Dipolbelegung* μ und erhält für μ eine Integralgleichung der Form

$$(id + K)\mu = f$$

mit einem kompakten Operator $K : \mathcal{C}(\partial G) \to \mathcal{C}(\partial G)$. Das setzt einen glatten Rand ∂G voraus. Von Fredholm, Hilbert und seinen Schülern wurde etwa ab 1900 eine Lösungstheorie für solche Gleichungen entwickelt. Die Resultate wurden dann abstrakter gefaßt, verallgemeinert, und es entstand die *Riesz-Schauder-Theorie*.

Beim Lösen solcher Gleichungen blieb man aber nicht stehen. Eine Theorie der Abbildungen entwickelte sich. Dabei spielen Darstellungssätze eine große Rolle. Denken Sie an die Jordansche Normalform für Matrizen. Entsprechend gibt es Spektraldarstellungen für allgemeinere Operatoren. Berühmt ist der *von Neumannsche Spektralsatz* für selbstadjungierte Abbildungen. Betrachten wir zum Beispiel in G eine Wellengleichung

$$u_{tt} + Au = o$$

mit den Anfangsbedingungen

$$u(0,x) = u^0(x) \quad \text{und} \quad u_t(0,x) = u^1(x).$$

Dabei seien u^0, u^1 vorgegeben und etwa $A = -\Delta$ mit homogener Dirichletscher Randbedingung. Mit Hilfe des Spektralsatzes lassen sich dann auch Funktionen von A definieren, und man erhält die Lösung der Wellengleichung durch

$$u(t, \cdot) = \big(\cos \sqrt{A}\, t\big) u^0 + \frac{\sin \sqrt{A}\, t}{\sqrt{A}} \, u^1.$$

Aus dieser Darstellung lassen sich wichtige Folgerungen ziehen.

Ich komme jetzt zur zweiten Methode, die ich Ihnen schildern möchte, nämlich der *Hilbertraummethode*. Auch sie geht auf Hilbert und seine Schule zurück. Hierbei geht es um die Fundierung des Dirichletschen Prinzips selbst. Beim Beweis des Dirichletschen Prinzips stehen zwei Gedanken, die von Hilbert klar herausgestellt wurden, im Vordergrund. Einmal muß die Klasse der zugelassenen Lösungen geeignet erweitert werden, um mehr Chancen zum Lösen zu haben und um vollständige Lösungsräume zu erhalten (Verallgemeinerung des Lösungsbegriffs). Es gibt nämlich sehr einfache Beispiele für Gebiete, für die die klassische Dirichletsche Randwertaufgabe nicht lösbar ist, das Dirichletsche Prinzip aber zu einer Lösung führt. Man denke etwa an den punktierten Einheitskreis mit den Randwerten $u = 0$ für $|x| = 1$ und $u = 1$ für $x = 0$.

Zum anderen müssen geeignete Einschränkungen an die Klasse der zugelassenen Randwerte f gefunden werden. Die klassische Voraussetzung $f \in \mathcal{C}(\partial G)$ ist ja wegen der aufgetretenen Gegenbeispiele nicht angemessen. Man wähle etwa den Einheitskreis und die stetigen Randwerte

$$f(\varphi) := \sum_{n=1}^{\infty} \frac{\sin n!\,\varphi}{n^2}.$$

Dann ist die Dirichletsche Randwertaufgabe zu diesen Daten mit Hilfe der klassischen Poissonschen Integralformel lösbar, das Dirichletsche Integral existiert aber für die Lösung nicht, wie man leicht nachrechnet. Die Methode des Dirichletschen Prinzips versagt also in diesem Fall.

Zur Verallgemeinerung des Lösungsbegriffs verallgemeinert man nun zunächst den Begriff der Ableitung. Es seien u und g_i aus $\mathcal{L}^2(G)$ sowie für $i = 1, 2, \ldots, n$

$$\forall\, \varphi \in \mathring{\mathcal{C}}_\infty(G) \qquad (u, \partial_i \varphi) = -(g_i, \varphi).$$

Dann nennt man die $g_i =: \partial_i u$ *schwache Ableitungen* von u und erhält so den *Hilbertraum* $\mathcal{W}_1(G)$ mit dem Skalarprodukt

$$(u,v)_1 := (u,v) + (\nabla u, \nabla v).$$

$\mathcal{W}_1(G)$ ist also die Klasse der in G einmal schwach differenzierbaren Funktionen. Die Menge der Abbildungen $\varphi \in \mathring{\mathcal{C}}_\infty(G)$ ist in $\mathcal{L}^2(G)$ dicht; man nennt sie auch *Testfunktionen* und sagt, die schwachen

Ableitungen seien durch *Testen* gewonnen worden. In vielen Anwendungen möchte man aber lieber *approximieren*, ähnlich wie beim Approximieren reeller Zahlen durch rationale. Man definiert deshalb auch *starke Ableitungen* und die Klasse der einmal stark differenzierbaren Funktionen durch

$$\mathcal{H}_1(G) := \Big\{ \mathcal{C}_1(G) \cap \mathcal{W}_1(G);\ \|\cdot\|_1 \Big\}^{\sim}.$$

Der Raum $\mathcal{H}_1(G)$ ist also die Vervollständigung von $\mathcal{C}_1(G) \cap \mathcal{W}_1(G)$ unter der Norm $\|u\|_1 := \sqrt{(u,u)_1}$. Entsprechend verfährt man mit den höheren Ableitungen. Wichtig ist nun die Aussage

$$\mathcal{H}_1(G) = \mathcal{W}_1(G),$$

also *stark gleich schwach*. Diese Aussage wurde in voller Allgemeinheit erst sehr spät von *Meyers & Serrin* bewiesen (1964). Ich skizziere für unseren einfachen Fall den Beweis: $\mathcal{H}_1(G)$ ist ein abgeschlossener Teilraum von $\mathcal{W}_1(G)$. Aus dem Projektionssatz folgt daher

$$\mathcal{W}_1(G) = \mathcal{H}_1(G) \oplus \mathcal{H}_1(G)^{\perp}.$$

Es sei nun $u \in \mathcal{H}_1(G)^{\perp}$. Dann folgt

$$\forall \varphi \in \mathring{\mathcal{C}}_\infty(G) \qquad (u,\varphi)_1 = 0.$$

Die Abbildung u ist mithin eine *schwache Lösung* von $(-\Delta + 1)u = o$, und aus *Regularitätssätzen* (zum Beispiel dem Weylschen Lemma aus dem Jahre 1940) folgt – wie in der Funktionentheorie – $u \in \mathcal{C}_\infty(G)$. Damit ist aber $u \in \mathcal{H}_1(G)$, also $u = o$.

Der Raum $\mathcal{H}_1(G)$ ist daher ein geeigneter Kandidat für den Lösungsraum. Es fehlt aber noch die Formulierung des Annehmens der Randwerte. Dazu wählt man zunächst

$$\mathring{\mathcal{H}}_1(G) := \Big\{ \mathring{\mathcal{C}}_\infty(G);\ \|\cdot\|_1 \Big\}^{\sim},$$

also die Vervollständigung der Testfunktionen unter der $\|\cdot\|_1$-Norm. Ist u ein Element von $\mathring{\mathcal{H}}_1(G)$, dann verallgemeinert das die klassische Aussage $u|\partial G = o$.

Als letztes muß man noch die Klasse der Randwerte f festlegen. Nach Hilbert wählt man nun $f \in \mathcal{H}_1(G)$. Die Funktion f soll also in G erklärt sein, und das Dirichletsche Integral von f soll existieren.

Zur Verdeutlichung der Grundidee möchte ich nun in dieser Einführung die einfachere Gleichung

$$(-\Delta + 1)u = o$$

mit $u|\partial G = f$ lösen. Es sei $f \in \mathcal{H}_1(G)$ vorgegeben. Wir suchen dann eine Lösung $u \in \mathcal{H}_1(G)$ mit

$$\forall \varphi \in \mathring{\mathcal{C}}_\infty(G) \quad (u,\varphi)_1 = 0 \quad \text{und} \quad u - f \in \mathring{\mathcal{H}}_1(G).$$

Diese Randwertaufgabe ist eindeutig lösbar. Das folgt unmittelbar aus dem Projektionssatz

$$\begin{aligned} \mathcal{H}_1(G) &= \mathring{\mathcal{H}}_1(G) \oplus \mathring{\mathcal{H}}_1(G)^{\perp}, \\ f \quad &= \quad v \quad + \quad u. \end{aligned}$$

Die Projektion u von f auf $\mathring{\mathcal{H}}_1(G)^{\perp}$ ist die gesuchte Lösung. Dem Beweis des Projektionssatzes aus dem Approximationssatz entnimmt man, daß u das Dirichletsche Integral minimiert. Ihre Herkunft aus der Variationsrechnung, also das Minimieren des Dirichletschen Integrals, sieht man dieser Lösungsmethode heute kaum noch an.

Die bisher vorgestellten Probleme waren zum guten Teil linear. Ich möchte aber betonen, daß in der Funktionalanalysis auch viele interessante nichtlineare Fragen behandelt werden. Das gilt natürlich auch für die Theorie der Differentialgleichungen. Wählt man bei der Saite die potentielle Energie proportional zur Längenänderung, dann erhält man eine nichtlineare Gleichung. Die Wellengleichung ist nur ihre lineare Näherung.

Zum Abschluß dieser Einführung möchte ich einige Bücher nennen, die ich bei der Vorbereitung dieser Vorlesung benutzt habe. Es gibt viel mehr, bitte stöbern Sie selbst in der Bibliothek.

H. W. Alt: Lineare Funktionalanalysis. Springer-Verlag 1985.

R. Brown: A Topological Introduction to Nonlinear Analysis. Birkhäuser Verlag 1993.

J. Dieudonné: History of Functional Analysis. North-Holland 1981.

P. Halmos: A Hilbert Space Problem Book. Springer-Verlag 1974.

H. Heuser: Funktionalanalysis. B.G. Teubner Verlag 1986.

F. Hirzebruch und W. Scharlau: Einführung in die Funktionalanalysis. Bibliographisches Institut 1971.

T. Kato: Perturbation Theory for Linear Operators. Springer-Verlag 1966.

G. Pedersen: Analysis Now. Springer-Verlag 1989.

M. Reed and B. Simon: Methods of Modern Mathematical Physics I; Functional Analysis. Academic Press 1980.

F. Riesz und B. Sz.-Nagy: Vorlesungen über Funktionalanalysis. VEB Deutscher Verlag der Wissenschaften 1956.

W. Rudin: Real and complex analysis. McGraw-Hill Publ. Comp. 1987.

A. Taylor and D. Lay: Introduction to Functional Analysis. John Wiley 1980.

J. Weidmann: Lineare Operatoren in Hilberträumen. B.G. Teubner Verlag 1976.

K. Yosida: Functional Analysis. Springer-Verlag 1965.

Weitere Literaturhinweise finden Sie im Literaturverzeichnis. Jahreszahlen in eckigen Klammern deuten auf diese Verweise.

1 Grundlagen

1.1 Algebraische Strukturen

Wir beginnen mit einer kurzen Wiederholung algebraischer Begriffe:

Eine *Gruppe* G ist eine Menge mit einer Abbildung $\circ$, die jedem Paar $x, y \in G$ ein Element von G zuordnet,

$$(x, y) \longmapsto x \circ y \in G,$$

und die die Eigenschaften

(1) $\forall\, x, y, z \in G \quad (x \circ y) \circ z = x \circ (y \circ z)$ (Die Abbildung ist *assoziativ*),

(2) $\exists\, n \in G \quad \forall\, x \in G \quad x \circ n = x$ (Ein rechtsseitiges *neutrales Element* existiert),

(3) $\forall\, x \in G \quad \exists\, y \in G \quad x \circ y = n$ (Ein rechtsseitiges *inverses Element* existiert)

besitzt. Dann ist auch $n \circ x = x$ und $x^{-1} \circ x = n$ mit $x^{-1} := y$. Die Gruppe heißt *abelsch*, wenn zusätzlich

(4) $\forall\, x, y \in G \quad x \circ y = y \circ x$

gilt.

Es sei $\mathbb{K}$ eine abelsche Gruppe mit der *Addition* $+$ und dem neutralen Element o. Es sei $\mathbb{K}\backslash\{o\}$ außerdem eine abelsche Gruppe bezüglich der *Multiplikation* $\cdot$, und es gelte das *Distributivgesetz*

$$\forall\, x, y, z \in \mathbb{K} \qquad x \cdot (y + z) = (x \cdot y) + (x \cdot z).$$

Dann ist $\mathbb{K}$ ein *Körper*. Wenn nichts anderes gesagt wird, steht im folgenden $\mathbb{K}$ für $\mathbb{R}$ oder $\mathbb{C}$, die reellen oder komplexen Zahlen.

Ein *linearer Raum* oder *Vektorraum* über dem Körper $\mathbb{K}$ ist nun eine Menge $\mathcal{X}$ mit zwei Abbildungen

$$\begin{aligned} + \quad &: \quad \mathcal{X} \times \mathcal{X} \longrightarrow \mathcal{X}, \\ \cdot \quad &: \quad \mathbb{K} \times \mathcal{X} \longrightarrow \mathcal{X} \end{aligned}$$

mit

(1) $\mathcal{X}$ ist eine abelsche Gruppe bezüglich $+$.

(2) Die *Multiplikation* mit den *Skalaren* $\alpha, \beta \in \mathbb{K}$ hat die Eigenschaften

$$\begin{aligned} \alpha(x + y) &= \alpha x + \alpha y, \\ (\alpha + \beta)x &= \alpha x + \beta x, \\ \alpha(\beta x) &= (\alpha\beta)x, \\ 1 \cdot x &= x, \\ 0 \cdot x &= o. \end{aligned}$$

Mit $o \in \mathcal{X}$ bezeichnet man das neutrale Element bezüglich der Addition.

Beispiele für Vektorräume sind natürlich der $\mathbb{R}^n$, der $\mathbb{C}^n$ oder auch

$$\ell^2 := \left\{ x = (\xi_1, \xi_2, \dots) \;\middle|\; \sum_{i=1}^{\infty} |\xi_i|^2 < \infty, \quad \xi_i \in \mathbb{K} \right\}.$$

Eine Teilmenge $\mathcal{M} \subset \mathcal{X}$ heißt *Unterraum*, *Teilraum* oder *lineare Mannigfaltigkeit*, wenn $\mathcal{M}$ selbst ein linearer Raum ist. Es sei $\mathcal{S} \subset \mathcal{X}$ eine Teilmenge von $\mathcal{X}$. Die Menge aller (endlichen) Linearkombinationen

$$\sum_{i=1}^{n} \alpha_i s_i \qquad \text{mit } n \in \mathbb{N},\ \alpha_i \in \mathbb{K} \text{ und } s_i \in \mathcal{S}$$

von Elementen aus $\mathcal{S}$ heißt der von $\mathcal{S}$ *erzeugte (aufgespannte) Teilraum* $\operatorname{span}(\mathcal{S})$. Dieser Teilraum ist der kleinste Teilraum von $\mathcal{X}$, der $\mathcal{S}$ enthält.

Die Elemente $x_1, \dots, x_n \in \mathcal{X}$ heißen *linear unabhängig*, wenn

$$\Big(\sum_{i=1}^{n} \alpha_i x_i = o \Big) \implies \alpha_1 = \alpha_2 = \dots = \alpha_n = 0$$

gilt. Es sei Λ eine beliebige Indexmenge. Dann heißen die $(x_\lambda)_{\lambda\in\Lambda}$ linear unabhängig, wenn jede endliche Teilmenge linear unabhängig ist. Es sei $n \in \mathbb{N}$ mit

(1) Es gibt n linear unabhängige Elemente $x_1, \ldots, x_n$.

(2) Alle $n+1$ Elemente $x_1, \ldots, x_{n+1}$ sind linear abhängig.

Dann ist $n =: \dim\mathcal{X}$ die Dimension von $\mathcal{X}$. Im folgenden werden uns natürlich Räume unendlicher Dimension, wie ℓ^2 oder der Raum der quadratintegrierbaren Funktionen, besonders interessieren.

Eine Teilmenge $\mathcal{H}$ linear unabhängiger Elemente von $\mathcal{X}$ heißt *Hamel-Basis* von $\mathcal{X}$, wenn sich jedes $x \in \mathcal{X}$ als Linearkombination von Elementen von $\mathcal{H}$ darstellen läßt. Diese Basis ist nach GEORGE HAMEL, 1877–1954, benannt. Man sollte sie nicht mit der *Schauder-Basis* verwechseln, die in Bemerkung 1.3.4.9 eingeführt wird. Die Existenz einer Hamelbasis folgt aus dem Zornschen Lemma. Im Verlauf dieses Abschnitts werden wir darauf noch näher eingehen.

Lineare Abbildungen: Es seien nun $\mathcal{X}$ und $\mathcal{Y}$ lineare Räume über $\mathbb{K}$. Wir betrachten Abbildungen

$$A \;:\; \mathcal{X} \longrightarrow \mathcal{Y}$$

oder allgemeiner

$$A \;:\; \mathcal{D}(A) \subset \mathcal{X} \longrightarrow \mathcal{Y}.$$

Dann sind

$\mathcal{D}(A) \subset \mathcal{X}$ der *Definitionsbereich* von A,
$\mathcal{R}(A) \subset \mathcal{Y}$ der *Wertebereich (range)* von A,
$\mathcal{N}(A) \subset \mathcal{X}$ der *Kern (Nullraum)* von A.

Eine solche Abbildung heißt *linear*, wenn $\mathcal{D}(A)$ ein Unterraum von $\mathcal{X}$ ist und

(1) $A(x_1 + x_2) = Ax_1 + Ax_2$,

(2) $A(\alpha x) = \alpha Ax$

gilt. Es sei $\mathcal{L}(\mathcal{X},\mathcal{Y})$ die Menge der linearen Abbildungen auf $\mathcal{X}$ in $\mathcal{Y}$ (synonym: von $\mathcal{X}$ nach $\mathcal{Y}$). Es soll also $\mathcal{D}(A) = \mathcal{X}$ sein. Man spricht auch von *linearen Transformationen*, *linearen Operatoren* oder *Homomorphismen* und verwendet $\mathrm{Hom}(\mathcal{X},\mathcal{Y})$ für $\mathcal{L}(\mathcal{X},\mathcal{Y})$. Eine lineare Abbildung auf $\mathcal{X}$ in $\mathcal{X}$ ist ein *Endomorphismus*. $\mathcal{L}(\mathcal{X},\mathcal{Y})$ selbst ist wieder ein Vektorraum über $\mathbb{K}$. Mit $A, B \in \mathcal{L}(\mathcal{X},\mathcal{Y})$ und $\alpha \in \mathbb{K}$ ist

$$(A+B)x := Ax + Bx,$$
$$(\alpha A)x := A(\alpha x).$$

Man sieht leicht, daß $\mathcal{R}(A) \subset \mathcal{Y}$ und $\mathcal{N}(A) \subset \mathcal{X}$ Unterräume sind. Wir zeigen

Satz 1.1.1: *Es sei $A \in \mathcal{L}(\mathcal{X},\mathcal{Y})$. Dann gilt*

(1) $A^{-1} : \mathcal{R}(A) \subset \mathcal{Y} \to \mathcal{X}$ *existiert* $\Longleftrightarrow \;\; \mathcal{N}(A) = \{o\}$.

(2) Wenn A^{-1} existiert, ist $A^{-1} \in \mathcal{L}(\mathcal{R}(A),\mathcal{X})$.

Beweis:

Zu (1) $\Rightarrow$: $\quad Ax = o \rightsquigarrow x = o$ wegen $x = A^{-1}Ax = o$.
$\Leftarrow$: $\quad Ax_1 = Ax_2 \rightsquigarrow A(x_1 - x_2) = o \rightsquigarrow x_1 = x_2$.

Zu (2): Es sei $Ax = y$, also $x = A^{-1}y$. Dann folgt aus der Linearität von A

$$A^{-1}(y_1+y_2) = A^{-1}(Ax_1 + Ax_2) = A^{-1}A(x_1+x_2) = x_1 + x_2 = A^{-1}y_1 + A^{-1}y_2$$

sowie

$$A^{-1}(\alpha y) = A^{-1}(\alpha Ax) = A^{-1}A\alpha x = \alpha x = \alpha A^{-1}y.$$

□

Im folgenden wird uns immer wieder die Frage beschäftigen, für welche $y \in \mathcal{Y}$ es ein $x \in \mathcal{X}$ mit

$$Ax = y \qquad (*)$$

gibt. Man möchte also Gleichungen lösen. Das ist natürlich besonders einfach, wenn A^{-1} existiert. Allgemein stellen sich die Frage nach der Existenz einer Lösung (was ist $\mathcal{R}(A)$?) und die Frage nach ihrer

Eindeutigkeit (was ist $\mathcal{N}(A)$?). Hat man neben der algebraischen auch noch eine topologische Struktur zur Verfügung, dann möchte man natürlich mehr über die Lösung erfahren, zum Beispiel wissen, ob sie stetig von vorgegebenen Daten abhängt. Im Augenblick soll uns das aber noch nicht interessieren.

Im Falle $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ zeigt man in der linearen Algebra

1. $\dim\mathcal{N}(A) = \dim\mathcal{N}(A^{tr})$.
2. $\mathbb{R}^n = \mathcal{R}(A) \oplus \mathcal{N}(A^{tr})$ orthogonal.

Dabei ist A^{tr} die zu A transponierte Abbildung (Matrix), und die Orthogonalität ist bezüglich des euklidischen Skalarprodukts

$$(x, y) = \sum_{i=1}^{n} \xi_i \eta_i$$

in $\mathbb{R}^n$ gemeint; $x = (\xi_1, \ldots, \xi_n)$, $y = (\eta_1, \ldots, \eta_n)$.

Man sagt auch, für Gleichung $(*)$ gelte die *Fredholmsche Alternative*, benannt nach Ivar Fredholm (1866–1927), nämlich: Entweder gilt

(1) Gl. $(*)$ ist für alle $y \in \mathbb{R}^n$ eindeutig lösbar

oder

(2) es ist $\dim\mathcal{N}(A) = \dim\mathcal{N}(A^{tr}) \neq 0$. Dann ist $\mathcal{R}(A) \neq \mathbb{R}^n$, und notwendig und hinreichend für die Lösbarkeit von Gl. $(*)$ ist die Bedingung $y \in \mathcal{R}(A) = \mathcal{N}(A^{tr})^{\perp}$, also

$$\forall\, z \in \mathcal{N}(A^{tr}) \qquad (y, z) = 0. \tag{$**$}$$

Im Fall (2) ist eine Lösung zunächst natürlich nicht eindeutig festgelegt; Eindeutigkeit läßt sich zum Beispiel durch die Forderung $x \in \mathcal{N}(A)^{\perp}$ erzwingen.

Wir werden dieses Ergebnis später verallgemeinern. Jetzt erinnern wir an

Beispiel 1.1.2: *Mit* $J := [0, 1]$ *seien* $\mathcal{X} := \left\{x \in \mathcal{C}_2(J) \;\middle|\; x(0) = x'(0) = 0\right\}$; $\mathcal{Y} := \mathcal{C}(J)$; $p, q \in \mathcal{C}(J)$ *und*

$$\begin{aligned} A \;:\; & \mathcal{X} \longrightarrow \mathcal{Y}, \\ & x \longmapsto x'' + px' + qx. \end{aligned}$$

Unsere Gleichung $Ax = y$ ist dann also eine gewöhnliche Differentialgleichung. Sie ist linear, und die Lipschitzbedingung ist erfüllt. Der Satz von Picard-Lindelöf aus der Theorie der gewöhnlichen Differentialgleichungen besagt dann

$$\mathcal{N}(A) = \{o\} \quad \text{und} \quad \mathcal{R}(A) = \mathcal{Y}.$$

In diesem Beispiel existiert A^{-1} also. Betrachten wir andererseits

Beispiel 1.1.3: *Es seien* $\mathcal{X} = \mathcal{Y} = \mathcal{C}_\infty(\mathbb{R})$ *und*

$$\begin{aligned} A \;:\; & \mathcal{X} \longrightarrow \mathcal{Y}, \\ & x \longmapsto x'. \end{aligned}$$

Nun löst

$$x(t) = \int_0^t y(s)\, ds$$

unsere Gleichung $Ax = y$. Es ist also $\mathcal{R}(A) = \mathcal{Y}$, aber $\mathcal{N}(A) = \text{span}(1)$. Die Fredholmsche Alternative gilt jetzt also nicht mehr.

Lineare Funktionale: Wir fahren nun mit der Diskussion algebraischer Strukturen fort.

Definition 1.1.4: *Die Vektorräume* $\mathcal{X}$, $\mathcal{Y}$ *über* $\mathbb{K}$ *heißen* isomorph, *wenn eine lineare bijektive Abbildung*

$$A \;:\; \mathcal{X} \longrightarrow \mathcal{Y}$$

existiert.

Definition 1.1.5: $A \in \mathcal{L}(\mathcal{X}, \mathbb{K})$ *heißt* lineares Funktional.

$$\mathcal{X}^f := \mathcal{L}(\mathcal{X}, \mathbb{K})$$

ist der algebraisch konjugierte Raum *zu* $\mathcal{X}$.

Mit $x \in \mathcal{X}$ und $x' \in \mathcal{X}^f$ schreiben wir

$$x'(x) =: x'x.$$

Beispiel 1.1.6: *Es seien* $\mathcal{X} = \mathbb{R}^n$, $x_1, \ldots, x_n$ *eine Basis von* $\mathcal{X}$ *und für ein* $x \in \mathcal{X}$, $\xi_i \in \mathbb{R}$

$$x = \sum_{i=1}^{n} \xi_i x_i.$$

Es seien ferner $x'x_j =: \alpha_j \in \mathbb{R}$ *für ein* $x' \in \mathcal{X}^f$. *Dann ist*

$$x'x = \sum_{j=1}^{n} \alpha_j \xi_j.$$

Betrachten wir die lineare Abbildung

$$\begin{aligned} A \;:\; & \mathcal{X}^f \longrightarrow \mathcal{X}, \\ & x' \longmapsto (\alpha_1, \ldots, \alpha_n). \end{aligned}$$

Dann gilt $\mathcal{D}(A) = \mathcal{X}^f$ *und* $\mathcal{R}(A) = \mathcal{X}$, *denn jedes* $(\alpha_1, \ldots, \alpha_n)$ *definiert ein* x' *durch*

$$x'x := \sum_{j=1}^{n} \alpha_j \xi_j.$$

Außerdem ist $\mathcal{N}(A) = \{o\}$, *denn* $Ax' = o$ *bedeutet*

$$\forall\, x \in \mathcal{X} \qquad x'x = 0 \iff x' = o.$$

Folgerung 1.1.7: *Es sei* $\mathcal{X} = \mathbb{R}^n$. *Dann sind* $\mathcal{X}$ *und* $\mathcal{X}^f$ *isomorph.*

Im Falle ∞-dimensionaler Räume $\mathcal{X}$ ist $\mathcal{X}^f$ im allgemeinen zu groß. Man betrachtet deshalb Teilräume, die eine zusätzliche topologische Struktur besitzen (stetige lineare Funktionale). Darüber folgt im nächsten Abschnitt mehr.

Zu $\mathcal{X}^f$ kann man nun wieder den algebraisch konjugierten Raum $\mathcal{X}^{ff} := \mathcal{L}(\mathcal{X}^f, \mathbb{K})$ bilden. Dieser Raum wird auch *Bidualraum* genannt. Es sei $x \in \mathcal{X}$ fest. Dann ist $x'x \in \mathbb{K}$ für jedes $x' \in \mathcal{X}^f$, definiert also eine Abbildung $x'' \in \mathcal{X}^{ff}$:

$$\forall\, x' \in \mathcal{X}^f \qquad x''x' := x'x$$

oder

$$\forall\, x \in \mathcal{X} \quad \overset{1}{\exists}\, x'' \in \mathcal{X}^{ff} \quad \forall\, x' \in \mathcal{X}^f \qquad x''x' = x'x.$$

Es sei

$$\begin{aligned} J \;:\; & \mathcal{X} \longrightarrow \mathcal{X}^{ff}, \\ & x \longmapsto x'' \end{aligned}$$

diese *kanonische Abbildung* von $\mathcal{X}$ in $\mathcal{X}^{ff}$. Sie ist linear. Aus $Jx_i = x_i''$, $i = 1, 2$, und $J(x_1 + x_2) = x_3''$ folgt zum Beispiel

$$x_3''x' = x'(x_1 + x_2) = x'x_1 + x'x_2 = x_1''x' + x_2''x' = (x_1'' + x_2'')x',$$

also

$$J(x_1 + x_2) = Jx_1 + Jx_2.$$

Es sei $Jx = o$, also

$$\forall\, x' \in \mathcal{X}^f \qquad x'x = 0.$$

Daraus möchten wir $x = o$, also $\mathcal{N}(J) = \{o\}$, schließen. Das ist nicht trivial, denn es ist noch nicht klar, ob $\mathcal{X}^f$ „genügend groß ist“. Vielmehr folgt das erst als Folgerung 1.1.15 aus dem Zornschen Lemma. Wir nehmen das Ergebnis vorweg:

Folgerung 1.1.8: $\mathcal{R}(J)$ *ist zu* $\mathcal{X}$ *isomorph.*

Schließlich definieren wir

Definition 1.1.9: $\mathcal{X}$ *heißt* algebraisch reflexiv, *wenn*

$$\mathcal{R}(J) = \mathcal{X}^{ff}$$

ist.

Transponierte Abbildungen: Es seien $\mathcal{X}$, $\mathcal{Y}$ lineare Räume und $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$. Dann definiert man die zu A *transponierte Abbildung* A^{tr} durch

$$\begin{aligned} A^{tr} \quad : \quad & \mathcal{Y}^f \longrightarrow \mathcal{X}^f, \\ & y' \longmapsto x' := y' \circ A. \end{aligned}$$

Es ist $A^{tr} \in \mathcal{L}(\mathcal{Y}^f, \mathcal{X}^f)$ und

$$\forall\, y' \in \mathcal{Y}^f \quad \forall\, x \in \mathcal{X} \qquad (A^{tr} y')x = y'(Ax).$$

1.1.1 Das Zornsche Lemma

In der Vorlesung über Infinitesimalrechnung benutzt man Axiome zur Einführung der Zahlen, dabei besonders das *Vollständigkeitsaxiom*. Nur an einer Stelle wird das *Auswahlaxiom* benötigt; nämlich, um eine im Lebesgueschen Sinne nicht meßbare Menge zu konstruieren, die Vitali-Menge. Diese Menge wird jedoch in der Infinitesimalrechnung meist nicht weiter verwendet.

Von nun an möchten wir aber das *Auswahlaxiom* wirklich benutzen, vor allem um Induktionsbeweise führen zu können, die über Teilmengen von $\mathbb{N}$ hinausgehen (sogenannte *transfinite Induktionen*). Das Auswahlaxiom wurde 1904 von Ernst Zermelo, 1871–1953, formuliert und ist heute weitgehend unumstritten. Lebesgue und viele andere Mathematiker haben es jedoch in den zwanziger Jahren abgelehnt. Für weitere Einzelheiten muß auf eine Vorlesung über Grundlagenfragen oder Mengenlehre verwiesen werden.

Es gibt verschiedene äquivalente Formulierungen des Auswahlaxioms; neben dem Auswahlaxiom selbst den *Wohlordnungssatz* oder eben das *Zornsche Lemma*. Letzteres wird nach Max Zorn, 1906–1993, benannt. Bevor die Formulierungen angegeben werden, folgen einige Vorbemerkungen:

Es sei zunächst daran erinnert, daß eine *Relation* eine Teilmenge $R \subset X \times X$, X eine Menge, mit folgenden Eigenschaften ist:

(1) $x \sim x$ (Die Relation ist *reflexiv*).

(2) $\big((x \sim y) \wedge (y \sim z)\big) \implies x \sim z$ (Die Relation ist *transitiv*).

(3′) $x \sim y \implies y \sim x$ (Die Relation ist *symmetrisch*).

oder

(3″) $(x \sim y) \wedge (y \sim x) \implies x = y$ (Die Relation ist *antisymmetrisch*).

Dabei haben wir $x \sim y$ für $(x, y) \in R$ geschrieben.

Definition 1.1.10: *Es sei M eine Menge und $\leq$ eine antisymmetrische Relation von $M \times M$. Dann heißt M* halb (partiell) geordnet *und $\leq$ eine* Halbordnung.

Beispiel 1.1.11:

1. *Es seien $z, w \in \mathbb{C}$ mit $z = x + \mathrm{i}y$ und $w = u + \mathrm{i}v$. Für $x \leq u$ und $y \leq v$ sei $z \leq w$.*
2. *Es sei $M := \mathcal{P}(\mathcal{X})$ die Potenzmenge von $\mathcal{X}$. Dann definiert $A \subset B$ eine Halbordnung für M.*
3. *Es seien M eine halbgeordnete Menge und $S \subset M$. Dann ist auch S halbgeordnet.*

Definition 1.1.12: *Eine Menge M heißt* vollständig geordnet, *wenn für je zwei Elemente $a, b \in M$ entweder $a \leq b$ oder $b \leq a$ gilt.*

Definition 1.1.13: *Es seien die Menge M halbgeordnet und $S \subset M$. Dann heißt $m \in M$* obere Schranke *von S, wenn*

$$\forall\, s \in S \qquad s \leq m$$

gilt.

Definition 1.1.14: *Die Menge M sei halbgeordnet. Dann heißt $m \in M$* maximal, *wenn*

$$\Big(a \in M \ \wedge \ m \leq a\Big) \implies m = a$$

gilt.

Nach diesen Vorbereitungen folgen das Auswahlaxiom und äquivalente Formulierungen:

Auswahlaxiom: *Es sei M eine nichtleere Menge. Dann gibt es eine Abbildung (Auswahlfunktion)*

$$f \;:\; \mathcal{P}(M)\backslash\{\emptyset\} \longrightarrow M,$$

mit

$$\forall\, P \in \mathcal{P}(M)\backslash\{\emptyset\} \qquad f(P) \in P.$$

Äquivalente Formulierungen sind

Wohlordnungssatz: *Jede Menge kann wohlgeordnet werden.*

Damit ist folgendes gemeint: Zu jeder Menge M gibt es eine Halbordnung $\leq$ mit

$$\forall\, A \subset M,\, A \neq \emptyset, \quad \exists\, k \in A \quad \forall\, a \in A \qquad k \leq a.$$

Man nennt k *kleinstes Element* von A. Die Menge M ist dann vollständig geordnet; zu $x, y \in M$ wähle man nur $A := \{x, y\}$.

Das Zornsche Lemma: *M sei eine nichtleere halbgeordnete Menge. Jede vollständig geordnete Teilmenge von M besitze eine obere Schranke in M. Dann enthält M wenigstens ein maximales Element.*

Schließlich soll noch das Prinzip der *transfiniten Induktion* angegeben werden: Es seien X eine wohlgeordnete Menge und $P(x)$ für $x \in X$ eine Aussage mit

(1) $P(x_0)$ ist wahr für das kleinste Element $x_0 \in X$,

(2) $\big(\forall\, y < x \quad P(y)$ ist wahr$\big) \implies P(x)$ ist wahr.

Dann gilt $P(x)$ für alle $x \in X$.

Es sei nämlich

$$S := \big\{ x \in X \;\big|\; x > x_0 \;\text{ und } P(x) \text{ ist falsch} \big\}.$$

Wäre $S \neq \emptyset$, dann besäße S ein kleinstes Element s_0. Die Aussage $P(y)$ wäre dann für alle $y < s_0$ wahr, aber für s_0 falsch. Das widerspricht (2).

1.1.2 Fortsetzung linearer Abbildungen

Es seien nun wieder $\mathcal{X}$, $\mathcal{Y}$ lineare Räume, $\mathcal{M} \subset \mathcal{X}$ ein Teilraum und $f \in \mathcal{L}(\mathcal{M}, \mathcal{Y})$. Wir wollen f auf $\mathcal{X}$ fortsetzen; wir suchen also ein

$$F \in \mathcal{L}(\mathcal{X}, \mathcal{Y}) \;\text{ mit } F \,\big|\, \mathcal{M} = f.$$

Fortsetzungssatz: *Es seien $\mathcal{X}, \mathcal{Y}$ lineare Räume, $\mathcal{M} \subset \mathcal{X}$ ein Teilraum und $f \in \mathcal{L}(\mathcal{M}, \mathcal{Y})$. Dann gibt es ein $F \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ mit $F \,\big|\, \mathcal{M} = f$.*

Den Beweis führen wir in drei Schritten:

1. Es seien $x_1 \in \mathcal{X}\backslash\mathcal{M}$ und $\mathcal{M}_1 := \operatorname{span}(\mathcal{M}, x_1)$. Es sei $z \in \mathcal{M}_1$. Dann kann man

$$z = m + \alpha\, x_1 \quad \text{mit } m \in \mathcal{M},\; \alpha \in \mathbb{K}$$

eindeutig darstellen. Aus

$$m_1 + \alpha_1 x_1 = m_2 + \alpha_2 x_1$$

folgt nämlich

$$\mathcal{M} \ni m_1 - m_2 = (\alpha_2 - \alpha_1) x_1 \notin \mathcal{M}\backslash\{o\},$$

also $m_1 = m_2$ und $\alpha_1 = \alpha_2$. Man wähle nun ein beliebiges $y_1 \in \mathcal{Y}$ und

$$\begin{aligned} F_1 \;:\; \mathcal{M}_1 &\longrightarrow \mathcal{Y}, \\ z &\longmapsto f(m) + \alpha\, y_1. \end{aligned}$$

Dann ist F_1 eine Fortsetzung von f.

2. Für ein $k \in \mathbb{N}$ sei $\mathcal{X} = \operatorname{span}(\mathcal{M}, x_1, \ldots, x_k)$. Dann wiederholen wir den ersten Schritt, bis wir eine Fortsetzung auf den ganzen Raum erhalten haben.

3. Es gebe kein $k \in \mathbb{N}$ mit $\mathcal{X} = \operatorname{span}(\mathcal{M}, x_1, \ldots, x_k)$, $\mathcal{X}$ entstehe also nicht aus $\mathcal{M}$ durch Hinzufügen endlich vieler Elemente. Dies ist der eigentlich interessante Teil des Satzes, und wir benötigen zum Beweis das Zornsche Lemma. Für eine lineare Abbildung

$$g \;:\; \mathcal{D}(g) \subset \mathcal{X} \longrightarrow \mathcal{Y}$$

ist $\mathcal{D}(g)$ ein Teilraum von $\mathcal{X}$. Es sei

$$G := \{ g \in \mathcal{L}(\mathcal{D}(g), \mathcal{Y}) \;\mid\; \mathcal{M} \subset \mathcal{D}(g) \subset \mathcal{X},\; g|\mathcal{M} = f \}$$

die Menge aller Fortsetzungen von f. Wir definieren in G eine Halbordnung durch

$$g \leq h \;:\Longleftrightarrow\; \mathcal{D}(g) \subset \mathcal{D}(h) \;\wedge\; h|\mathcal{D}(g) = g.$$

Es sei nun Q eine vollständig geordnete Teilmenge von G. Wir wollen zeigen, daß Q eine obere Schranke $q \in G$ besitzt:

a) Es sei

$$\mathcal{D}(q) := \bigcup_{g \in Q} \mathcal{D}(g).$$

Man sieht sofort, daß $\mathcal{D}(q)$ ein Teilraum von $\mathcal{X}$ ist, denn zu $x_1, x_2 \in \mathcal{D}(q)$ gibt es $g_1, g_2 \in Q$ mit $x_i \in \mathcal{D}(g_i)$. Es sei etwa $g_1 \leq g_2$. Dann ist $\mathcal{D}(g_1) \subset \mathcal{D}(g_2)$, und es folgt

$$\begin{aligned} x_1 + x_2 &\in \mathcal{D}(g_2) \subset \mathcal{D}(q), \\ \alpha\, x_1 &\in \mathcal{D}(g_2) \subset \mathcal{D}(q). \end{aligned}$$

b) Es sei $x \in \mathcal{D}(q)$. Dann gibt es ein $g \in Q$ mit $x \in \mathcal{D}(g)$. Es sei

$$q(x) := g(x).$$

Diese Definition ist eindeutig, denn aus $x \in \mathcal{D}(g_1)$ und $x \in \mathcal{D}(g_2)$ folgt $g_1(x) = g_2(x)$, weil Q vollständig geordnet ist.

c) Es ist

$$q \in \mathcal{L}(\mathcal{D}(q), \mathcal{Y}).$$

Es seien nämlich $x_1, x_2 \in \mathcal{D}(q)$ und etwa $x_1 \in \mathcal{D}(g_1)$, $x_2 \in \mathcal{D}(g_2)$ mit $g_1 \leq g_2$. Dann ist $x_1, x_2 \in \mathcal{D}(g_2)$ und

$$\begin{aligned} q(x_1 + x_2) = g_2(x_1 + x_2) = g_2(x_1) + g_2(x_2) = q(x_1) + q(x_2), \\ q(\alpha\, x_1) = g_2(\alpha\, x_1) = \alpha\, g_2(x_1) = \alpha\, q(x_1). \end{aligned}$$

d) Es ist nach Konstruktion

$$q \in G \quad \text{und} \quad \forall\, g \in Q \quad g \leq q.$$

Mithin ist q obere Schranke von Q.

Das Zornsche Lemma impliziert daher die Existenz eines maximalen Elements F in G. Dann ist $\mathcal{D}(F) = \mathcal{X}$. Denn sonst gäbe es ein $x_1 \in \mathcal{X} \backslash \mathcal{D}(F)$, und F könnte wie im ersten Schritt fortgesetzt werden. Mithin ist $F \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ eine Fortsetzung von f. □

Damit haben wir ein fundamentales Resultat hergeleitet. Wir ziehen einige wichtige Folgerungen:

Trennungssatz: *Es seien $\mathcal{X}_0 \subsetneq \mathcal{X}$ ein Teilraum und $x_1 \in \mathcal{X} \backslash \mathcal{X}_0$. Dann gilt*

$$\exists\, x' \in \mathcal{X}^f \qquad x'x = \begin{cases} 0 & \text{für } x \in \mathcal{X}_0 \\ 1 & \text{für } x = x_1. \end{cases}$$

Beweis: Es seien $\mathcal{M} := \operatorname{span}(\mathcal{X}_0, x_1)$ und $\mathcal{Y} := \mathbb{K}$. Ein Element $m \in \mathcal{M}$ läßt sich dann durch

$$m = x + \alpha\, x_1 \quad \text{mit} \quad x \in \mathcal{X}_0 \quad \text{und} \quad \alpha \in \mathbb{K}$$

eindeutig darstellen. Es seien

$$\begin{aligned} f \;:\; \mathcal{M} &\longrightarrow \mathcal{Y}, \\ m &\longmapsto \alpha \end{aligned}$$

und x_1' eine Fortsetzung von f auf $\mathcal{X}$. □

Folgerung 1.1.15: *Es seien $x \in \mathcal{X}$ und*

$$\forall\, x' \in \mathcal{X}^f \qquad x'x = 0.$$

Dann ist $x = o$.

$\mathcal{X}^f$ ist also „genügend groß“, man kann *testen* und so zum Beispiel Eindeutigkeitssätze beweisen.

Der Beweis folgt aus dem Trennungssatz mit $\mathcal{X}_0 := \{o\}$. Wäre $x \neq o$, dann würde es ein x' mit $x'x = 1$ geben. □

Wir können jetzt auch zeigen, daß die kanonische Abbildung J injektiv ist: Es war

$$\begin{aligned} J \;:\; \mathcal{X} &\longrightarrow \mathcal{X}^{ff}, \\ x &\longmapsto x'' \end{aligned}$$

mit

$$\forall\, x \in \mathcal{X} \quad \overset{1}{\exists}\, x'' \in \mathcal{X}^{ff} \quad \forall\, x' \in \mathcal{X}^f \qquad x''x' = x'x.$$

Es sei $Jx = o$, also

$$\forall\, x' \in \mathcal{X}^f \qquad x'x = 0.$$

Dann ist $x = o$. Das war nachzutragen.

Zum Abschluß dieses Abschnitts zeigen wir noch die Existenz einer speziellen Fortsetzung, die später interessant sein wird. Wir beschränken uns hier auf den wichtigen Fall $\mathbb{K} = \mathbb{R}$. Die Übertragung auf das Komplexe ist eine leichte Verallgemeinerung.

Definition 1.1.16: *Es sei $\mathcal{X}$ ein linearer Raum über $\mathbb{R}$. Dann heißt*

$$p \;:\; \mathcal{X} \longrightarrow \mathbb{R}$$

ein sublineares Funktional *auf $\mathcal{X}$, wenn folgendes gilt:*

(1) $p(x + y) \leq p(x) + p(y)$.

(2) $p(\alpha x) = \alpha\, p(x)$ *für alle* $\alpha \geq 0$.

Beispielsweise ist jede Norm ein sublineares Funktional. Man fragt nun nach Fortsetzungen, die durch ein sublineares Funktional beschränkt werden.

Satz 1.1.17: *Es seien $\mathcal{X}$ ein linearer Raum über $\mathbb{R}$, $\mathcal{M} \subset \mathcal{X}$ ein Teilraum, p ein sublineares Funktional auf $\mathcal{X}$ und $f \in \mathcal{L}(\mathcal{M}, \mathbb{R})$ mit*

$$\forall\, x \in \mathcal{M} \qquad f(x) \leq p(x).$$

Dann gilt

$$\exists\, F \in \mathcal{X}^f \quad \text{mit } F \,\big|\, \mathcal{M} = f \quad \text{und } F(x) \leq p(x) \text{ für alle } x \in \mathcal{X}.$$

Der Beweis erfolgt analog zum Beweis des Fortsetzungssatzes aus dem Zornschen Lemma. Jetzt ist aber der erste Schritt wichtig:

Es seien $x_1 \in \mathcal{X} \backslash \mathcal{M}$ und $\mathcal{M}_1 := \operatorname{span}(\mathcal{M}, x_1)$. Wir suchen ein $\xi_1 \in \mathbb{R}$ mit

$$F_1(m + \alpha x_1) := f(m) + \alpha\, \xi_1$$

und

$$\forall\, m \in \mathcal{M} \quad \forall\, \alpha \in \mathbb{R} \qquad F_1(m + \alpha x_1) \leq p(m + \alpha x_1).$$

1. Wir leiten eine notwendige Bedingung für ξ_1 her: Für $\alpha = 1$ soll sein

$$\xi_1 \leq p(m + x_1) - f(m)$$

und für $\alpha = -1$ und $m \rightsquigarrow -m$

$$f(-m) - p(-m - x_1) \leq \xi_1,$$

zusammen also

$$\forall\, m \in \mathcal{M} \qquad -p(-m - x_1) - f(m) \leq \xi_1 \leq p(m + x_1) - f(m). \qquad (*)$$

2. Wir zeigen, daß diese Bedingung $(*)$ auch hinreichend ist: Sie sei also erfüllt. Dann folgt mit $\alpha > 0$ und $m \rightsquigarrow \frac{m}{\alpha}$ rechts

$$\xi_1 \le p\Big(\frac{m}{\alpha} + x_1\Big) - f\Big(\frac{m}{\alpha}\Big) = \frac{1}{\alpha}\big(p(m + \alpha x_1) - f(m)\big)$$

oder

$$F_1(m + \alpha x_1) = f(m) + \alpha\,\xi_1 \le p(m + \alpha x_1).$$

Es sei nun $\alpha < 0$. Dann folgt mit $m \rightsquigarrow \frac{m}{\alpha}$ links

$$\xi_1 \ge -p\Big(-\frac{m}{\alpha} - x_1\Big) - f\Big(\frac{m}{\alpha}\Big) = \frac{1}{\alpha}\big(p(m + \alpha x_1) - f(m)\big),$$

also für alle α

$$F_1(m + \alpha x_1) = f(m) + \alpha\,\xi_1 \le p(m + \alpha x_1).$$

3. Wir zeigen die Existenz eines ξ_1, das Bedingung $(*)$ erfüllt: Es seien $m_1, m_2 \in \mathcal{M}$. Dann ist

$$f(m_2) - f(m_1) = f(m_2 - m_1) \le p(m_2 - m_1) = p\big((m_2 + x_1) - (m_1 + x_1)\big) \le p(m_2 + x_1) + p(-m_1 - x_1)$$

oder

$$-p(-m_1 - x_1) - f(m_1) \le p(m_2 + x_1) - f(m_2).$$

Es seien

$$c := \sup_{m \in \mathcal{M}} \big(-p(-m - x_1) - f(m)\big)\,,$$
$$C := \inf_{m \in \mathcal{M}} \big(p(m + x_1) - f(m)\big)\,.$$

Beide Größen existieren, und es gilt

$$-\infty < c \le C < \infty.$$

Man wähle daher ξ_1 mit

$$c \le \xi_1 \le C.$$

Dann ist Bedingung $(*)$ erfüllt.

Der Beweis verläuft nun weiter wie der zum Fortsetzungssatz. □

1.2 Topologische Grundlagen

Wir wiederholen nun die hier benötigten topologischen Grundlagen und besprechen vor allem metrische Räume.

1.2.1 Topologische Räume

Definition 1.2.1: *Es sei X eine Menge und $\mathcal{T} \subset \mathcal{P}(X)$ ein System von Teilmengen. Dann heißt $\mathcal{T}$* Topologie *auf X und $(X, \mathcal{T})$* topologischer Raum*, wenn folgendes gilt:*

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

(2) Die Vereinigung beliebiger Elemente $O \in \mathcal{T}$ gehört zu $\mathcal{T}$,

$$\forall\, M \subset \mathcal{T} \qquad \bigcup_{O \in M} O \;\in \mathcal{T}.$$

(3) Für alle $O_1, O_2 \in \mathcal{T}$ ist $O_1 \cap O_2 \in \mathcal{T}$.

Die Menge $O \subset X$ heißt *offen* (bezüglich der Topologie), wenn $O \in \mathcal{T}$ ist. Eine beliebige Vereinigung und ein endlicher Durchschnitt offener Mengen ist also offen. $A \subset X$ heißt *abgeschlossen*, wenn $X \backslash A$ offen ist. Ein Punkt $y \in Y \subset X$ heißt *innerer* Punkt von Y, wenn es eine offene Menge O mit

$$y \in O \subset Y$$

gibt. Die Menge der inneren Punkte von Y bezeichnet man mit $\mathring{Y}$. Der *offene Kern* $\mathring{Y}$ ist die größte offene Menge, die in Y enthalten ist. Der *Abschluß* $\bar{Y}$ von Y ist der Durchschnitt aller abgeschlossenen Mengen $A \supset Y$. Die Teilmenge $Y \subset X$ heißt *dicht* in X, wenn $\bar{Y} = X$ ist.

Die leere Menge $\emptyset$ und X selbst sind stets offen und abgeschlossen.

$$\mathcal{T}_t := \{\emptyset, X\}$$

heißt *triviale* Topologie und

$$\mathcal{T}_d := \mathcal{P}(X)$$

diskrete Topologie. $\mathcal{T}_d$ ist die *stärkste* oder *feinste* Topologie von X, $\mathcal{T}_t$ die *schwächste* oder *gröbste*. $\mathcal{T}_1$ ist feiner als $\mathcal{T}_2$, wenn

$$\mathcal{T}_t \subset \mathcal{T}_2 \subset \mathcal{T}_1 \subset \mathcal{T}_d$$

gilt.

Definition 1.2.2: *Es seien $(X, \mathcal{T})$ ein topologischer Raum und $x \in X$. Dann heißt $U \subset X$* Umgebung *von x, wenn*

$$\exists\, O \in \mathcal{T} \qquad x \in O \subset U$$

gilt. Die Menge der Umgebungen von x heißt Umgebungsfilter *$O(x)$.*

In $(X, \mathcal{T}_d)$ ist $\{x\}$ Umgebung von x. Jeder Punkt liegt also *isoliert*, daher auch der Name diskrete Topologie. In $X = \mathbb{R}^n$ wählen wir die *euklidische* Topologie $\mathcal{T}_e$. Elemente von $\mathcal{T}_e$ sind die üblichen offenen Mengen O des $\mathbb{R}^n$, also

$$\forall\, o \in O \quad \exists\, B(o, \varepsilon) \subset O \qquad \text{mit } B(x, \varepsilon) := \left\{ y \in X \;\middle|\; |y - x| < \varepsilon \right\}.$$

Offenbar gilt

$$\mathcal{T}_t \subset \mathcal{T}_e \subset \mathcal{T}_d.$$

Man sagt auch, $\mathcal{T}_e$ werde von den Kugeln $B(x, \varepsilon)$ *erzeugt*. Es gilt nämlich

Satz 1.2.3: *Zu $F \subset \mathcal{P}(X)$ gibt es eine schwächste Topologie $\mathcal{F}$ auf X mit $F \subset \mathcal{F}$.*

Beweis: Es sei Ω die Familie aller Topologien $\mathcal{T}$ auf X, die F enthalten. Ω ist nicht leer wegen $\mathcal{T}_d \subset \Omega$. Es sei $\mathcal{F}$ der Durchschnitt aller $\omega \in \Omega$. Dann ist $F \in \mathcal{F}$, und es gilt

$$\forall\, \omega \in \Omega \qquad \mathcal{F} \subset \omega.$$

Wir müssen also nur noch überprüfen, ob $\mathcal{F}$ selbst eine Topologie ist. Es seien etwa $O_\lambda \in \mathcal{F}$. Dann gilt

$$\forall\, \omega \in \Omega \qquad O_\lambda \in \omega$$

und damit $\cup O_\lambda \in \omega$, folglich $\cup O_\lambda \in \mathcal{F}$. Der Nachweis der anderen Eigenschaften ist ebenfalls klar. □

Die Menge F heißt dann *Subbasis* von $\mathcal{F}$. Ist jede Menge in $\mathcal{F}$ eine Vereinigung von Mengen in F, dann heißt F *Basis* von $\mathcal{F}$. Die Topologie $\mathcal{F}$ besteht aus $\emptyset, X$ und Mengen der Form

$$\bigcup \left(\bigcap_{i=1}^{k} F_i \right) \qquad \text{mit } F_i \in F \text{ und } k \in \mathbb{N}.$$

Eine Menge $F \subset \mathcal{P}(X)$ enthält eine *Umgebungsbasis* für den Punkt $x \in X$, wenn

$$\forall\, A \in O(x) \quad \exists\, B \in F \cap O(x) \qquad B \subset A$$

gilt. F ist Basis, wenn F für jeden Punkt $x \in X$ eine Umgebungsbasis enthält.

Der topologische Raum $(X, \mathcal{T})$ heißt *Hausdorffraum* (benannt nach Felix Hausdorff, 1868–1942), wenn das folgende „Trennungsaxiom“ gilt: Zu $x, y \in X$ mit $x \neq y$ gibt es offene Mengen O_x, O_y mit $x \in O_x$, $y \in O_y$ und $O_x \cap O_y = \emptyset$.

Eine Folge $(x_n)_{n \in \mathbb{N}}$ heißt gegen $x \in X$ konvergent, wenn

$$\forall\, O \subset \mathcal{T},\, x \in O, \quad \exists\, N > 0 \quad \forall\, n \geq N \qquad x_n \in O$$

gilt. Ist X hausdorffsch, dann ist der Grenzwert x eindeutig bestimmt, wenn er existiert.

Wir benutzen im folgenden die in der Analysis üblichen Bezeichnungen. Es sei $Y \subset X$. Dann ist $\bar{Y}$ die Menge der *Berührungspunkte* von Y.

$$\partial Y := \bar{Y} \backslash \mathring{Y}$$

ist die Menge der *Randpunkte*.

Es seien $(X, \mathcal{T})$ ein topologischer Raum. Dann definiert man die *relative Topologie* für eine Teilmenge $M \subset X$ als System aller Mengen der Form

$$T \cap M \quad \text{mit } T \in \mathcal{T}.$$

Eine Teilmenge $Z \subset M$ der Form $Z = A \cap M$ mit in X abgeschlossenem A ist also bezüglich dieser Topologie abgeschlossen. Wir sprechen in solchen Fällen von relativ offenen oder abgeschlossenen Mengen.

Die relative Topologie nennt man auch die *von X induzierte Topologie*, genauer, die durch $id : M \to X$ induzierte Topologie. Es sei nämlich $f : M \to X$ eine Abbildung und

$$\mathcal{T}_f := \left\{ f^{-1}(O) \;\middle|\; O \subset X \text{ offen} \right\}.$$

Dann ist $\mathcal{T}_f$ die durch f induzierte Topologie von M.

Es seien nun $(X, \mathcal{T}_X)$, $(Y, \mathcal{T}_Y)$ topologische Räume und

$$f : X \longrightarrow Y$$

eine Abbildung. Dann heißt f *offen*, wenn

$$\forall\, O \in \mathcal{T}_X \qquad f(O) \in \mathcal{T}_Y$$

ist. Bilder offener Mengen sollen also offen sein. f heißt *stetig*, wenn die Urbilder offener Mengen offen sind, wenn also

$$\forall\, O \in \mathcal{T}_Y \qquad f^{-1}(O) \in \mathcal{T}_X$$

gilt. Die Räume $(X, \mathcal{T}_X)$ und $(Y, \mathcal{T}_Y)$ heißen *homöomorph*, wenn es eine bijektive, offene und stetige Abbildung

$$f : X \longrightarrow Y$$

gibt. Die Funktionen f und f^{-1} sollen also stetig sein. Aus topologischer Sicht sind $(X, \mathcal{T}_X)$ und $(Y, \mathcal{T}_Y)$ dann nicht zu unterscheiden.

1.2.2 Metrische Räume

Wir betrachten nun spezielle topologische Räume, nämlich die metrischen.

Definition 1.2.4: *Ein* metrischer Raum (M, d) *ist eine Menge M mit einer Abbildung*

$$d : M \times M \longrightarrow \mathbb{R}_0^+$$

mit

(1) $d(x, y) = 0 \iff x = y$,

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

(3) $d(x, y) \leq d(x, z) + d(z, y)$ (Dreiecksungleichung).

d heißt Metrik *auf M.*

Jeder metrische Raum wird durch die von den Kugeln

$$B(x, \varepsilon) := \left\{ y \in M \;\middle|\; d(x, y) < \varepsilon \right\}$$

erzeugte Topologie zu einem topologischen Raum, aber nicht jeder topologische Raum ist metrisierbar. Es gibt natürlich viele Beispiele für Metriken. Denken Sie an

1. Die diskrete Metrik:

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

2. Metriken in $\mathbb{R}^n$:

$$d_e(x,y) := d_2(x,y) := |x-y| = \sqrt{\sum_{i=1}^{n} |\xi_i - \eta_i|^2},$$

$$d_p(x,y) := \Big(\sum_{i=1}^{n} |\xi_i - \eta_i|^p \Big)^{1/p} \qquad \text{mit } 1 \le p < \infty,$$

$$d_\infty(x,y) := \sup_{i=1,\dots,n} |\xi_i - \eta_i|.$$

Dabei seien $x = (\xi_1, \dots, \xi_n)$ und $y = (\eta_1, \dots, \eta_n)$. Es gilt

$$d_\infty \le d_p \le \sqrt[p]{n}\, d_\infty .$$

3. Metriken in $\mathcal{C}([0,1])$:

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

$$d_p(f,g) := \|f-g\|_p := \Big(\int_0^1 |f(x)-g(x)|^p\, dx \Big)^{1/p} \qquad \text{mit } 1 \le p < \infty.$$

Aus der Dreiecksgleichung folgt

$$d(x,x') \le d(x,y) + d(y,y') + d(y',x'),$$

also

$$d(x,x') - d(y,y') \le d(x,y) + d(y',x').$$

Vertauscht man (x,x') mit (y,y'), dann folgen

$$|d(x,x') - d(y,y')| \le d(x,y) + d(x',y') \qquad (*)$$

und

$$|d(x,z) - d(y,z)| \le d(x,y) \le d(x,z) + d(z,y). \qquad (**)$$

Damit können wir in metrischen Räumen in gewohnter Weise Analysis treiben. Alle Inklusionen

$$id \;:\; (\mathbb{R}^n, d_i) \longrightarrow (\mathbb{R}^n, d_j)$$

sind stetig. Im Falle $d_i = d_\infty$, $d_j = d_p$ ist

$$d_p \le \sqrt[p]{n}\, d_\infty ,$$

zu vorgegebenem $\varepsilon > 0$ hat man also nur $\delta := \varepsilon / \sqrt[p]{n}$ zu wählen. Für solche stetigen Einbettungen schreibt man auch

$$(\mathbb{R}^n, d_i) \hookrightarrow (\mathbb{R}^n, d_j).$$

Die Inklusion

$$id \;:\; \Big(\mathcal{C}([0,1]), d_\infty\Big) \longrightarrow \Big(\mathcal{C}([0,1]), d_p\Big)$$

ist wegen $d_p \le d_\infty$ ebenfalls stetig, nicht aber

$$id \;:\; \Big(\mathcal{C}([0,1]), d_p\Big) \longrightarrow \Big(\mathcal{C}([0,1]), d_\infty\Big).$$

Man wähle nur

$$f_n(x) := \begin{cases} 1 - nx & \text{für } 0 \le x < \frac{1}{n} \\ 0 & \text{sonst.} \end{cases}$$

Dann ist $d_\infty(f_n, o) = 1$ und

$$d_p(f_n, o) \le \Big(\int_0^{1/n} dx \Big)^{1/p} = \Big(\frac{1}{n}\Big)^{1/p} \to 0$$

für $n \to \infty$. Für $\varepsilon = 1$ gilt also

$$\forall\, \delta > 0 \quad \exists\, f_\delta,\ d_p(f_\delta, o) < \delta, \qquad d_\infty(f_\delta, o) \ge \varepsilon,$$

mithin ist in diesem Falle die Einbettung nicht stetig.

Analog zur Homöomorphie definieren wir

Definition 1.2.5: *Zwei metrische Räume (X, d_X) und (Y, d_Y) heißen* isometrisch, *wenn es eine bijektive Abbildung*

$$f \;:\; X \longrightarrow Y$$

gibt mit

$$\forall\, x_1, x_2 \in X \qquad d_X(x_1, x_2) = d_Y\big(f(x_1), f(x_2)\big).$$

f heißt dann Isometrie.

Isometrische Räume werden also in diesem Sinne nicht mehr unterschieden. Im folgenden wollen wir das Vollständigkeitsaxiom verwenden. Die Menge der Äquivalenzklassen von Cauchyfolgen soll also wohldefiniert sein. Wie in der Infinitesimalrechnung gilt dann der auf Felix Hausdorff zurückgehende

Satz 1.2.6: *Es sei (X, d_X) ein metrischer Raum. Dann gibt es einen vollständigen metrischen Raum (Y, d_Y) und einen dichten Teilraum (Y_0, d_Y), $Y_0 \subset Y$, mit*

$$(X, d_X) \quad \textit{und} \quad (Y_0, d_Y) \;\; \textit{sind isometrisch.}$$

(Y, d_Y) heißt *Vervollständigung* von (X, d_X). In diesem Sinne unterscheiden wir X und Y_0 nicht mehr.

Der Begriff der Metrik läßt sich durch Abschwächung der Voraussetzungen erweitern. Hierzu folgende

Definition 1.2.7:

1. *Eine Abbildung*
$$d \;:\; M \times M \longrightarrow [0, \infty]$$
mit den Eigenschaften (1)-(3) einer Metrik heißt Fastmetrik.
2. *Eine Abbildung $d : M \times M \to \mathbb{R}_0^+$ mit den Eigenschaften*
(1′) $x = y \Longrightarrow d(x, y) = 0$
sowie (2) und (3) heißt Halbmetrik.

Eine Menge mit einer Fast- oder Halbmetrik wird *fastmetrischer* oder *halbmetrischer Raum* genannt.

Es folgen noch zwei Bemerkungen:

1. Aus einem fastmetrischen Raum kann man leicht einen metrischen machen, man wähle nur

$$d^* := \frac{d}{1+d}.$$

Dann ist d^* eine Metrik. Eigenschaften (1) und (2) einer Metrik sind erfüllt. Wir zeigen die Dreiecksungleichung: Für $r \geq 0$ sei

$$f(r) := \frac{r}{1+r}.$$

Dann gilt für alle $h \geq 0$

$$f(r+h) = \frac{r}{1+r+h} + \frac{h}{1+r+h} \leq f(r) + f(h)$$

und

$$f'(r) = \frac{1}{(1+r)^2} > 0.$$

Also wächst f monoton, und es folgt

$$d^*(x,y) = f\big(d(x,y)\big) \leq f\big(d(x,z) + d(z,y)\big) \leq f\big(d(x,z)\big) + f\big(d(z,y)\big) = d^*(x,z) + d^*(z,y).$$

Die Räume (M, d) und (M, d^*) sind homöomorph. Es ist nämlich

$$d^* \leq d$$

und

$$\forall\, d^*,\; d^* \leq \frac{1}{2}, \qquad d = \frac{d^*}{1-d^*} \leq 2d^*.$$

Daraus folgt die Stetigkeit von

$$id \;:\; (M, d) \longrightarrow (M, d^*)$$

und

$$id \;:\; (M,d^*) \longrightarrow (M,d),$$

denn im ersten Fall wähle man $\delta := \varepsilon$ und im zweiten $\delta := \frac{1}{2}\min\left(\frac{1}{2},\varepsilon\right)$.

2. Es seien (X,d_X) und (Y,d_Y) zwei metrische Räume. Dann definiert man für $Z := X \times Y$

$$d_Z(z_1,z_2) := d_X(x_1,x_2) + d_Y(y_1,y_2)$$

mit $z_i = \langle x_i, y_i\rangle$. Die spitzen Klammern sollen auf das direkte Produkt hindeuten.

Beispiel 1.2.8: *Es seien (X,d_X) ein metrischer Raum und $Z := X \times X$. Dann ist*

$$\begin{aligned} f \;:\; Z &\longrightarrow [0,\infty), \\ \langle x,y\rangle &\longmapsto d_X(x,y) \end{aligned}$$

gleichmäßig stetig.

Beweis: Es sei $z_i := \langle x_i, y_i\rangle$ für $i = 1,2$. Dann ist

$$\big|f(z_1) - f(z_2)\big| = \big|d_X(x_1,y_1) - d_X(x_2,y_2)\big| \le d_X(x_1,x_2) + d_X(y_1,y_2) = d_Z(z_1,z_2),$$

und es folgt mit $\delta = \varepsilon$

$$\forall\,\varepsilon > 0 \quad \exists\,\delta > 0 \quad \forall\, z_1,z_2;\, d_Z(z_1,z_2) < \delta, \qquad \big|f(z_1) - f(z_2)\big| < \varepsilon. \qquad \square$$

1.2.3 Räume von Abbildungen

Es seien E eine Menge, (Y,d_Y) ein metrischer Raum und $\mathcal{F}(E,Y)$ die Menge der Abbildungen

$$f \;:\; E \longrightarrow Y.$$

Wir benutzen die topologische Struktur von Y, um auch auf $\mathcal{F}(E,Y)$ eine Struktur zu erklären, und leiten Ergebnisse her. Anschließend werden wir für den Definitionsbereich E ebenfalls eine topologische Struktur voraussetzen. Wir wählen

$$\begin{aligned} d_{\mathcal{F}} \;:\; \mathcal{F}\times\mathcal{F} &\longrightarrow [0,\infty], \\ \langle f,g\rangle &\longmapsto \sup_{e\in E} d_Y\big(f(e),g(e)\big), \end{aligned}$$

also die *Supremumsmetrik*. Sie ist eine Fastmetrik und $(\mathcal{F},d_{\mathcal{F}})$ damit ein fastmetrischer Raum.

Definition 1.2.9: *Eine Abbildung $f \in \mathcal{F}(E,Y)$ heißt* beschränkt, *wenn*

$$\forall\, y \in Y \quad \exists\, r > 0 \quad \forall\, e \in E \qquad d_Y(f(e),y) \le r$$

gilt. Es sei $\mathcal{B}(E,Y) \subset \mathcal{F}(E,Y)$ die Menge der beschränkten Abbildungen.

Bemerkung 1.2.10:

1. Aus $f \in \mathcal{B}(E,Y)$ folgt

$$\forall\, z \in Y \quad \exists\, R > 0 \quad \forall\, e \in E \qquad f(e) \in \bar{B}(z,R)$$

und

$$\sup_{e_1,e_2\in E} d_Y\big(f(e_1),f(e_2)\big) < \infty.$$

2. $\mathcal{B}(E,Y)$ ist mit der von $\mathcal{F}(E,Y)$ induzierten Metrik ein metrischer Raum.

3. Y läßt sich mittels der konstanten Abbildung isometrisch in $\mathcal{F}(E,Y)$ einbetten.

Der Beweis der beiden ersten Aussagen folgt aus

$$d\big(f(e),z\big) \le d\big(f(e),y\big) + d(y,z) \le r + d(y,z) =: R_f(z)$$

und

$$d\big(f(e_1),f(e_2)\big) \le d\big(f(e_1),z\big) + d\big(z,f(e_2)\big) \le R_f(z) + R_g(z).$$

Zum Nachweis der dritten Aussage holen wir etwas aus: Es sei

$$\begin{aligned} c \;:\; Y &\longrightarrow \mathcal{F}(E,Y), \\ y &\longmapsto c(y) \end{aligned}$$

mit

$$\forall\, e \in E \qquad \big(c(y)\big)(e) := y$$

die konstante Abbildung. Wegen

$$d_{\mathcal{F}}\big(c(y_1),c(y_2)\big) = d_Y(y_1,y_2)$$

ist c eine Isometrie, und wir können Y mit $c(Y) \subset \mathcal{F}(E,Y)$ identifizieren. □

Lemma 1.2.11: *$c(Y)$ ist in $\mathcal{F}(E,Y)$ abgeschlossen.*

Zum Beweis zeigen wir, daß $\mathcal{F}(E,Y)\backslash c(Y)$ offen ist. Es sei $f \in \mathcal{F}(E,Y)\backslash c(Y)$. Dann gibt es $e_1, e_2 \in E$ mit $f(e_1) \neq f(e_2)$ und

$$\varepsilon := \frac{1}{2} d_Y\big(f(e_1),f(e_2)\big) > 0.$$

Es sei ferner $g \in \mathcal{F}(E,Y)$ mit $d_{\mathcal{F}}(f,g) < \varepsilon$. Dann folgt aus

$$2\varepsilon = d\big(f(e_1),f(e_2)\big) \le d\big(f(e_1),g(e_1)\big) + d\big(g(e_1),g(e_2)\big) + d\big(g(e_2),f(e_2)\big) < 2\varepsilon + d\big(g(e_1),g(e_2)\big)$$

$$d\big(g(e_1),g(e_2)\big) > 0,$$

also $g \notin c(Y)$. □

Folgerung 1.2.12: *Es sei $\mathcal{F}(E,Y)$ vollständig. Dann ist auch Y vollständig.*

Zum Beweis identifiziere man Y mit $c(Y)$ in $\mathcal{F}(E,Y)$. □

Es gilt auch die Umkehrung und damit

Satz 1.2.13: *$\mathcal{F}(E,Y)$ ist genau dann vollständig, wenn Y es ist.*

Zum Beweis haben wir nur noch eine Richtung zu zeigen. Y sei also vollständig und (f_n) eine Cauchyfolge in $\mathcal{F}(E,Y)$. Für jedes $e \in E$ ist dann erst recht $\big(f_n(e)\big)$ eine Cauchyfolge in Y, und es gibt ein $f(e) \in Y$ mit $f_n(e) \to f(e)$. Es sei

$$\begin{aligned} f \;:\; E &\longrightarrow Y, \\ e &\longmapsto f(e). \end{aligned}$$

Dann folgt aus

$$\forall\, \varepsilon > 0 \quad \exists\, n_0 \in \mathbb{N} \quad \forall\, n,m \ge n_0 \quad \forall\, e \in E \qquad d_Y\big(f_n(e),\, f_m(e)\big) < 2\varepsilon$$

für $m \to \infty$

$$\forall\, \varepsilon > 0 \quad \exists\, n_0 \in \mathbb{N} \quad \forall\, n \ge n_0 \quad \forall\, e \in E \qquad d_Y\big(f_n(e),\, f(e)\big) < \varepsilon,$$

also $d_{\mathcal{F}}(f_n,f) < \varepsilon$. □

Im folgenden benötigen wir das einfache

Lemma 1.2.14: *Es seien $b \in \mathcal{B}(E,Y)$ und $f \in \mathcal{F}(E,Y)$ mit $d_{\mathcal{F}}(b,f) < \infty$. Dann ist auch $f \in \mathcal{B}(E,Y)$.*

Der Beweis folgt aus

$$d\big(f(e),b(e_0)\big) \le d\big(f(e),b(e)\big) + d\big(b(e),b(e_0)\big) \le d_{\mathcal{F}}(b,f) + R\big(b(e_0)\big).$$ □

Wir zeigen nun die wichtige Aussage

Satz 1.2.15: *$\mathcal{B}(E,Y)$ ist eine sowohl offene als auch abgeschlossene Teilmenge von $\mathcal{F}(E,Y)$.*

Beweis:

1. $\mathcal{B}(E,Y)$ ist in $\mathcal{F}(E,Y)$ offen: Es seien $b \in \mathcal{B}(E,Y)$ und $f \in B(b,\varepsilon) \subset \mathcal{F}(E,Y)$. Dann folgt

$$d\big(f,b(e_0)\big) \le d(f,b) + d\big(b,b(e_0)\big) \le \varepsilon + R\big(b(e_0)\big),$$

   also $f \in \mathcal{B}(E,Y)$.

2. $\mathcal{B}(E,Y)$ ist in $\mathcal{F}(E,Y)$ abgeschlossen: Zum Nachweis zeigen wir, daß $\mathcal{F}(E,Y)\backslash\mathcal{B}(E,Y)$ offen ist. Es seien $g \in \mathcal{F}(E,Y)\backslash\mathcal{B}(E,Y)$ und $f \in B(g,\varepsilon) \subset \mathcal{F}(E,Y)$. Dann ist $f \in \mathcal{F}(E,Y)\backslash\mathcal{B}(E,Y)$, denn sonst würde aus Lemma 1.2.14 $g \in \mathcal{B}(E,Y)$ folgen. □

Folgerung 1.2.16: *Y sei vollständig. Dann ist auch $\mathcal{B}(E,Y)$ vollständig.*

Wir betrachten nun Abbildungen, deren Definitionsbereich E ebenfalls eine topologische Struktur trägt. Es sei E also von nun an ein topologischer Raum und (Y,d_Y) wieder ein metrischer Raum. Dann erklären wir wie üblich:

Definition 1.2.17: *Eine Abbildung $f : E \to Y$ heißt in $e \in E$* stetig, *wenn*

$$\forall\,\varepsilon > 0 \quad \exists\, U \in O(e) \qquad f(U) \subset B\big(f(e),\varepsilon\big)$$

ist. Dafür schreibt man $f \in \mathcal{C}_e(E,Y)$. Gilt dies für alle $e \in E$, dann heißt f in E stetig, also

$$f \in \mathcal{C}(E,Y) := \bigcap_{e\in E} \mathcal{C}_e(E,Y).$$

Es gilt

Satz 1.2.18: *$\mathcal{C}_e(E,Y)$ ist in $\mathcal{F}(E,Y)$ abgeschlossen.*

Zum Beweis zeigen wir, daß $\mathcal{F}(E,Y)\backslash\mathcal{C}_e(E,Y)$ offen ist. Es sei $f \in \mathcal{F}(E,Y)\backslash\mathcal{C}_e(E,Y)$, also

$$\exists\,\varepsilon > 0 \quad \forall\, U(e) \in O(e) \quad \exists\, e_u \in U(e) \qquad f(e_u) \notin B(f(e),\varepsilon).$$

Es sei ferner $g \in B(f,\frac{\varepsilon}{3}) \subset \mathcal{F}(E,Y)$. Dann ist

$$\varepsilon \le d\big(f(e),f(e_u)\big) \le d\big(f(e),g(e)\big) + d\big(g(e),g(e_u)\big) + d\big(g(e_u),f(e_u)\big) \le \frac{2\varepsilon}{3} + d\big(g(e),g(e_u)\big),$$

also

$$d\big(g(e),g(e_u)\big) \ge \frac{\varepsilon}{3}.$$

Mithin ist auch $g \in \mathcal{F}(E,Y)\backslash\mathcal{C}_e(E,Y)$. □

Folgerung 1.2.19: *$\mathcal{C}(E,Y)$ ist in $\mathcal{F}(E,Y)$ abgeschlossen.*

Folgerung 1.2.20: *Y sei vollständig. Dann ist auch $\mathcal{C}(E,Y)$ vollständig.*

1.2.4 Kompakte Mengen

Der Begriff der *Kompaktheit* tritt schon in den Vorlesungen über Infinitesimalrechnung auf. Er spielt eine große Rolle in der Analysis. Wir wollen ihn deshalb auch hier besprechen. Es sei im folgenden (X,d) ein metrischer Raum.

Definition 1.2.21: *$A \subset X$ heißt beschränkt , wenn*

$$\exists\, k > 0 \qquad \delta(A) := \sup_{x,y\in A} d(x,y) \le k$$

gilt. Man bezeichnet mit $\delta(A)$ den Durchmesser *von A.*

Definition 1.2.22: *Es seien*

$$\operatorname{dist}(x, A) := \inf_{a \in A} d(x, a)$$

der Abstand *zwischen $x \in X$ und $A \subset X$, sowie*

$$\operatorname{dist}(B, A) := \inf_{b \in B, a \in A} d(b, a)$$

der Abstand *zwischen $B \subset X$ und $A \subset X$.*

Definition 1.2.23: *Es seien Λ eine Indexmenge, $G_\lambda \subset X$ und*

$$A \subset \bigcup_{\lambda \in \Lambda} G_\lambda .$$

Dann heißt $(G_\lambda)_{\lambda \in \Lambda}$ Überdeckung *von A. Sie ist* endlich, *wenn Λ endlich ist; sie ist* offen, *wenn alle G_λ offen sind.*

Definition 1.2.24:

(1) X heißt kompakt, *wenn jede offene Überdeckung von X eine endliche Teilüberdeckung besitzt.*

(2) $A \subset X$ heißt kompakt, wenn A in der relativen Topologie kompakt ist. Dafür schreiben wir $A \Subset X$.

Es folgt leicht

Satz 1.2.25: *Es seien X kompakt, (Y, d_Y) ein metrischer Raum und $f : X \to Y$ stetig. Dann ist auch $f(X)$ kompakt.*

Beweis: Es seien $G_\lambda \subset Y$ offen, $\lambda \in \Lambda$, mit

$$f(X) \subset \bigcup_{\lambda \in \Lambda} G_\lambda .$$

Dann gilt

$$X \subset f^{-1}(f(X)) \subset f^{-1}\left(\bigcup_{\lambda \in \Lambda} G_\lambda \right) = \bigcup_{\lambda \in \Lambda} f^{-1}(G_\lambda).$$

Nun ist $f^{-1}(G_\lambda)$ offen und X kompakt, also

$$X \subset \bigcup_{i=1}^{k} f^{-1}(G_{\lambda_i})$$

für eine endliche Teilmenge $\{\lambda_1, \ldots, \lambda_k\} \in \Lambda$. Daraus folgt

$$f(X) \subset f\left(\bigcup_{i=1}^{k} f^{-1}(G_{\lambda_i}) \right) \subset \bigcup_{i=1}^{k} f\left(f^{-1}(G_{\lambda_i})\right) \subset \bigcup_{i=1}^{k} G_{\lambda_i}.$$

Mithin ist $f(X)$ kompakt. □

In $\mathbb{R}^n$ ist eine Teilmenge genau dann kompakt, wenn sie beschränkt und abgeschlossen ist. Das ist im allgemeinen sicherlich falsch, wie die folgenden Beispiele zeigen.

Beispiel 1.2.26: *Es sei X eine unendliche Menge mit diskreter Metrik. X ist abgeschlossen und wegen $\delta(X) = 1$ beschränkt. Die Kugeln $B(x, \frac{1}{2}) = \{x\}$ sind offen, endlich viele überdecken X aber nicht.*

Etwas weniger ausgefallen ist

Beispiel 1.2.27: *Es sei S die Sphäre in ℓ^2. Sie ist beschränkt und abgeschlossen, jedoch nicht kompakt.*

Das sieht man leicht: Es seien wie üblich e_j die Einheits-Basisvektoren. Dann schreiben wir $x \in \ell^2$ als $x = (\xi_1, \xi_2, \ldots)$ oder

$$x = \sum_{i=1}^{\infty} \xi_i e_i,$$

und es ist

$$d(x,y) := \sqrt{\sum_{i=1}^{\infty} |\xi_i - \eta_i|^2}.$$

Es gilt

$$d(e_i, e_j) = \sqrt{2} \quad \text{für } i \neq j$$

und für beliebige $x, y \in S$

$$d(x,y) \leq d(x,o) + d(o,y) = 2.$$

Es sei nun $\varepsilon < \sqrt{2}/2$ gewählt und mit $\Lambda := S$

$$G_\lambda := B(\lambda, \varepsilon).$$

Offensichtlich überdecken diese G_λ die Sphäre S. Es gebe nun eine endliche Teilüberdeckung

$$G_{s_1}, G_{s_2}, \ldots, G_{s_k}.$$

Dann existieren ein $s \in \{s_1, s_2, \cdots s_k\}$ und zwei verschiedene e_i, e_j mit

$$e_i \in G_s \quad \text{und} \quad e_j \in G_s.$$

Das führt aber zum Widerspruch

$$\sqrt{2} = d(e_i, e_j) \leq d(e_i, s) + d(s, e_j) < 2\varepsilon < \sqrt{2}. \qquad \square$$

Die eine Richtung der Charakterisierung kompakter Mengen in $\mathbb{R}^n$ gilt jedoch weiterhin, nämlich

Satz 1.2.28: *Es sei $A \Subset X$. Dann ist A beschränkt und abgeschlossen.*

Beweis:

1. A ist beschränkt: Aus

$$A \subset \bigcup_{\lambda \in A} B(\lambda, 1)$$

folgt die Existenz von $\lambda_1, \ldots, \lambda_k \in A$ mit

$$A \subset \bigcup_{i=1}^{k} B(\lambda_i, 1).$$

Für alle $x, y \in A$ ist deshalb

$$d(x,y) \leq d(x, \lambda_i) + d(\lambda_i, \lambda_j) + d(\lambda_j, y) \leq 2 + \alpha$$

mit

$$\alpha := \max_{i,j=1,\ldots,k} d(\lambda_i, \lambda_j) < \infty.$$

Mithin ist A beschränkt.

2. A ist abgeschlossen: Wir zeigen, daß $X \backslash A$ offen ist: Es sei $x \in X \backslash A$. Dann gilt

$$\forall\, a \in A \quad \exists\, \varepsilon_a > 0 \qquad B(x, \varepsilon_a) \cap B(a, \varepsilon_a) = \emptyset,$$

man wähle etwa $\varepsilon_a := d(x,a)/2$. Weil die $B(a, \varepsilon_a)$ die kompakte Menge A überdecken, folgt mit geeigneten $a_1, \ldots, a_k$ und $\varepsilon_i := \varepsilon_{a_i}$

$$A \subset \bigcup_{i=1}^{k} B(a_i, \varepsilon_i).$$

Es sei nun

$$\varepsilon := \min_{i=1,\ldots,k} \varepsilon_i > 0.$$

Dann ist

$$\forall\, i \in \{1, \ldots, k\} \qquad B(x, \varepsilon) \cap B(a_i, \varepsilon_i) = \emptyset,$$

und es folgt $B(x, \varepsilon) \cap A = \emptyset$. Die offene Menge $B(x, \varepsilon)$ liegt also in $X \backslash A$. $\square$

Damit haben wir eine Richtung der Charakterisierung kompakter Mengen in $\mathbb{R}^n$ auf den allgemeineren Fall übertragen können. Wir wissen aber, daß die andere Richtung falsch ist. Um trotzdem ein analoges Kriterium formulieren zu können, wollen wir die auftretenden Begriffe etwas modifizieren.

Das ist leicht, soweit es die Abgeschlossenheit betrifft, sie läßt sich abschwächen.

Definition 1.2.29: *$A \subset X$ heißt* relativ kompakt, *wenn $\bar{A}$ kompakt ist.*

Schwieriger und wichtiger ist die Modifizierung der Beschränktheit. Beispiel 1.2.26 zeigt das.

Definition 1.2.30: *$A \subset X$ heißt* präkompakt (totalbeschränkt), *wenn*

$$\forall\, \varepsilon > 0 \quad \exists\, E \subset X, \textit{endlich}, \qquad A \subset \bigcup_{e \in E} B(e, \varepsilon)$$

gilt.

Man spricht auch von einem *endlichen ε-Netz* oder *ε-Gitter* $N_\varepsilon \subset X$ für $A \subset X$:

$$\forall\, a \in A \quad \exists\, n \in N_\varepsilon \qquad d(a,n) < \varepsilon.$$

Offenbar ist in $\mathbb{R}^n$ eine beschränkte Menge präkompakt. Man sieht sofort

Lemma 1.2.31: *A sei präkompakt. Dann ist A beschränkt.*

Zu $\varepsilon > 0$ gibt es nämlich eine endliche Überdeckung $\big(B(x_i, \varepsilon)\big)_{i=1,\dots,k}$ von A. Deshalb gilt für $a, b \in A$

$$d(a,b) \le d\big(a, x_{i_a}\big) + d\big(x_{i_a}, x_{i_b}\big) + d(x_{i_b}, b) \le 2\varepsilon + d(x_{i_a}, x_{i_b}) < \infty. \qquad \square$$

Schließlich definieren wir noch

Definition 1.2.32: *X heißt* folgenkompakt, *wenn jede Folge in X einen Häufungspunkt besitzt. $A \subset X$ heißt folgenkompakt, wenn A in der relativen Topologie folgenkompakt ist.*

Jede Folge in A soll also eine konvergente Teilfolge mit einem Grenzwert in A besitzen. Es sei hier schon bemerkt, daß in allgemeinen topologischen Räumen Begriffe wie kompakt – folgenkompakt oder vollständig – folgenvollständig verschieden sind. Das wird wichtig werden, wenn wir später schwache Topologien benutzen. In metrischen Räumen gilt aber der

Hauptsatz: *(X,d) sei ein metrischer Raum. Dann sind äquivalent:*

(1) X ist kompakt.

(2) X ist folgenkompakt.

(3) X ist vollständig und präkompakt.

Beweis:

(1) $\Rightarrow$ (2): Es sei daran erinnert, daß x Häufungspunkt der Folge (x_n) heißt, wenn für jedes $\varepsilon > 0$ die Menge

$$\big\{ n \in \mathbb{N} \;\big|\; x_n \in B(x, \varepsilon) \big\}$$

unendlich ist. Wir schließen indirekt. Es sei (x_n) eine Folge, die keinen Häufungspunkt besitzt. Das heißt,

$$\forall\, x \in X \quad \exists\, \varepsilon_x > 0 \qquad \big\{ n \in \mathbb{N} \;\big|\; x_n \in B(x, \varepsilon_x) \big\} \text{ ist endlich.}$$

Nun ist

$$\big\{ B(x, \varepsilon_x) \big\}_{x \in X}$$

eine offene Überdeckung von X. Mithin gibt es eine endliche Menge $E \subset X$ mit

$$X \subset \bigcup_{x \in E} B(x, \varepsilon_x).$$

Deshalb ist auch

$$\bigcup_{x \in E} \big\{ n \in \mathbb{N} \;\big|\; x_n \in B(x, \varepsilon_x) \big\}$$

endlich, andererseits aber gleich $\mathbb{N}$. Das ist ein Widerspruch.

(2) $\Rightarrow$ (3): Jede Cauchyfolge besitzt einen Grenzwert in X. Mithin ist X vollständig. Um zu zeigen, daß X präkompakt ist, nehmen wir an, es gäbe ein $\varepsilon > 0$ und dazu keine endliche ε-Überdeckung. Wir wählen x_n rekursiv mit $x_1 \in X$, $x_2 \notin B(x_1, \varepsilon)$ und

$$x_{n+1} \notin \bigcup_{k=1}^{n} B(x_k, \varepsilon).$$

Dann ist $d(x_i, x_j) \geq \varepsilon$ für $i \neq j$. Die Folge (x_n) besitzt also keinen Häufungspunkt.

(3) $\Rightarrow$ (1): Auch jetzt schließen wir indirekt. Es sei

$$(U_\lambda)_{\lambda \in \Lambda} \supset X$$

eine Überdeckung, die keine endliche Teilüberdeckung enthält. Wegen der Präkompaktheit gilt

$$\forall\, \varepsilon > 0 \quad \exists\, n_\varepsilon,\, x_i \qquad X = \bigcup_{i=1}^{n_\varepsilon} B(x_i, \varepsilon),$$

und wenigstens ein $B(x_{i_\varepsilon}, \varepsilon)$, $i \in \{1, \ldots, n_\varepsilon\}$ wird nicht endlich überdeckt. Wählen wir nun der Reihe nach $\varepsilon = 1, \frac{1}{2}, \frac{1}{3}, \ldots$, dann erhalten wir die Existenz folgender nicht endlich zu überdeckender Mengen

$$\begin{aligned}
B_1 &:= \quad B(y_1, 1) \qquad (\varepsilon = 1),\\
B_2 &:= \bigcap\nolimits_{j=1}^{2} B(y_j, \frac{1}{j}) \qquad (\varepsilon = \frac{1}{2}, y_2 \in B_1),\\
&\vdots\\
B_n &:= \bigcap\nolimits_{j=1}^{n} B(y_j, \frac{1}{j}) \qquad (\varepsilon = \frac{1}{n}, y_n \in B_{n-1}),\\
&\vdots
\end{aligned}$$

Für $m \leq n$ gilt dann

$$y_m, y_n \in B_m \subset B\left(y_m, \tfrac{1}{m}\right)$$

oder

$$d(y_m, y_n) \leq \frac{1}{m}.$$

Mithin ist (y_n) eine Cauchyfolge mit dem Grenzwert $y \in X$ und

$$\varepsilon_n := d(y_n, y) \to 0 \quad \text{für } n \to \infty.$$

Es sei $y \in U_{\lambda_0}$. Dann folgt für große n

$$B_n \subset B\left(y_n, \tfrac{1}{n}\right) \subset B\left(y, \tfrac{1}{n} + \varepsilon_n\right) \subset U_{\lambda_0}.$$

Das ist ein Widerspruch, denn B_n war nicht endlich zu überdecken. □

Folgerung 1.2.33: *(X, d) sei ein metrischer Raum und $A \subset X$. Dann sind äquivalent:*

(1) A ist relativ kompakt.

(2) A ist präkompakt und $\bar{A}$ vollständig.

Folgerung 1.2.34: *(X, d) sei ein vollständiger metrischer Raum und $A \subset X$. Dann sind äquivalent:*

(1) A ist relativ kompakt.

(2) A ist präkompakt.

Das ist die gewünschte Verallgemeinerung des entsprechenden Satzes für den $\mathbb{R}^n$.

1.2.5 Der Satz von Arzelà-Ascoli

In diesem Abschnitt wollen wir ein wichtiges Kompaktheitskriterium, nämlich den *Satz von Arzelà-Ascoli*, bereitstellen. Er ist nach Cesare Arzelà, 1847–1912, und Giulio Ascoli, 1843–1896, benannt. Es seien in diesem Abschnitt:

1. (X, d_X) ein kompakter metrischer Raum.
2. (Y, d_Y) ein metrischer Raum.
3. $\mathcal{C}(X, Y) \subset \mathcal{F}(X, Y)$ die stetigen Abbildungen von X nach Y mit

$$d_{\mathcal{C}}(f, g) := d_{\mathcal{F}}(f, g) := \sup_{x \in X} d_Y\big(f(x), g(x)\big).$$

Es seien ferner $F \subset \mathcal{C}(X, Y)$ und

$$F(x) := \big\{ f(x) \;\big|\; f \in F \big\} \subset Y.$$

Definition 1.2.35: *$F \subset \mathcal{C}(X,Y)$ heißt* gleichartig stetig, *wenn*

$$\forall\, x \in X \quad \forall\, \varepsilon > 0 \quad \exists\, \delta > 0 \quad \forall\, \xi \in B(x,\delta) \quad \forall\, f \in F \qquad f(\xi) \in B\big(f(x),\varepsilon\big)$$

gilt.

Dann zeigen wir den

Satz von Arzelà-Ascoli: *Es sei $F \subset \mathcal{C}(X,Y)$. Dann sind äquivalent:*

(1) F ist relativ kompakt.

(2) F ist gleichartig stetig, und für alle $x \in X$ ist $F(x)$ relativ kompakt.

Dieser Satz wurde gegen Ende des letzten Jahrhunderts für den $\mathbb{R}^n$ bewiesen. Man lernt ihn in einer Vorlesung über Infinitesimalrechnung etwa in folgender Form kennen:

Es seien $\bar{G} \Subset \mathbb{R}^n$, $f_n \in \mathcal{C}(\bar{G},\mathbb{R}^m)$ und (f_n) punktweise beschränkt und gleichartig stetig in $\bar{G}$. Dann gibt es eine Teilfolge $(f_{n'})$ und ein $f \in \mathcal{C}(\bar{G},\mathbb{R}^m)$ mit $\|f_{n'} - f\|_\infty \to 0$.

Wir geben nun einen Beweis des Satzes von Arzelà-Ascoli an:

(1) $\Rightarrow$ (2): Als erstes zeigen wir, daß F gleichartig stetig ist. Es seien $x_0 \in X$ und $\varepsilon > 0$ vorgegeben. Aus dem Hauptsatz folgt, daß $\bar{F}$ präkompakt ist. Es gibt also $f_1, \ldots, f_n \in \bar{F} \subset \mathcal{F}(X,Y)$ mit

$$F \subset \bar{F} \subset \bigcup_{i=1}^{n} B\left(f_i, \tfrac{\varepsilon}{3}\right).$$

Weil $\mathcal{C}(X,Y)$ in $\mathcal{F}(X,Y)$ abgeschlossen ist, ist auch $\bar{F} \subset \mathcal{C}(X,Y)$. Die f_i sind also in x_0 stetig. Das heißt, es gilt

$$\forall\, i \in \{1,\ldots,n\} \quad \exists\, \delta_i > 0 \quad \forall\, x,\, d(x,x_0) < \delta_i, \qquad d\big(f_i(x), f_i(x_0)\big) < \frac{\varepsilon}{3}.$$

Wir wählen nun $\delta := \min\{\delta_1,\ldots,\delta_n\} > 0$. Zusammen mit

$$\forall\, f \in F \quad \exists\, i \in \{1,\ldots,n\} \qquad f \in B\left(f_i, \tfrac{\varepsilon}{3}\right)$$

folgt dann für alle x mit $d(x,x_0) < \delta$ und alle $f \in F$

$$d\big(f(x),f(x_0)\big) \le d\big(f(x),f_i(x)\big) + d\big(f_i(x),f_i(x_0)\big) + d\big(f_i(x_0),f(x_0)\big) < \varepsilon.$$

F ist also gleichartig stetig. Für festes $x \in X$ sei nun

$$\begin{aligned} g \;:\; F &\longrightarrow F(x), \\ f &\longmapsto f(x). \end{aligned}$$

Wegen

$$d\big(g(f_1),g(f_2)\big) = d_Y\big(f_1(x),f_2(x)\big) \le d_{\mathcal{C}}(f_1,f_2)$$

ist g stetig. Aus Satz 1.2.25 folgt deshalb die relative Kompaktheit von $F(x)$.

(2) $\Rightarrow$ (1): Nach Folgerung 1.2.33 haben wir die Vollständigkeit von $\bar{F}$ und die Präkompaktheit von F zu zeigen. Wir beginnen mit der Vollständigkeit:

Es sei (f_n) eine Cauchyfolge in F. Für festes $x \in X$ ist $\overline{F(x)}$ kompakt. Die Folge $\big(f_n(x)\big)$ besitzt daher einen Häufungspunkt und als Cauchyfolge sogar einen Grenzwert in $\overline{F(x)} \subset Y$. Es sei

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

Es ist $f \in \mathcal{F}(X,Y)$, und weil (f_n) in F Cauchyfolge ist, folgt sogar

$$d_{\mathcal{F}}(f_n,f) \to 0.$$

Der Raum $\mathcal{C}(X,Y)$ ist in $\mathcal{F}(X,Y)$ abgeschlossen. Mithin ist $f \in \mathcal{C}(X,Y)$, und es gilt

$$f_n \to f \quad \text{in } \mathcal{C}(X,Y).$$

Wir kommen nun zum wichtigsten Teil des Beweises und zeigen die Präkompaktheit von F: Es sei $\varepsilon > 0$ vorgegeben. Dann folgt aus der gleichartigen Stetigkeit von F

$$\forall\, x \in X \quad \exists\, \delta_x > 0 \quad \forall\, y, d(y,x) < \delta_x, \quad \forall\, f \in F \qquad d\big(f(y), f(x)\big) < \frac{\varepsilon}{4}. \tag{$*$}$$

$\{B(x,\delta_x)\}_{x\in X}$ ist eine offene Überdeckung von X, und X ist kompakt. Mithin existieren $x_1, \ldots, x_n \in X$ mit

$$X \subset \bigcup_{i=1}^{n} B(x_i, \delta_{x_i}). \tag{$**$}$$

$F(x_i)$ ist relativ kompakt. Deshalb gilt

$$\exists\, y_1, \ldots, y_m \in \bigcup_{i=1}^{n} F(x_i) \qquad \bigcup_{i=1}^{n} F(x_i) \subset \bigcup_{j=1}^{m} B\big(y_j, \tfrac{\varepsilon}{4}\big). \tag{$***$}$$

Es sei nun Φ die Menge der Abbildungen

$$\varphi \;:\; \{1,\ldots,n\} \longmapsto \{1,\ldots,m\}.$$

Φ ist endlich. Es sei ferner

$$\begin{aligned} L \;:\; \Phi &\longrightarrow \mathcal{P}(F), \\ \varphi &\longmapsto L\varphi \end{aligned}$$

mit

$$L\varphi := \Big\{ f \in F \;\Big|\; \forall\, i,\, 1 \le i \le n, \quad d\big(f(x_i), y_{\varphi(i)}\big) < \frac{\varepsilon}{4} \Big\}.$$

Es sei nun $f \in F$. Wegen $(***)$ gibt es dann zu $f(x_i)$ ein y_j mit

$$d\big(f(x_i), y_j\big) < \frac{\varepsilon}{4}.$$

Es gibt also ein $\varphi \in \Phi$ mit $f \in L\varphi$, also

$$F \;=\; \bigcup_{\varphi\in\Phi} L\varphi.$$

Diese Vereinigung ist endlich. Damit haben wir eine endliche Überdeckung von F gefunden, aber noch keine Überdeckung mit Kugeln. Es fehlt also noch der Nachweis der Beschränktheit der einzelnen $L\varphi$. Es seien also $\varphi \in \Phi$, $f, g \in L\varphi$ und $x \in X$. Dann folgt aus $(**)$

$$\exists\, i \in \{1,\ldots,n\} \qquad d(x, x_i) < \delta_{x_i},$$

also zusammen mit $(*)$

$$d\big(f(x), g(x)\big) \le d\big(f(x), f(x_i)\big) + d\big(f(x_i), y_{\varphi(i)}\big) + d\big(y_{\varphi(i)}, g(x_i)\big) + d\big(g(x_i), g(x)\big) < \varepsilon.$$

Mithin ist $d_{\mathcal{C}}(f,g) < \varepsilon$, und das war noch zu zeigen. □

Der Satz von Arzelà-Ascoli zählt zu den wichtigsten Grundlagen der Analysis. Er wird vielfältig verwandt. Erinnert sei an die Beweise zum Peanoschen Existenzsatz in der Theorie gewöhnlicher Differentialgleichungen oder zum Riemannschen Abbildungssatz. Weitere Anwendungen werden im Laufe dieser Vorlesung folgen.

1.3 Normierte Räume und stetige lineare Abbildungen

Bisher haben wir voneinander getrennt algebraische und topologische Strukturen betrachtet. Es versteht sich von selbst, daß wir sehr viel mehr interessante Resultate erwarten, wenn wir beide Strukturen zusammenfügen. Das soll nun geschehen. Zu Beginn sei aber doch etwas gewarnt. Die entstehende schöne Theorie beschreibt weitgehend lineare Phänomene. Man sollte aber nicht vergessen, daß viele wichtige Fragen auf nichtlineare Probleme führen.

Wir beginnen mit

Definition 1.3.1: *Es seien X ein linearer Raum und $\mathcal{T}$ eine Topologie auf X. Dann heißt $(X, \mathcal{T})$ ein* topologisch linearer Raum, *wenn die Abbildungen*

$$\begin{aligned} + \;:\; & \mathcal{X} \times \mathcal{X} \longrightarrow \mathcal{X}, \\ & \langle x, y \rangle \longmapsto x + y \end{aligned}$$

und

$$\begin{aligned} \cdot \;:\; & \mathbb{K} \times \mathcal{X} \longrightarrow \mathcal{X}, \\ & \langle \alpha, x \rangle \longmapsto \alpha x \end{aligned}$$

stetig sind.

Dabei tragen $\mathbb{K}$ die übliche Topologie der reellen oder komplexen Zahlen und Produkträume die kartesische Produkttopologie. Weitere Eigenschaften topologisch linearer Räume findet man zum Beispiel in *Taylor & Lay [1980], S. 94f.*

1.3.1 Normierte Räume

Besonders wichtige topologisch lineare Räume sind die normierten Räume. Besitzt ein linearer Raum eine Metrik, dann erwarten wir von ihr die Eigenschaft

$$\begin{aligned} d(x + z, y + z) &= d(x, y), \\ d(\alpha x, \alpha y) &= |\alpha|\, d(x, y). \end{aligned}$$

Damit gelangt man zur *Norm*. Wir präzisieren:

Definition 1.3.2: *Es sei $\mathcal{X}$ ein Vektorraum über $\mathbb{K}$. Eine Abbildung*

$$\|\cdot\| \;:\; \mathcal{X} \longrightarrow \mathbb{R}_0^+$$

heißt Norm *auf $\mathcal{X}$, wenn*

(1) $\forall\, x \in \mathcal{X} \qquad \|x\| = 0 \iff x = o,$

(2) $\forall\, x \in \mathcal{X} \quad \forall\, \alpha \in \mathbb{K} \qquad \|\alpha x\| = |\alpha|\, \|x\|,$

(3) $\forall\, x, y \in \mathcal{X} \qquad \|x + y\| \leq \|x\| + \|y\|$

gilt. $(\mathcal{X}, \|\cdot\|)$ *heißt* normierter Vektorraum.

Entsprechend den Aussagen zur Metrik werden folgende Begriffe eingeführt.

Definition 1.3.3:

1. Man spricht von einer Halb- *oder* Seminorm, *wenn die Eigenschaft (1) der Norm nicht erfüllt ist.*

2. Gilt

$$\|\cdot\| \;:\; \mathcal{X} \longrightarrow [0, \infty],$$

dann spricht man von einer Fastnorm.

3. Eine Quasinorm *ist eine Abbildung*

$$\|\cdot\| \;:\; \mathcal{X} \longrightarrow [0, \infty)$$

mit den Eigenschaften (1), (3) der Norm und

(2′) $\|-x\| = \|x\|,\ \lim\limits_{\alpha_n \to 0} \|\alpha_n x\| = 0$ *und* $\lim\limits_{\|x_n\| \to 0} \|\alpha x_n\| = 0.$

Es werden teilweise auch Abschwächungen der Dreiecksungleichung (3) zugelassen, etwa

(3′) $\exists\, c \in \mathbb{R}^+ \quad \forall\, x, y \in \mathcal{X} \qquad \|x + y\| \leq c\,\{\|x\| + \|y\|\}.$

Bemerkung 1.3.4:

1. $(\mathcal{X}, \|\cdot\|)$ ist vermöge
$$d(x,y) := \|x-y\|$$
ein metrischer Raum. In diesem Sinne werden wir alle Ergebnisse über topologische Strukturen verwenden. Ein vollständiger normierter Raum heißt *Banachraum*, benannt nach STEFAN BANACH, 1892–1945.

2. Ein beliebiger metrischer Raum ist natürlich kein normierter Vektorraum. Augenfällige Beispiele sind
$$M := [0,1] \quad \text{mit } d(x,y) := |x-y|$$
oder
$$M := \mathbb{R} \quad \text{mit } d(x,y) := \frac{|x-y|}{1+|x-y|}.$$

3. Aus
$$|d(x,x_1) - d(y,y_1)| \le d(x,y) + d(x_1,y_1)$$
folgt mit $x_1 = y_1 = 0$
$$\big|\, \|x\| - \|y\| \,\big| \le \|x-y\| \le \|x\| + \|y\|.$$

4. Betrachten wir als Beispiel für eine Seminorm und eine Quasinorm
$$\mathcal{L}^1(\mathbb{R}) := \Big\{ \mathring{\mathcal{C}}_\infty(\mathbb{R});\ \|\cdot\| \Big\}^\sim \quad \text{mit } \|f\| := \int_{\mathbb{R}} |f(x)|\,dx.$$
$\mathcal{L}^1(\mathbb{R})$ ist ein Banachraum. Es sei
$$\mathcal{L}^1_{loc}(\mathbb{R}) := \Big\{ f \in \mathcal{F}(\mathbb{R},\mathbb{R}) \;\Big|\; \forall\, K \Subset \mathbb{R} \quad f \in \mathcal{L}^1(K) \Big\}.$$
Seminormen für $\mathcal{L}^1_{loc}(\mathbb{R})$ sind
$$p_n(f) := \int_{|x|<n} |f(x)|\,dx,$$
und eine Quasinorm ist
$$|\!|\!| f |\!|\!| := \sum_{n=1}^{\infty} \frac{p_n(f)}{2^n(1+p_n(f))}.$$
Ein vollständiger quasinormierter Raum ist ein *Fréchetraum*. Allgemeiner heißt ein vollständiger, lokal konvexer, metrischer linearer Raum *Fréchetraum*, benannt nach RENÉ MAURICE FRÉCHET, 1878–1973.

5. Unterräume eines normierten Vektorraumes werden in kanonischer Weise mit der induzierten Norm ausgestattet.

6. Es sei $\mathcal{M} \subset \mathcal{X}$ ein Unterraum. Dann kann man auf $\mathcal{X}$ eine Äquivalenzrelation definieren:
$$x \sim y \quad :\Longleftrightarrow \quad x - y \in \mathcal{M}.$$
Diese Relation ist reflexiv, transitiv und symmetrisch. Es seien $[x]$ die von x erzeugte Klasse und $\mathcal{X}/\mathcal{M}$ die Menge dieser Äquivalenzklassen. $\mathcal{X}/\mathcal{M}$ ist wieder ein Vektorraum,
$$\begin{aligned} [x]+[y] &:= [x+y], \\ \alpha[x] &:= [\alpha x] \end{aligned}$$
sind eindeutig definiert. Es ist
$$\begin{aligned} \varphi \;:\; \mathcal{X} &\longrightarrow \mathcal{X}/\mathcal{M}, \\ x &\longmapsto [x] \end{aligned}$$
die *kanonische Abbildung* von $\mathcal{X}$ auf den *Quotientenraum* $\mathcal{X}/\mathcal{M}$. Es sei wieder
$$\operatorname{dist}(x,\mathcal{M}) := \inf_{m\in\mathcal{M}} \|x-m\|$$
der Abstand von x zu $\mathcal{M}$. Dann gilt
$$\operatorname{dist}(x,\mathcal{M}) = 0 \iff x \in \bar{\mathcal{M}},$$
und wir erhalten

Lemma 1.3.5: *Es sei $\mathcal{M} \subset \mathcal{X}$ ein abgeschlossener Teilraum von $\mathcal{X}$. Dann ist*

$$\begin{aligned} \left\| \cdot \right\| \; : \; \mathcal{X}/\mathcal{M} &\longrightarrow \mathbb{R}_0^+, \\ x &\longmapsto \operatorname{dist}(x, \mathcal{M}) \end{aligned}$$

eine Norm auf $\mathcal{X}/\mathcal{M}$.

In diesem Sinne wird $\mathcal{X}/\mathcal{M}$ zu einem normierten Vektorraum.

Beweis: Die erste Eigenschaft einer Norm ist wegen $\mathcal{M} = \bar{\mathcal{M}}$ klar, (2) folgt mit $\alpha \neq 0$ aus

$$\left\| [\alpha x] \right\| = \inf_m \|\alpha x + m\| = \inf_m \|\alpha x + \alpha m\| = |\alpha| \inf_m \|x + m\| = |\alpha| \left\| [x] \right\|$$

und Eigenschaft (3) aus

$$\begin{aligned} \left\| [x] + [y] \right\| &= \inf_m \|x + y + m\| = \\ &= \inf_{m_x} \inf_{m_y} \|(x + m_x) + (y + m_y)\| \leq \inf_{m_x} \inf_{m_y} \Big\{ \|x + m_x\| + \|y + m_y\| \Big\} = \left\| [x] \right\| + \left\| [y] \right\|. \end{aligned}$$

□

Es gilt auch

Lemma 1.3.6: *Es seien $\mathcal{M} \subset \mathcal{X}$ ein abgeschlossener Teilraum von $\mathcal{X}$ und $\varphi : \mathcal{X} \to \mathcal{X}/\mathcal{M}$ die kanonische Abbildung. Dann gilt:*

(1) φ ist offen.

(2) φ ist stetig.

(3) Es sei $\mathcal{X}$ ein Banachraum. Dann ist $\mathcal{X}/\mathcal{M}$ ein Banachraum.

7. Für $i = 1, 2$ seien $(\mathcal{X}_i, \|\cdot\|_i)$ normierte Räume. Dann sind

$$\begin{aligned} \left\| \cdot \right\| \; : \; \mathcal{X}_1 \times \mathcal{X}_2 &\longrightarrow \mathbb{R}_0^+, \\ \langle x_1, x_2 \rangle &\longmapsto \|x_1\|_1 + \|x_2\|_2 \end{aligned}$$

oder

$$\langle x_1, x_2 \rangle \longmapsto \sqrt{\|x_1\|_1^2 + \|x_2\|_2^2}$$

Normen auf $\mathcal{X}_1 \times \mathcal{X}_2$.

8. Wir kennen bereits isomorphe und homöomorphe Räume. Entsprechend nennt man zwei normierte Vektorräume $\mathcal{X}, \mathcal{Y}$ *linear homöomorph* oder *topologisch isomorph*, wenn es eine homöomorphe Abbildung

$$A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$$

gibt. Die Räume $\mathcal{X}, \mathcal{Y}$ heißen *isometrisch isomorph* oder *kongruent*, wenn $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ isometrisch ist.

9. Eine Banachraum $\mathcal{X}$ besitzt eine *abzählbare Basis* oder *Schauder-Basis* $\{u_n\}_{n \in \mathbb{N}}$, wenn sich jedes $x \in \mathcal{X}$ eindeutig in der Form

$$x = \sum_{i=1}^{\infty} \xi_i u_i \quad \text{mit } \xi_i \in \mathbb{K}$$

darstellen läßt. Ein Banachraum $\mathcal{X}$ mit Schauder-Basis ist *separabel*, das heißt, $\mathcal{X}$ besitzt eine abzählbare dichte Teilmenge. Umgekehrt besitzen die meisten separablen Banachräume eine Schauder-Basis; es gibt jedoch separable Banachräume ohne Schauderbasis. Die Schauderbasis ist nach Juliusz Pawel Schauder, 1896–1943, benannt.

1.3.2 Beispiele

1. Es sei $\mathcal{X}$ der $\mathbb{R}^n$ oder $\mathbb{C}^n$. Es seien ferner

$$\{e_1, e_2, \ldots, e_n\}$$

eine Basis von $\mathcal{X}$ und

$$x = \sum_{i=1}^{n} \xi_i \, e_i, \quad y = \sum_{i=1}^{n} \eta_i \, e_i, \qquad \xi_i \, , \eta_i \in \mathbb{K}$$

Darstellungen von $x, y \in \mathcal{X}$. In $\mathcal{X}$ hat man die Normen

$$\begin{aligned} \|x\|_p &:= \left(|\xi_1|^p + \cdots + |\xi_n|^p\right)^{1/p} \quad \text{für } 1 \le p < \infty, \\ \|x\|_\infty &:= \max\{|\xi_1|, \ldots, |\xi_n|\} \quad \text{für } p = \infty. \end{aligned}$$

Es ist

$$\|x\|_\infty \le \|x\|_p \le \sqrt[p]{n}\, \|x\|_\infty \underset{p\to\infty}{\longrightarrow} \|x\|_\infty.$$

Mit

$$1 = \frac{1}{p} + \frac{1}{q}$$

gilt die *Höldersche Ungleichung*

$$\sum_{i=1}^{n} |\xi_i \eta_i| \le \left(\sum_{i=1}^{n} |\xi_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} |\eta_i|^q\right)^{\frac{1}{q}}.$$

Für $p = 2$ ist das die *Schwarzsche Ungleichung*. Daraus folgt die Dreiecksungleichung

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

Man nennt sie auch *Minkowskische Ungleichung*.

2. Entsprechend werden die normierten Folgenräume ℓ^p über $\mathbb{R}$ oder $\mathbb{C}$ definiert, nämlich für $x = (\xi_1, \xi_2, \ldots)$

$$\ell^p = \left\{ x = (\xi_i) \;\middle|\; \|x\|_p < \infty \right\}$$

mit

$$\|x\|_p := \begin{cases} \left(\sum_{i=1}^{\infty} |\xi_i|^p\right)^{1/p} & \text{für } 1 \le p < \infty, \\ \sup_{i\in\mathbb{N}} |\xi_i| & \text{für } p = \infty. \end{cases}$$

3. Es seien E eine Menge und $\mathcal{Y}$ ein normierter $\mathbb{K}$-Vektorraum. Dann wird $\mathcal{F}(E, \mathcal{Y})$ vermöge

$$(f + g)(x) := f(x) + g(x), \qquad (\alpha f)(x) := \alpha f(x)$$

zu einem $\mathbb{K}$-Vektorraum. Eine Fastnorm wird mit

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

auf $\mathcal{F}(E, \mathcal{Y})$ definiert. Mit $\mathcal{B}(E, \mathcal{Y}) \subset \mathcal{F}(E, \mathcal{Y})$ bezeichnet man den normierten Raum der *beschränkten Abbildungen*, speziell $\ell^\infty = \mathcal{B}(\mathbb{N}, \mathbb{K})$.

Ist E ein topologischer Raum, dann sind $\mathcal{C}(E, \mathcal{Y})$ bzw. $\mathcal{BC}(E, \mathcal{Y})$ die Räume der stetigen bzw. beschränkten und stetigen Abbildungen. $\mathcal{BC}(E, \mathcal{Y})$ wird mit der Norm $\|f\|_{\mathcal{BC}} := \|f\|_\infty$ versehen.

4. Es sei $\Omega \subset \mathbb{R}^n$ eine offene Menge. Dann ist $\mathcal{L}^p(\Omega, \mathbb{K}) \subset \mathcal{F}(\Omega, \mathbb{K})$ die Menge der *im Lebesgueschen Sinne über Ω integrablen Funktionen* mit

$$\|f\|_{\mathcal{L}^p} := \begin{cases} \left(\int_\Omega |f(x)|^p dx\right)^{1/p} & \text{für } 1 \le p < \infty, \\ \operatorname{ess\,sup}_{x\in\Omega} |f(x)| & \text{für } p = \infty. \end{cases}$$

Mit ess sup ist das fast überall genommene Supremum gemeint, also

$$\operatorname*{ess\,sup}_{x\in\Omega} |f(x)| := \inf_{N,\, |N|=0} \sup_{x\in\Omega\setminus N} |f(x)|.$$

5. $\mathcal{BV}([a, b], \mathbb{K})$ sind die Funktionen von *beschränkter Variation* über $[a, b]$. Es seien $f \in \mathcal{F}([a, b], \mathbb{K})$,

$$P_n := \left\{ (x_0, \ldots, x_n) \in \mathbb{R}^{n+1} \;\middle|\; a = x_0 < x_1 < \cdots < x_n = b \right\}$$

die Menge der Partitionen n-ten Grades von $[a,b]$ und mit $p \in P_n$

$$V(f,p) := \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|$$

Variation von f. Dann heißt f von *beschränkter Variation*, wenn die *totale Variation*

$$V(f) := \sup_{n\in\mathbb{N}} \sup_{p\in P_n} V(f,p) < \infty$$

ist. $V(f)$ ist eine Halbnorm.

$$\|f\|_{\mathcal{BV}} := |f(a)| + V(f)$$

ist eine Norm für $\mathcal{BV}([a,b],\mathbb{K})$.

1.3.3 Stetige lineare Abbildungen

Wir betrachten nun stetige lineare Abbildungen. Die Grundlage bildet

Satz 1.3.7: *Es seien $\mathcal{X},\mathcal{Y}$ normierte Räume und $A \in \mathcal{L}(\mathcal{X},\mathcal{Y})$. Dann gelten:*

(1) A ist überall stetig oder nirgends.
(2) A ist genau dann stetig, wenn folgendes gilt

$$\exists\, c \geq 0 \quad \forall\, x \in \mathcal{X} \qquad \|Ax\|_{\mathcal{Y}} \leq c\,\|x\|_{\mathcal{X}}. \tag{$*$}$$

Beweis:

1. A sei in $x_0 \in X$ stetig, das heißt

$$\forall\,\varepsilon > 0 \quad \exists\,\delta > 0 \quad \forall\, x,\ \|x - x_0\| < \delta, \qquad \|Ax - Ax_0\| < \varepsilon.$$

Es sei $x_1 \in X$ beliebig und $x \in B(x_1,\delta)$. Dann gilt

$$\|Ax - Ax_1\| = \|A(x - x_1 + x_0) - Ax_0\| < \varepsilon$$

wegen $\|(x - x_1 + x_0) - x_0\| = \|x - x_1\| < \delta$. A ist also auch in x_1 stetig.

2. Ist $(*)$ erfüllt, dann ist A im Nullpunkt stetig, also überall.

Umgekehrt, es sei A im Nullpunkt stetig. Dann gibt es ein $\delta = \delta(1)$ mit

$$\forall\, x,\ \|x\| < \delta, \qquad \|Ax\| < 1.$$

Es sei $x \neq 0$ und

$$x_0 := \frac{\delta}{2\|x\|} x.$$

Dann ist $\|x_0\| < \delta$, $\|Ax_0\| < 1$, und es folgt

$$\|Ax\| = \frac{2\|x\|}{\delta}\,\|Ax_0\| < \frac{2}{\delta}\,\|x\|.$$

Das heißt, es ist $\|A\| < 2/\delta$. □

Üblicherweise nennt man die Menge der stetigen linearen Abbildungen deshalb *beschränkt linear*. Man schreibt $\mathcal{L}_b(\mathcal{X},\mathcal{Y})$ oder sogar $\mathcal{L}(\mathcal{X},\mathcal{Y})$ dafür. Wir werden jedoch auch in der Notation die linearen Abbildungen von den stetigen linearen unterscheiden und $\mathcal{CL}(\mathcal{X},\mathcal{Y})$ dafür verwenden. Es ist also

$$\mathcal{CL}(\mathcal{X},\mathcal{Y}) := \Big\{ A \in \mathcal{L}(\mathcal{X},\mathcal{Y}) \;\Big|\; \sup_{x\in S_{\mathcal{X}}} \|Ax\|_{\mathcal{Y}} < \infty \Big\}.$$

Dabei ist $S_{\mathcal{X}} := \{x \in \mathcal{X} \mid |x| = 1\}$ die Einheitssphäre in $\mathcal{X}$. Mit

$$\|A\|_{\mathcal{CL}} := \sup_{x\in S_{\mathcal{X}}} \|Ax\|_{\mathcal{Y}}$$

wird $\mathcal{CL}(\mathcal{X},\mathcal{Y})$ selbst zu einem normierten Raum. $\|\cdot\|_{\mathcal{CL}}$ nennt man auch *Operatornorm*.

Natürlich gibt es unbeschränkt lineare Abbildungen. Gerade die Differentialoperatoren sind unbeschränkt. Denken Sie zum Beispiel an $\mathcal{X} := \mathcal{C}_1([0,1],\mathbb{R})$ und $\mathcal{Y} := \mathcal{C}([0,1],\mathbb{R})$, beide versehen mit der Supremumsnorm. Dann ist $A \in \mathcal{L}(\mathcal{X},\mathcal{Y})$, $Ax := x'$, ein unbeschränkt linearer Operator. Mit $f_n(x) := x^n$ ist nämlich

$$\|f_n\| = 1 \quad \text{und} \quad \|Af_n\| = n.$$

Aus Satz 1.3.7 folgt

Satz 1.3.8: *Es sei $A \in \mathcal{L}(\mathcal{X},\mathcal{Y})$. Dann sind äquivalent:*

(1) $A^{-1} \in \mathcal{L}(\mathcal{R}(A),\mathcal{X})$ *existiert und ist stetig.*
(2) $\exists\, m>0 \quad \forall\, x \in \mathcal{X} \quad m\|x\|_{\mathcal{X}} \le \|Ax\|_{\mathcal{Y}}$.

Beweis:
(1) $\Rightarrow$ (2) folgt mit $y = Ax$ aus

$$\|A^{-1}y\| \le c\,\|y\| \quad \text{mit } c>0.$$

(2) $\Rightarrow$ (1): A ist injektiv. Deshalb existiert A^{-1}, und es folgt $\|A^{-1}y\| \le \frac{1}{m}\|y\|$.

Folgerung 1.3.9: *Die normierten Räume $\mathcal{X},\mathcal{Y}$ sind genau dann* linear homöomorph, *wenn es eine surjektive Abbildung $A \in \mathcal{L}(\mathcal{X},\mathcal{Y})$ und $m,M>0$ mit*

$$\forall\, x \in \mathcal{X} \qquad m\,\|x\|_{\mathcal{X}} \le \|Ax\|_{\mathcal{Y}} \le M\,\|x\|_{\mathcal{X}}$$

gibt.

Definition 1.3.10: *Zwei Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ auf $\mathcal{X}$ heißen* äquivalent, *wenn*

$$\exists\, m,M>0, \quad \forall\, x \in \mathcal{X} \qquad m\,\|x\|_1 \le \|x\|_2 \le M\,\|x\|_1$$

gilt.

Folgerung 1.3.11: *Die Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ auf $\mathcal{X}$ definieren genau dann dieselbe Topologie, wenn sie äquivalent sind.*

Es sei noch folgendes bemerkt:

1. Für $A \in \mathcal{CL}(\mathcal{X},\mathcal{Y})$ gilt auch

$$\|A\| = \sup_{x\neq o} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\|\le 1} \|Ax\| \quad \text{und} \quad \|A\| = \inf\left\{ c \in \mathbb{R}_0^+ \;\middle|\; \forall\, x \in \mathcal{X} \quad \|Ax\| \le c\,\|x\| \right\}.$$

2. Aus $A \in \mathcal{CL}(\mathcal{X},\mathcal{Y})$ und $B \in \mathcal{CL}(\mathcal{Y},\mathcal{Z})$ folgen $BA \in \mathcal{CL}(\mathcal{X},\mathcal{Z})$ und

$$\|BA\| \le \|A\| \cdot \|B\|.$$

3. $\mathcal{Y}$ sei ein Banachraum. Dann ist auch $\mathcal{CL}(\mathcal{X},\mathcal{Y})$ ein Banachraum. Es sei nämlich (A_n) eine $\mathcal{CL}$-Cauchyfolge. Dann existiert zunächst punktweise

$$\lim_{n\to\infty} A_n x =: Ax \in \mathcal{Y}.$$

A ist linear, und aus

$$\forall\, \varepsilon>0 \quad \exists\, n_0 \in \mathbb{N} \quad \forall\, n,m \ge n_0, \qquad \sup_{\|x\|=1} \|A_n x - A_m x\| < \varepsilon$$

folgt

$$\forall\, \varepsilon>0 \quad \exists\, n_0 \in \mathbb{N} \quad \forall n \ge n_0 \qquad \sup_{\|x\|=1} \|A_n x - Ax\| < \varepsilon.$$

Es sei nun $\mathcal{M}$ ein Unterraum von $\mathcal{X}$. Dann ist auch $\bar{\mathcal{M}}$ ein Unterraum. Es gibt natürlich nicht abgeschlossene Unterräume, denken Sie an $\mathcal{C}_1([0,1]) \subset \mathcal{C}([0,1])$. Wichtig ist nun das folgende Resultat über *stetige Fortsetzungen*. Dabei sei an Kapitel 1.1.2 erinnert.

Satz 1.3.12: *Es seien $\mathcal{X}$ ein normierter Raum, $\mathcal{M} \subset \mathcal{X}$ ein Unterraum, $\mathcal{Y}$ ein Banachraum und $A \in \mathcal{CL}(\mathcal{M},\mathcal{Y})$. Dann gilt*

$$\overset{1}{\exists} \;\; \bar{A} \in \mathcal{CL}(\bar{\mathcal{M}},\mathcal{Y}) \quad \textit{mit } \bar{A}\,|\,\mathcal{M} = A \textit{ und } \|\bar{A}\| = \|A\|.$$

Ein analoges Resultat wird in der Infinitesimalrechnung bewiesen. Wir skizzieren den Beweis: Zu $x \in \bar{\mathcal{M}}$ gibt es eine gegen x konvergente Folge (x_n) mit $x_n \in \mathcal{M}$. Wegen

$$\|Ax_n - Ax_m\| \le \|A\| \cdot \|x_n - x_m\|$$

ist (Ax_n) ebenfalls eine Cauchyfolge und

$$\bar{A}x := \lim_{n\to\infty} Ax_n$$

existiert. Dann ist $\bar{A}$ wohldefiniert und linear, insbesondere gilt

$$\bar{A} \,|\, \mathcal{M} = A.$$

Wir zeigen $\|\bar{A}\| = \|A\|$ und damit die Stetigkeit von $\bar{A}$: Aus

$$\|Ax_n\| \le \|A\| \cdot \|x_n\|$$

und $x_n \to x$ folgt $\|\bar{A}\| \le \|A\|$. Andererseits ist $\|A\| \le \|\bar{A}\|$, weil $\bar{A}$ die Abbildung A fortsetzt. □

Bemerkung 1.3.13: *Ist $\mathcal{Y}$ nur Unterraum eines Banachraumes, dann folgt $\bar{A} \in \mathcal{CL}(\bar{\mathcal{M}}, \bar{\mathcal{Y}})$.*

In diesem Zusammenhang sei noch an folgendes erinnert: Es sei $\mathcal{X}$ ein Vektorraum. Dann ist in $\mathcal{L}(\mathcal{X}, \mathcal{X})$ durch die Verknüpfung zweier Abbildungen auch eine Multiplikation erklärt. Ein Vektorraum $\mathcal{X}$, in dem eine Multiplikation so definiert ist, daß $(\mathcal{X}, +, \cdot\,)$ ein Ring mit Einselement id und

$$\forall\, x, y \in \mathcal{X} \quad \forall\, \alpha \in \mathbb{K} \qquad \alpha(x\,y) = (\alpha x)y = x(\alpha y)$$

ist, heißt *Algebra*. Eine Algebra ist *kommutativ*, wenn

$$\forall\, x, y \in \mathcal{X} \qquad xy = yx$$

gilt.

Eine *normierte Algebra* ist eine Algebra, versehen mit einer Norm $\|\cdot\|$, die die Eigenschaften

(1) $\forall\, x, y \in \mathcal{X} \qquad \|xy\| \le \|x\| \cdot \|y\|$,

(2) $\|id\| = 1$

hat. Eine *Banachalgebra* ist eine vollständige normierte Algebra.

Ist $\mathcal{X}$ ein Banachraum, dann ist $\mathcal{CL}(\mathcal{X}, \mathcal{X})$ also eine Banachalgebra.

1.3.4 Lineare Homöomorphismen

In diesem Abschnitt sei $\mathcal{X}$ ein Banachraum. $\mathcal{CL}(\mathcal{X}) := \mathcal{CL}(\mathcal{X}, \mathcal{X})$ ist dann eine Banachalgebra. Wir wollen uns im folgenden mit der Frage nach der Existenz von Inversen beschäftigen und beginnen mit dem Nachweis der Konvergenz der nach Carl Neumann, 1832–1925, benannten Reihe.

Satz 1.3.14: *Es seien $\mathcal{X}$ ein Banachraum und $A \in \mathcal{CL}(\mathcal{X})$ mit $\|A\| < 1$. Dann ist*

$$id - A$$

ein linearer Homöomorphismus, und es gilt

$$\left\|(id - A)^{-1}\right\| \le \frac{1}{1 - \|A\|}.$$

Dabei ist $(id - A)^{-1} = \sum_{i=0}^{\infty} A^i$ die Neumannsche Reihe *zu A. Sie konvergiert in der Operatornorm.*

Beweis: Es seien $A^0 := id$, $A^n := A \circ A^{n-1}$ und

$$S_n := \sum_{i=0}^{n} A^i.$$

Dann ist (S_n) in $\mathcal{CL}(\mathcal{X})$ eine Cauchyfolge. Der Raum $\mathcal{CL}(\mathcal{X})$ ist selbst ein Banachraum, folglich existiert ein $S \in \mathcal{CL}(\mathcal{X})$ mit $S_n \to S$. Aus

$$S_n \circ (id - A) = \sum_{i=0}^{n} A^i - \sum_{i=1}^{n+1} A^i = id - A^{n+1} \to id$$

und

$$(id - A) \circ S_n \to id$$

erhalten wir deshalb

$$S \circ (id - A) = (id - A) \circ S = id.$$

Das beweist den Satz. □

Satz 1.3.15: *Es sei Ω die Menge der linearen Homöomorphismen von $\mathcal{X}$ auf $\mathcal{X}$. Dann gilt:*
(1) $\Omega \subset \mathcal{CL}(\mathcal{X})$ ist offen.
(2) Die Abbildung

$$\begin{aligned} \Omega &\longrightarrow \Omega, \\ A &\longmapsto A^{-1} \end{aligned}$$

ist stetig.

Beweis: Es seien $S \in \Omega$ und $T \in \mathcal{CL}(\mathcal{X}, \mathcal{X})$ mit $T \in B\left(S, \|S^{-1}\|^{-1}\right)$. Dann folgt

$$\|id - S^{-1}T\| \leq \|S^{-1}\| \cdot \|S - T\| < 1, \qquad (*)$$

also

$$S^{-1}T = id - (id - S^{-1}T) \in \Omega$$

oder $T = SS^{-1}T \in \Omega$. Die Menge Ω ist also offen.

Zum Nachweis der zweiten Aussage seien $S, T \in \Omega$. Dann ist

$$\|T^{-1} - S^{-1}\| = \|T^{-1}(S - T)S^{-1}\| \leq \|T^{-1}\| \cdot \|S^{-1}\| \cdot \|S - T\| < \varepsilon$$

für alle T mit

$$T \in B\Big(S, \frac{1}{2\|S^{-1}\|}\Big) \bigcap B\Big(S, \frac{\varepsilon}{2\|S^{-1}\|^2}\Big).$$

Das folgt aus

$$\begin{aligned} \|T^{-1}\| \leq \|S^{-1}\| \cdot \|T^{-1}S\| = \|S^{-1}\| \cdot \|(S^{-1}T)^{-1}\| = \\ = \|S^{-1}\| \cdot \left\| \Big(id - (id - S^{-1}T)\Big)^{-1} \right\| < \frac{\|S^{-1}\|}{1 - \frac{1}{2}} = 2\left\|S^{-1}\right\|. \end{aligned}$$

□

Bemerkung 1.3.16: *In Satz 1.3.14 ist die Voraussetzung $\|A\| < 1$ für die erste Aussage nicht notwendig. Zur Konvergenz der Reihe $\sum A^n$ genügt wie bei Potenzreihen*

$$\limsup_{n \to \infty} \left(\|A^n\|\right)^{1/n} < 1.$$

Hierzu geben wir eine Anwendung.

Die Volterrasche Integralgleichung: Es seien $\mathcal{X} := \mathcal{C}([a, b], \mathbb{R})$,

$$D := \big\{(s, t) \;\big|\; a \leq t \leq s \leq b\big\} \subset \mathbb{R}^2$$

sowie $k \in \mathcal{C}(D, \mathbb{R})$, $M := \|k\|_\infty < \infty$ und

$$\begin{aligned} K \;:\; \mathcal{X} &\longrightarrow \mathcal{X}, \\ x &\longmapsto Kx \end{aligned}$$

mit

$$(Kx)(s) := \int_a^s k(s, t)x(t)\, dt.$$

Offenbar ist $K \in \mathcal{CL}(\mathcal{X})$ mit $\|K\| \leq M(b - a)$.

Es sei nun ein $y \in \mathcal{X}$ gegeben. Gesucht ist ein $x \in \mathcal{X}$ mit

$$(id - K)x = y. \qquad (*)$$

Zum Nachweis der Existenz einer Lösung x zeigen wir, daß die Reihe

$$x := \sum_{n=0}^{\infty} K^n y$$

konvergiert. Dazu beweist man zunächst durch vollständige Induktion

$$\forall\, n \in \mathbb{N} \quad \forall\, x \in \mathcal{X} \quad \forall\, s \in [a,b] \qquad |(K^n x)(s)| \le M^n \frac{(s-a)^n}{n!} \|x\|_\infty$$

oder

$$\forall\, n \in \mathbb{N} \qquad \left\|K^n\right\| \le M^n \, \frac{(b-a)^n}{n!}.$$

Die Neumannsche Reihe zu K wird also von der e-Reihe majorisiert. Sie konvergiert deshalb, und wir erhalten

$$\|x\|_\infty \le e^{M(b-a)} \, \|y\|_\infty .$$

Damit haben wir zugleich die Eindeutigkeit der Lösung und die Abschätzung

$$\|(id - K)^{-1}\| \le e^{M(b-a)}$$

gewonnen. Benannt werden diese Integralgleichungen nach Vito Volterra, 1860–1940.

1.3.5 Endlichdimensionale Räume

Aus Vorlesungen über Infinitesimalrechnung ist bekannt, daß die Sphären

$$\left\{ x \in \mathbb{K}^{\mathrm{n}} \;\middle|\; |x| = 1 \right\},$$

$$\left\{ x \in \mathbb{K}^{\mathrm{n}} \;\middle|\; \|x\|_\infty = 1 \right\}$$

abgeschlossen und präkompakt sind. $\mathbb{K}^{\mathrm{n}}$ ist vollständig. Damit sind diese Sphären kompakt. Wir wollen zeigen, daß alle Sphären in endlich dimensionalen Räumen diese Eigenschaft besitzen und daß umgekehrt aus der Kompaktheit der Sphäre in einem normierten Raum $\mathcal{X}$ auch $\dim \mathcal{X} < \infty$ folgt. Wir beginnen mit

Satz 1.3.17: *Es seien $(\mathcal{X}, \|\cdot\|_\mathcal{X})$ und $(\mathcal{Y}, \|\cdot\|_\mathcal{Y})$ normierte Räume gleicher endlicher Dimension n über $\mathbb{K}$. Dann sind $\mathcal{X}$ und $\mathcal{Y}$ linear homöomorph.*

Beweis: O.B.d.A. genügt es, den Satz für $(\mathcal{Y}, \|\cdot\|_\mathcal{Y}) := (\mathbb{K}^{\mathrm{n}}, \|\cdot\|_\infty)$ zu zeigen. Es seien

$$(k_1, \ldots, k_n) \text{ bzw. } (x_1, \ldots, x_n)$$

Basen von $\mathbb{K}^{\mathrm{n}}$ bzw. $\mathcal{X}$ und $T \in \mathcal{L}(\mathbb{K}^{\mathrm{n}}, \mathcal{X})$ definiert durch

$$T\, k_i := x_i .$$

Die Abbildung T ist ein Isomorphismus. Mit

$$k = \sum_{i=1}^{n} \kappa_i k_i \quad \text{und} \quad x = \sum_{i=1}^{n} \xi_i x_i$$

und $M := \sum_{i=1}^{n} \|x_i\|$ folgt

$$\|Tk\| \le \sum_{i=1}^{n} |\kappa_i| \, \|Tk_i\| \le \|k\|_\infty \left(\sum_{i=1}^{n} \|x_i\| \right) = M \, \|k\|_\infty .$$

T ist also stetig. Wir betrachten auf der Sphäre

$$S := \left\{ k \in \mathbb{K}^{\mathrm{n}} \;\middle|\; \|k\|_\infty = 1 \right\}$$

die Abbildung

$$\begin{aligned} S &\longrightarrow \mathbb{R}_0^+, \\ k &\longmapsto \|Tk\|_\mathcal{X}. \end{aligned}$$

Diese Abbildung ist als Verknüpfung zweier stetiger Abbildungen selbst stetig, nimmt somit auf dem Kompaktum S ihr Minimum an. Es gilt also

$$\exists\, m \ge 0 \quad \forall\, k \in S \qquad \|Tk\| \ge m.$$

Es ist sogar $m > 0$, denn aus $m = 0$ würde die Existenz eines $k_0 = \sum_{i=1}^{n} \kappa_i k_i \in S$ mit $Tk_0 = o$ folgen. Dann wäre aber $\sum_{i=1}^{n} \kappa_i x_i = o$ und damit $\kappa_i = 0$ für alle i im Widerspruch zu $k_0 \in S$. Es folgt also

$$\exists\, m > 0 \quad \forall\, k \in \mathbb{K}^{\mathrm{n}} \qquad \|Tk\| \ge m \|k\|_\infty ,$$

mithin ist auch T^{-1} stetig. □

Folgerung 1.3.18:

(1) Es sei $\mathcal{X}$ ein endlichdimensionaler Vektorraum. Dann sind alle Normen auf $\mathcal{X}$ äquivalent und erzeugen dieselbe Topologie.

(2) Es seien $\mathcal{X}, \mathcal{Y}$ normierte Räume und $\dim \mathcal{X} < \infty$. Dann ist $\mathcal{L}(\mathcal{X}, \mathcal{Y}) = \mathcal{CL}(\mathcal{X}, \mathcal{Y})$.

(3) Es seien $\mathcal{X}$ ein normierter Raum und $\mathcal{M}$ ein Teilraum mit $\dim \mathcal{M} = n$. Dann ist $\mathcal{M}$ topologisch isomorph zu $\mathbb{K}^n$, also vollständig und abgeschlossen.

Es sei noch einmal herausgestellt

Satz 1.3.19: *Es sei $(\mathcal{X}, \|\cdot\|)$ ein endlichdimensionaler normierter Raum. Dann ist die Sphäre*

$$S_{\mathcal{X}} := \{x \in \mathcal{X} \mid \|x\| = 1\}$$

kompakt.

$S_{\mathcal{X}}$ ist nämlich ein homöomorphes Bild der entsprechenden Sphäre in $\mathbb{K}^n$. Wichtig ist nun, daß auch die Umkehrung dieser Aussage gilt, zusammen also

Satz 1.3.20: *Es sei $(\mathcal{X}, \|\cdot\|)$ ein normierter Raum. Dann gilt*

$$\dim \mathcal{X} < \infty \iff S_{\mathcal{X}} \text{ ist kompakt.}$$

Zum Beweis holen wir etwas aus und zeigen zunächst

Das Rieszsche Lemma: *Es seien $\mathcal{X}$ ein normierter Raum und $\mathcal{X}_0 \subsetneq \mathcal{X}$ ein abgeschlossener Unterraum. Dann gilt*

$$\forall\, q \in (0,1) \quad \exists\, x_q \in S_{\mathcal{X}} \subset \mathcal{X} \qquad q \le \operatorname{dist}(x_q, \mathcal{X}_0) \le 1.$$

Beweis: Es seien $x \in \mathcal{X} \backslash \mathcal{X}_0$ und $d := \operatorname{dist}(x, \mathcal{X}_0) > 0$. Dann gibt es ein $x_0 \in \mathcal{X}_0$ mit

$$d \le \|x - x_0\| \le \frac{d}{q}.$$

Wir wählen

$$x_q := \frac{x - x_0}{\|x - x_0\|}.$$

Dann ist $x_q \in S_{\mathcal{X}}$, und es folgt für alle $y \in \mathcal{X}_0$

$$\|x_q - y\| = \frac{1}{\|x - x_0\|} \left\| x - \left(x_0 + y\,\|x - x_0\|\right) \right\| \ge \frac{q}{d} \cdot d = q. \qquad \square$$

Wir beweisen nun den noch offenen Teil von Satz 1.3.20: $S_{\mathcal{X}}$ sei also kompakt. Wir nehmen $\dim \mathcal{X} = \infty$ an und wählen ein $x_1 \in S_{\mathcal{X}}$. Es sei $\mathcal{X}_1 := \operatorname{span}(x_1)$. $\mathcal{X}_1$ ist ein abgeschlossener Unterraum von $\mathcal{X}$. Folglich existiert ein $x_2 \in S_{\mathcal{X}}$ mit $\operatorname{dist}(x_2, \mathcal{X}_1) \ge \frac{1}{2}$. Es sei $\mathcal{X}_2 := \operatorname{span}(x_1, x_2)$. So fahren wir fort und wählen $x_{n+1} \in S_{\mathcal{X}}$ mit

$$\operatorname{dist}(x_{n+1}, \mathcal{X}_n) \ge \frac{1}{2}.$$

Es gibt also eine Folge (x_n) mit $x_n \in S_{\mathcal{X}}$ und

$$\forall\, n, m \in \mathbb{N}; n \ne m, \qquad \|x_n - x_m\| \ge \frac{1}{2}.$$

Diese Folge enthält keine konvergente Teilfolge. Das steht im Widerspruch zur Kompaktheit von $S_{\mathcal{X}}$. $\square$

Schließlich zeigen wir an einem Beispiel, daß die Aussage im Rieszschen Lemma für $q = 1$ allgemein nicht mehr gilt.

Beispiel 1.3.21: *Es sei wieder $\mathcal{X}_0 \subsetneq \mathcal{X}$ ein abgeschlossener Teilraum. Dann existiert im allgemeinen kein $x_1 \in S_{\mathcal{X}}$ mit*

$$\operatorname{dist}(x_1, \mathcal{X}_0) = 1.$$

Es sei nämlich

$$\mathcal{X} := \{x \in \mathcal{C}([0,1],\mathbb{K}) \mid x(0) = 0\}$$

und

$$\begin{aligned} I \;:\; \mathcal{X} &\longrightarrow \mathbb{K}, \\ x &\longmapsto \int_0^1 x(t)\,dt. \end{aligned}$$

Dann ist I eine lineare stetige Abbildung mit $\|I\| \leq 1$, also ist

$$\mathcal{X}_0 := \mathcal{N}(I)$$

ein abgeschlossener Teilraum von $\mathcal{X}$. Es gilt:

1. $\|I\| = 1$. Denn für $x_n(t) := t^{1/n}$ gilt $Ix_n = n/(n+1) \to 1$.
2. $\forall\, x \in S_{\mathcal{X}}$ ist $|Ix| < 1$. Denn x ist stetig mit $x(0) = 0$.
3. $\forall\, x \in \mathcal{X}$ ist $|Ix| = \operatorname{dist}(x, \mathcal{X}_0)$. Für $x_0 \in \mathcal{X}_0$ ist nämlich

$$|Ix| = |I(x - x_0)| \leq \|x - x_0\|,$$

also $|Ix| \leq \operatorname{dist}(x, \mathcal{X}_0)$. Es sei wieder $x_n(t) := t^{1/n}$. Dann ist

$$x - \frac{Ix}{Ix_n} x_n \in \mathcal{X}_0$$

und

$$\left\| x - \left(x - \frac{Ix}{Ix_n} x_n \right) \right\| = \frac{|Ix|}{|Ix_n|} \|x_n\| = \frac{n+1}{n}\,|Ix|,$$

also $\operatorname{dist}(x, \mathcal{X}_0) \leq |Ix|$.

4. Es gebe nun ein $x_1 \in \mathcal{X}$ mit $\|x_1\| = 1$ und

$$\operatorname{dist}(x_1, \mathcal{X}_0) = 1.$$

Dann folgt der Widerspruch

$$1 = \operatorname{dist}(x_1, \mathcal{X}_0) = |Ix_1| < 1.$$

□

1.4 Hilberträume

In diesem Abschnitt wollen wir spezielle normierte Räume betrachten, deren Normen sich aus einem Skalarprodukt ableiten.

1.4.1 Skalarprodukte

Definition 1.4.1: *Es sei $\mathcal{X}$ ein $\mathbb{K}$-Vektorraum. Dann heißt*

$$(\cdot,\cdot) \;:\; \mathcal{X} \times \mathcal{X} \longrightarrow \mathbb{K}$$

ein Skalarprodukt, *wenn folgendes gilt:*

(1) $\forall\, x, y, z \in \mathcal{X} \quad \forall\, \alpha, \beta \in \mathbb{K} \quad (\alpha x + \beta y, z) = \alpha(x,z) + \beta(y,z)$.

(2) $\forall\, x, y \in \mathcal{X} \quad (x,y) = \overline{(y,x)}$.

(3) $\forall\, x \in \mathcal{X} \quad (x,x) \geq 0$.

(4) $\forall\, x \in \mathcal{X},\, x \neq o, \quad (x,x) > 0$.

$\big(\mathcal{X}, (\cdot,\cdot)\big)$ *heißt* Prähilbertraum.

Das Skalarprodukt ist eine *Sesquilinearform*. Wir notieren einige Eigenschaften:

1. $(\cdot,\cdot)$ ist im zweiten Argument konjugiert linear.
2. Im Falle $\mathbb{K} = \mathbb{R}$ ist $(\cdot,\cdot)$ symmetrisch.

3. Durch
$$\|x\| := \sqrt{(x,x)}$$
wird auf $\mathcal{X}$ eine Norm definiert. In diesem Sinne ist $\mathcal{X}$ dann ein normierter Vektorraum, und es gilt die *Cauchy-Schwarzsche Ungleichung*
$$\forall\, x,y \in \mathcal{X} \qquad |(x,y)| \le \|x\| \cdot \|y\|.$$

4. Es gilt der *Satz von Pythagoras*
$$\forall\, x,y \in \mathcal{X}, (x,y) = 0, \qquad \|x+y\|^2 = \|x\|^2 + \|y\|^2.$$

5. Es gilt die *Parallelogrammgleichung*
$$\forall\, x,y \in \mathcal{X} \qquad \|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right).$$

6. Es sei $\mathbb{K} = \mathbb{R}$. Dann gilt die *Polarisationsformel*
$$\forall\, x,y \in \mathcal{X} \qquad (x,y) = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right).$$
Es sei $\mathbb{K} = \mathbb{C}$. Dann gilt entsprechend
$$\forall\, x,y \in \mathcal{X} \qquad (x,y) = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + \mathrm{i}\|x+\mathrm{i}y\|^2 - \mathrm{i}\|x-\mathrm{i}y\|^2\right).$$
Das Skalarprodukt läßt sich also durch *Polarisieren* aus der *induzierten Norm* zurückgewinnen.

7. Es sei $\mathcal{X}$ ein normierter Raum, dessen Norm der Parallelogrammgleichung genügt. Dann wird durch das Polarisieren ein Skalarprodukt erklärt, das wiederum diese Norm induziert.

Definition 1.4.2: *Zwei Elemente $x,y \in \mathcal{X}$ heißen* orthogonal, *$x \perp y$, wenn $(x,y) = 0$ ist. Es sei $S \subset \mathcal{X}$ eine beliebige Teilmenge. Dann heißt*
$$S^\perp := \left\{x \in \mathcal{X} \;\middle|\; \forall\, s \in S \quad (x,s) = 0\right\}$$
das orthogonale Komplement *von S.*

Man sieht sofort, daß $S^\perp$ ein abgeschlossener Teilraum von $\mathcal{X}$ ist und daß $S^{\perp\perp} = S$ für abgeschlossene Teilräume S gilt. Der Satz von Pythagoras besagt
$$x \perp y \implies \|x+y\|^2 = \|x\|^2 + \|y\|^2.$$
Aus
$$\|x+y\|^2 - \|x\|^2 - \|y\|^2 = 2\,\mathrm{Re}\,(x,y)$$
folgt im Falle $\mathbb{K} = \mathbb{R}$ auch die Umkehrung dieser Aussage. Im Falle $\mathbb{K} = \mathbb{C}$ ist das aber falsch, man wähle nur $x = 1$ und $y = \mathrm{i}$. Dann ist $(x,y) = -\mathrm{i}$ und
$$|x+y|^2 = |1+\mathrm{i}|^2 = 2 = |x|^2 + |y|^2.$$

Wir halten noch fest:

Lemma 1.4.3: *Es sei $\mathcal{X}$ ein Prähilbertraum. Dann ist das Skalarprodukt*
$$(\cdot,\cdot) \;:\; \mathcal{X} \times \mathcal{X} \longrightarrow \mathbb{K}$$
eine stetige Abbildung auf $\mathcal{X} \times \mathcal{X}$.

Beweis: Es sei $\langle x,y\rangle \in \mathcal{X} \times \mathcal{X}$ und $\mathcal{X} \times \mathcal{X}$ kanonisch mit der Norm
$$\|\langle x,y\rangle\|_{\mathcal{X}\times\mathcal{X}} := \sqrt{\|x\|_{\mathcal{X}}^2 + \|y\|_{\mathcal{X}}^2}$$

versehen. Dann gilt

$$\begin{aligned}\big|(x,y)-(x_0,y_0)\big| &= \big|(x-x_0,y-y_0)+(x-x_0,y_0)+(x_0,y-y_0)\big| \\ &\le \|x-x_0\|\Big\{\frac{1}{2}\|y-y_0\|+\|y_0\|\Big\}+\|y-y_0\|\Big\{\frac{1}{2}\|x-x_0\|+\|x_0\|\Big\} \\ &\le \Big(\|x-x_0\|+\|y-y_0\|\Big)\Big\{\|x_0\|+\|y_0\|+\frac{1}{2}\Big[\|x-x_0\|+\|y-y_0\|\Big]\Big\}.\end{aligned}$$

Es sei nun $\|\langle x-x_0,y-y_0\rangle\| \le \sqrt{2}$. Dann ist mit $k := 1+\|x_0\|+\|y_0\|$

$$\big|(x,y)-(x_0,y_0)\big| \le \sqrt{2}k\,\|\langle x-x_0,y-y_0\rangle\| = \sqrt{2}k\,\|\langle x,y\rangle-\langle x_0,y_0\rangle\| < \varepsilon$$

für alle $\langle x,y\rangle$ mit $\|\langle x,y\rangle-\langle x_0,y_0\rangle\| < \delta := \sqrt{2}\min\big(1,\frac{\varepsilon}{2k}\big)$. □

Definition 1.4.4: *Ein bezüglich der induzierten Norm vollständiger Prähilbertraum heißt* Hilbertraum.

Natürlich gibt es viele Beispiele für Hilbert- oder Prähilberträume. Erinnert sei nur an

1. $\mathcal{X} = \mathbb{R}^n$ oder $\mathcal{X} = \mathbb{C}^n$ mit $(x,y) := \sum_{i=1}^n x_i\,\bar{y}_i$,
2. ℓ^2 oder $\ell^2(n)$,
3. $\mathcal{L}^2(I)$ mit $(f,g) := \int_I f(x)\overline{g(x)}\,dx$,
4. $\mathcal{C}([0,1])$ mit dem $\mathcal{L}^2$-Skalarprodukt. Dies ist nur ein Prähilbertraum, wie wir wissen. Die vom $\mathcal{L}^2$-Skalarprodukt induzierte Norm ist nicht zur $\|\cdot\|_\infty$-Norm von $\mathcal{C}([0,1])$ äquivalent (vgl. §1.2.2).

1.4.2 Der Approximationssatz

In der Theorie der Hilberträume und für viele Anwendungen spielt der Approximationssatz eine außerordentlich große Rolle. Obwohl er aus einer Vorlesung über Infinitesimalrechnung bekannt sein dürfte, beweisen wir ihn deshalb auch hier.

Definition 1.4.5: *Es sei $\mathcal{X}$ ein Vektorraum. Dann heißt eine Teilmenge $M \subset \mathcal{X}$* konvex, *wenn*

$$\forall\, x,y \in M \quad \forall\, \lambda \in [0,1] \qquad \lambda x + (1-\lambda)y \in M$$

gilt.

Der Durchschnitt konvexer Mengen ist konvex, jeder Unterraum von $\mathcal{X}$ ist konvex.

Approximationssatz: *Es seien $\mathcal{X}$ ein Hilbertraum und M eine abgeschlossene konvexe Teilmenge von $\mathcal{X}$. Dann gilt*

$$\forall\, x \in \mathcal{X} \quad \overset{1}{\exists}\, m \in M \qquad \|x-m\| = \operatorname{dist}(x,M).$$

Zum Beweis des Satzes wählen wir eine *Minimalfolge* (m_i) aus M mit

$$\lim_{i\to\infty}\|m_i - x\| = d := \operatorname{dist}(x,M).$$

Aus der Parallelogrammgleichung

$$\|u+v\|^2 + \|u-v\|^2 = 2\left(\|u\|^2+\|v\|^2\right)$$

folgt dann mit $u := x-m_i$ und $v := x-m_j$

$$\underbrace{4\Big\|x-\frac{m_i+m_j}{2}\Big\|^2}_{\ge 4d^2} + \|m_i-m_j\|^2 = 2\left(\|x-m_i\|^2+\|x-m_j\|^2\right) \to 4\,d^2.$$

Mithin ist (m_i) eine Cauchyfolge. Weil $\mathcal{X}$ vollständig und M abgeschlossen ist, existiert ein $m \in M$ mit $m_i \to m$.

Der Grenzwert m ist eindeutig bestimmt. Es seien nämlich m_1 und m_2 zwei Bestapproximierende. Dann folgt wieder aus der Parallelogrammgleichung

$$\underbrace{4\Big\|x-\frac{m_1+m_2}{2}\Big\|^2}_{\ge 4d^2} + \|m_1-m_2\|^2 = 4\,d^2.$$

Also ist $m_1 = m_2$. □

Aus dem Approximationssatz folgt unmittelbar der

Projektionssatz: *Es seien $\mathcal{X}$ ein Hilbertraum und M ein abgeschlossener Teilraum. Dann gilt*

$$\mathcal{X} = M \oplus M^\perp,$$

also

$$\forall\, x \in \mathcal{X} \quad \overset{1}{\exists}\, m \in M \quad \overset{1}{\exists}\, m^\perp \in M^\perp \qquad x = m + m^\perp.$$

Zum Beweis wähle man m aus dem Approximationssatz und $m^\perp := x - m$. Dann folgt für alle $y \in M$ mit $\|y\| = 1$ und mit $\lambda := (m^\perp, y)$

$$||m^\perp||^2 \leq ||m^\perp - \lambda y||^2 = ||m^\perp||^2 - 2\,\mathrm{Re}\,\bar{\lambda}(m^\perp, y) + |\lambda|^2,$$

also

$$0 \leq -|(m^\perp, y)|^2.$$

Mithin ist $m^\perp \in M^\perp$. Aus

$$x = m_1 + m_1^\perp = m_2 + m_2^\perp$$

folgt

$$0 = \underbrace{(m_1 - m_2)}_{\in M} + \underbrace{(m_1^\perp - m_2^\perp)}_{\in M^\perp},$$

also

$$0 = \|m_1 - m_2\|^2 + \|m_1^\perp - m_2^\perp\|^2,$$

und damit die Eindeutigkeit. □

1.4.3 Orthonormalsysteme und separable Räume

Kenntnisse über Orthonormalsysteme aus einer Vorlesung über Infinitesimalrechnung werden im folgenden als bekannt vorausgesetzt. Hier sollen nur einige Resultate kurz zusammengestellt werden.

Definition 1.4.6: *Es seien $\mathcal{X}$ ein Prähilbertraum, Λ eine Indexmenge und $u_\lambda \in \mathcal{X}$ für $\lambda \in \Lambda$. Dann heißt*

$$\{u_\lambda\}_{\lambda \in \Lambda} \subset \mathcal{X}$$

ein Orthonormalsystem, *wenn*

$$\forall\, \mu, \nu \in \Lambda \qquad (u_\mu, u_\nu) = \delta_{\mu,\nu} := \begin{cases} 1 & \text{für } \mu = \nu \\ 0 & \text{für } \mu \neq \nu \end{cases}$$

ist.

Verschiedene Elemente eines Orthonormalsystems sind stets linear unabhängig. Erinnert sei an das *Schmidtsche Orthogonalisierungsverfahren*, benannt nach Erhard Schmidt, 1876–1959, nämlich

Satz 1.4.7: *Es seien $\mathcal{X}$ ein Prähilbertraum und $\{x_i\}_{i \in \mathbb{N}}$ eine Familie linear unabhängiger Elemente in $\mathcal{X}$. Dann gibt es ein Orthonormalsystem*

$$\{u_i\}_{i \in \mathbb{N}}$$

mit

$$\forall\, n \in \mathbb{N} \qquad \mathrm{span}(x_1, \ldots, x_n) = \mathrm{span}(u_1, \ldots, u_n).$$

Beweis: Es seien $u_1 := x_1/\|x_1\|$ und $u_2, \ldots, u_n$ bereits konstruiert. Dann betrachten wir

$$y_{n+1} := x_{n+1} - \sum_{i=1}^{n} (x_{n+1}, u_i) u_i.$$

Es ist $y_{n+1} \neq 0$ und

$$\forall\, i = 1, 2, \ldots, n \qquad (y_{n+1}, u_i) = 0.$$

Mithin können wir definieren

$$u_{n+1} := \frac{y_{n+1}}{\|y_{n+1}\|}.$$

□

Die Besselsche Ungleichung: Es seien nun $OS \subset \mathcal{X}$ ein Orthonormalsystem und $(u_i)_{i\in\mathbb{N}}$ eine Folge aus OS. Dann gilt die *Besselsche Ungleichung*

$$\forall\, x \in \mathcal{X} \quad \forall\, n \in \mathbb{N} \qquad \sum_{i=1}^{n} |(x,u_i)|^2 \le \|x\|^2.$$

Mit $\xi_i := (x,u_i)$ folgt der Beweis unmittelbar aus

$$0 \le \left\| x - \sum_{i=1}^{n} \xi_i u_i \right\|^2 = \|x\|^2 - 2\,\mathrm{Re} \sum_{i=1}^{n} \bar{\xi}_i (x,u_i) + \sum_{i=1}^{n} |\xi_i|^2 = \|x\|^2 - \sum_{i=1}^{n} |\xi_i|^2 \,. \qquad \square$$

Folgerung 1.4.8: *Für jedes $x \in \mathcal{X}$ ist die Menge*

$$\left\{ u \in OS \;\middle|\; (x,u) \neq 0 \right\}$$

höchstens abzählbar unendlich.

Beweis: Weil $\sum_{i=1}^{n} |\xi_i|^2$ konvergiert, sind die Mengen

$$A_n := \left\{ u \in OS \;\middle|\; |(x,u)| \ge \frac{1}{n} \right\}$$

endlich. Mithin ist

$$\left\{ u \in OS \;\middle|\; (x,u) \neq 0 \right\} = \bigcup_{n=1}^{\infty} A_n$$

als abzählbare Vereinigung endlicher Mengen höchstens abzählbar. $\square$

Deshalb ist

$$\sum_{u\in OS} (x,u)\, u$$

höchstens eine abzählbare Summe und wegen der Besselschen Ungleichung damit wohldefiniert.

Es seien nun $x, y \in \mathcal{X}$. Dann gibt es $(u_i)_{i\in\mathbb{N}}$ mit

$$\left\{ u \in OS \;\middle|\; (x,u) \neq 0 \text{ und } (y,u) \neq 0 \right\} = \left\{ u_1, u_2, u_3, \dots \right\}.$$

Es ist also auch

$$\sum_{u\in OS} |(x,u)\cdot\overline{(y,u)}| = \sum_{i=1}^{\infty} |(x,u_i)\cdot\overline{(y,u_i)}| \le \sqrt{\sum_{i=1}^{\infty} |(x,u_i)|^2} \cdot \sqrt{\sum_{i=1}^{\infty} |(y,u_i)|^2} \le \|x\|\cdot\|y\|.$$

Im speziellen Fall *separabler Räume* sind alle Orthonormalsysteme abzählbar.

Definition 1.4.9: *Ein metrischer Raum $(\mathcal{X}, d)$ heißt* separabel, *wenn $\mathcal{X}$ eine abzählbare dichte Teilmenge enthält.*

Die Räume $\mathbb{R}^n$, ℓ^2, $\mathcal{L}^2$ und $\mathcal{C}([a,b])$ sind separabel, Beispiele für nichtseparable Räume werden im zweiten Teil dieses Abschnitts angegeben. Zunächst zeigen wir aber

Satz 1.4.10: *Es seien $\mathcal{X}$ ein separabler Prähilbertraum und OS ein Orthonormalsystem in $\mathcal{X}$. Dann ist OS höchstens abzählbar unendlich.*

Beweis: Es seien $x, y \in OS$, $x \neq y$. Dann ist $\|x-y\| = \sqrt{2}$. Es sei $\{y_n\}_{n\in\mathbb{N}}$ dicht in $\mathcal{X}$. Dann gilt

$$\forall\, s \in OS \quad \exists\, y_s \qquad \|s - y_s\| < \frac{\sqrt{2}}{2}.$$

Jedem $s \in OS$ wird also ein $y_s \in \{y_n\}_{n\in\mathbb{N}}$ zugeordnet. Es sei

$$T \;:\; OS \longrightarrow \{y_1, y_2, \dots\}$$

diese Abbildung. Dann ist T injektiv, denn aus

$$Tx_1 = y_i \quad \text{und} \quad Tx_2 = y_i$$

folgt

$$\|x_1 - x_2\| \leq \|x_1 - y_i\| + \|x_2 - y_i\| < \sqrt{2},$$

also $x_1 = x_2$. Mithin ist OS höchstens abzählbar. □

Wir fassen nun unsere bisherigen Ergebnisse zusammen.

Satz 1.4.11: *Es seien $\mathcal{X}$ ein Hilbertraum, OS ein Orthonormalsystem in $\mathcal{X}$, $M := \overline{\operatorname{span} OS}$ und*

$$\forall\, x \in \mathcal{X} \qquad m_x := \sum_{u \in OS} (x, u)\, u.$$

Dann gilt:

(1) $m_x \in M$.

(2) $x - m_x \in M^{\perp}$.

(3) m_x *ist das bestapproximierende Element zu* x, *also* $\|x - m_x\| = \operatorname{dist}(x, M)$.

Beweis: Nach dem Vorangegangenen gibt es zu $x \in \mathcal{X}$ eine Menge $\{u_i\}_{i \in \mathbb{N}}$ mit $(x, u_i) \neq 0$ und

$$m_x = \sum_{i=1}^{\infty} (x, u_i)\, u_i.$$

Wegen

$$\left\| \sum_{i=1}^{n} (x, u_i) u_i \right\|^2 = \sum_{i=1}^{n} |(x, u_i)|^2 \leq \|x\|^2$$

ist $m_x \in M$ deshalb wohldefiniert, und es folgt

$$\forall\, v \in OS \qquad (x - m_x, v) = (x, v) - \sum_{i=1}^{\infty} (x, u_i)(u_i, v) = 0.$$

Das beweist (1) und (2). Für alle $m \in M$ ist nun

$$\|x - m\|^2 = \|(x - m_x) + (m_x - m)\|^2 = \|x - m_x\|^2 + \|m_x - m\|^2.$$

Das beweist (3). □

Es bleibt natürlich die wichtige Frage, ob bzw. wann $m_x = x$ ist.

Definition 1.4.12: *Es sei $\mathcal{X}$ ein Hilbertraum. Dann heißt ein Orthonormalsystem OS in $\mathcal{X}$* vollständig, *wenn*

$$M^{\perp} := \overline{\operatorname{span} OS}^{\perp} = \{o\}$$

ist.

Es geht also um die Frage nach der Vollständigkeit eines Orthonormalsystems. Wir zeigen

Satz 1.4.13: *Es seien $\mathcal{X}$ ein Hilbertraum und OS ein Orthogonalsystem in $\mathcal{X}$. Dann läßt sich OS zu einem vollständigen Orthogonalsystem ergänzen. Insbesondere enthält jeder Hilbertraum ein vollständiges Orthogonalsystem.*

Der Beweis folgt aus dem Zornschen Lemma. Es sei G die Menge aller Orthogonalsysteme in $\mathcal{X}$, welche OS enthalten. Diese Menge G ist vermöge der Inklusion partiell geordnet. Es sei $T \subset G$ eine vollständig geordnete Teilmenge. Dann besitzt T in G eine obere Schranke. Es sei nämlich

$$T = \left\{ t_\lambda \supset OS,\ \ \lambda \in \Lambda \right\}.$$

Dann ist

$$g := \bigcup_{\lambda \in \Lambda} t_\lambda \in G$$

eine obere Schranke von T. Mithin enthält G ein maximales Element h, und es gilt

$$\overline{\operatorname{span} h}^{\perp} = \{o\},$$

denn sonst wäre h nicht maximal. □

Wir fassen zusammen.

Satz 1.4.14: *Es seien $\mathcal{X}$ ein Hilbertraum und OS ein Orthonormalsystem in $\mathcal{X}$. Dann sind äquivalent:*

(1) *OS ist vollständig.*

(2) $\overline{\operatorname{span} OS} = \mathcal{X}$.

(3) $\forall\, x \in \mathcal{X} \quad x = \sum_{u \in OS} (x,u)\, u$.

(4) $\forall\, x, y \in \mathcal{X} \quad (x,y) = \sum_{u \in OS} (x,u) \cdot \overline{(y,u)}$.

(5) $\forall\, x \in \mathcal{X} \quad \|x\|^2 = \sum_{u \in OS} |(x,u)|^2$.

Aussage (5) nennt man *Parsevalsche Gleichung*. Wir skizzieren den Beweis:

(1) $\Rightarrow$ (2) folgt unmittelbar aus der Definition.

(2) $\Rightarrow$ (3): Es ist $x - \sum_{u \in OS} (x,u)\, u \in M^{\perp} = \{o\}$.

(3) $\Rightarrow$ (4): Die Rechnung wurde bereits vorgeführt; es treten dabei nur Summanden über $i \in \mathbb{N}$ auf.

(4) $\Rightarrow$ (5) folgt für $y = x$.

(5) $\Rightarrow$ (1): Es sei $x \in M^{\perp}$. Dann folgt $x = o$. □

Separable Hilberträume: Die separablen Hilberträume bilden eine wichtige Teilklasse. Deshalb wollen wir noch etwas näher auf sie eingehen. Wir wissen bereits, daß in separablen Hilberträumen alle Orthonormalsysteme höchstens abzählbar unendlich sind. Wir zeigen jetzt

Satz 1.4.15: *Es sei $\mathcal{X}$ ein unendlichdimensionaler separabler Hilbertraum. Dann ist $\mathcal{X}$ zu ℓ^2 kongruent.*

Es sei daran erinnert, daß im Falle $\dim \mathcal{X} = n < \infty$ der Hilbertraum $\mathcal{X}$ zu $\ell^2(n)$ kongruent ist.

Beweis: Wir wissen bereits, daß $\mathcal{X}$ ein abzählbares vollständiges Orthonormalsystem $OS = \{u_i\}_{i \in \mathbb{N}}$ enthält. Es seien $x \in \mathcal{X}$, $\xi_i := (x, u_i)$ und

$$\begin{aligned} T \;:\; & \mathcal{X} \longrightarrow \ell^2, \\ & x \longmapsto \xi := (\xi_1, \xi_2, \ldots). \end{aligned}$$

Aus der Parsevalschen Gleichung folgt dann

$$\|x\|^2 = \sum_{i=1}^{\infty} |\xi_i|^2 = \|\xi\|^2.$$

T ist linear. Aus $Tx = o$ folgt $x = o$. T ist also bijektiv und isometrisch. □

Folgerung 1.4.16: *ℓ^2 ist separabel.*

Das folgt unmittelbar aus

$$\|x - x_n\| < \varepsilon \iff \|Tx - Tx_n\| < \varepsilon.$$

Die Kongruenz erhält also die Separabilität. Natürlich läßt sich die Separabilität von ℓ^2 auch leicht direkt beweisen. Denn jedes $x \in \ell^2$ läßt sich durch ein $\xi \in \ell^2$ approximieren, welches nur endlich viele von Null verschiedene Komponenten hat. Diese approximiere man durch rationale Zahlen. □

Analog folgt die Separabilität der ℓ^p-Räume, $1 \le p < \infty$, und auch die Separabilität von $\mathcal{L}^p(\Omega)$ für $1 \le p < \infty$. Man approximiere ein $f \in \mathcal{L}^p(\Omega)$ zunächst durch stetige Funktionen mit kompaktem Träger und diese durch Polynome (Weierstraßscher Approximationssatz).

Es folgen Beispiele nicht separabler Räume:

1. Eine überabzählbare Menge $\mathcal{X}$ mit der diskreten Metrik.

2. Der Folgenraum ℓ^∞. Es sei nämlich $\{x_n\}_{n\in\mathbb{N}}$ dicht in ℓ^∞. Dann bilden wir mit $x_k := (\xi_{k1}, \xi_{k2}, \ldots)$

$$\eta_k := \begin{cases} \xi_{kk} + 1 & \text{für } |\xi_{kk}| \leq 1 \\ 0 & \text{sonst.} \end{cases}$$

Dann ist $|\eta_k| \leq 2$ und damit $y := (\eta_1, \eta_2, \ldots) \in \ell^\infty$, aber

$$|\eta_k - \xi_{kk}| \geq 1,$$

also

$$\forall\, n \in \mathbb{N} \qquad \|y - x_n\| \geq 1.$$

3. Die Räume $\mathcal{L}^\infty(\Omega)$ und $\mathcal{BV}([a,b],\mathbb{K})$.

4. Es sei $\mathcal{Y}$ die Menge der meßbaren Abbildungen

$$f \;:\; \mathbb{R} \longrightarrow \mathbb{R}$$

mit

$$\forall\, T > 0 \qquad f\big|(-T,T) \in \mathcal{L}^2\big((-T,T),\mathbb{R}\big)$$

und

$$\big|f\big|^2 := \lim_{T\to\infty} \frac{1}{T} \int_{-T}^{T} |f(x)|^2\, dx < \infty.$$

Die Menge $\mathcal{Y}$ ist ein Vektorraum mit der Halbnorm $\big|\cdot\big|$. Es sei

$$\mathcal{N} := \Big\{ f \in \mathcal{Y} \;\Big|\; \big|f\big| = 0 \Big\}.$$

Zum Beispiel gehören die Treppenfunktionen zum Unterraum $\mathcal{N}$. Es sei nun

$$\mathcal{X} := \mathcal{Y}/\mathcal{N}$$

und $\varphi : \mathcal{Y} \to \mathcal{X}$ die kanonische Abbildung. Auf $\mathcal{X}$ definieren wir das Skalarprodukt

$$(F,G)_{\mathcal{X}} := \lim_{T\to\infty} \frac{1}{T} \int_{-T}^{T} f(x)g(x)\, dx,$$

mit $F := \varphi(f)$, $G := \varphi(g)$. Diese Definition ist unabhängig von der Wahl der Repräsentanten f, g. Damit wird $\mathcal{X}$ zu einem Prähilbertraum.

Für $\alpha > 0$ sei nun

$$u_\alpha(x) := \sin \alpha x \in \mathcal{Y}.$$

Wegen

$$\sin(\alpha x) \cdot \sin(\beta x) = \frac{1}{2} \Big\{ \cos(\alpha - \beta)x - \cos(\alpha + \beta)x \Big\}$$

ist dann

$$\int_{-T}^{T} u_\alpha(x) u_\beta(x)\, dx = \frac{1}{2} \left\{ \frac{\sin(\alpha-\beta)x}{\alpha-\beta} - \frac{\sin(\alpha+\beta)x}{\alpha+\beta} \right\} \Bigg|_{-T}^{T} = \begin{cases} \frac{\sin(\alpha-\beta)T}{\alpha-\beta} - \frac{\sin(\alpha+\beta)T}{\alpha+\beta} & \text{für } \alpha \neq \beta \\ T - \frac{\sin(\alpha+\beta)T}{\alpha+\beta} & \text{für } \alpha = \beta. \end{cases}$$

Mit $U_\alpha := \varphi(u_\alpha)$ und $U_\beta := \varphi(u_\beta)$ folgt also

$$(U_\alpha, U_\beta) = \delta_{\alpha\beta}.$$

Damit enthält $\mathcal{X}$ ein überabzählbares Orthonormalsystem. $\mathcal{X}$ ist also nicht separabel.

1.4.4 Projektoren

Die Zerlegung

$$\mathcal{X} = M \oplus M^{\perp}$$

eines Hilbertraumes wie im Projektionssatz gibt Anlaß, *Projektoren* zu definieren. Wir lassen zunächst auch allgemeinere Räume zu.

Definition 1.4.17: *Es sei $\mathcal{X}$ ein Vektorraum. Dann heißt $P \in \mathcal{L}(\mathcal{X}, \mathcal{X})$* Projektor *von $\mathcal{X}$ auf $\mathcal{R}(P)$, wenn $P \circ P = P$ ist.*

Wir zeigen den

Satz 1.4.18: *Es seien $\mathcal{X}$ ein Vektorraum und P ein Projektor. Dann gilt:*

(1) $\mathcal{X} = \mathcal{R}(P) \oplus \mathcal{N}(P)$ *(direkte Summe).*

(2) $id - P$ *ist ein Projektor auf $\mathcal{N}(P)$ mit*

$$\mathcal{N}(id - P) = \mathcal{R}(P).$$

Beweis:

Zu (1): Es ist $x - Px \in \mathcal{N}(P)$ für alle $x \in \mathcal{X}$. Es sei mit $r_i := Px_i$ für $i = 1, 2$

$$x = r_1 + n_1 = r_2 + n_2.$$

Dann ist $n_i \in \mathcal{N}(P)$ und

$$r_1 - r_2 = n_2 - n_1,$$

also

$$0 = P(r_1 - r_2) = P^2(x_1 - x_2) = P(x_1 - x_2) = r_1 - r_2.$$

Daraus folgt (1).

Zu (2): Wegen

$$(id - P)^2 = id - 2P + P^2 = id - P$$

ist $id - P$ ebenfalls ein Projektor. Es gilt also auch

$$\mathcal{X} = \mathcal{R}(id - P) \oplus \mathcal{N}(id - P).$$

Ist $y \in \mathcal{R}(id - P)$, dann gibt es ein x mit $y = (id - P)x$; es folgt $Py = o$, also $y \in \mathcal{N}(P)$. Ist $y \in \mathcal{N}(P)$, dann folgt $y = (id - P)y \in \mathcal{R}(id - P)$. Also gilt

$$\mathcal{R}(id - P) = \mathcal{N}(P)$$

und $\mathcal{N}(id - P) = \mathcal{R}(P)$. □

Satz 1.4.19: *Es sei $\mathcal{X}$ ein Vektorraum und $\mathcal{Y} \subset \mathcal{X}$ ein Unterraum. Dann gibt es einen Projektor von $\mathcal{X}$ auf $\mathcal{Y}$.*

Beweis: Zu $id \in \mathcal{L}(\mathcal{Y}, \mathcal{Y})$ gibt es eine Fortsetzung $F \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ aufgrund des Fortsetzungssatzes in §1.1.2. Dann ist F ein Projektor, denn es folgt aus $Fx = y \in \mathcal{Y}$ für $x \in \mathcal{X}$

$$F^2 x = Fy = y = Fx,$$

also $F^2 = F$. □

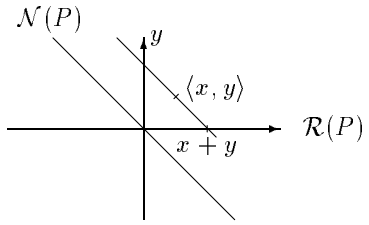
Dieses Ergebnis ist also eine triviale Folgerung aus dem Fortsetzungssatz. Was man eigentlich wissen möchte ist aber, ob es auch einen stetigen Projektor gibt. Diese Frage stellt sich in endlichdimensionalen Räumen natürlich noch nicht. Zum Beispiel sei $\mathcal{X} = \mathbb{R}^2$. Dann ist

$$P(\langle x, y \rangle) = \langle x + y, 0 \rangle$$

ein Projektor auf die x-Achse mit dem Nullraum

$$\mathcal{N}(P) = \{ \langle x, y \rangle \mid x + y = 0 \}$$

und $\|P\| = \sqrt{2}$. Man vergleiche die folgende Abbildung.



Wir zeigen zuerst

Satz 1.4.20: *Es seien $\mathcal{X}$ ein normierter Raum und P ein stetiger Projektor. Dann gilt:*

(1) Es ist $P = O$ oder $\|P\| \geq 1$, dabei ist O die Nullabbildung.

(2) $\mathcal{N}(P)$ und $\mathcal{R}(P)$ sind in $\mathcal{X}$ abgeschlossen.

Beweis: Die erste Behauptung folgt aus

$$\forall\, r \in \mathcal{R}(P) \qquad Pr = r.$$

Es seien $n_i \in \mathcal{N}(P)$ mit $n_i \to n$. Dann ist

$$Pn = \lim_{i \to \infty} Pn_i = o,$$

also $n \in \mathcal{N}(P)$. Mithin ist $\mathcal{N}(P)$ abgeschlossen. Damit ist auch $\mathcal{R}(P) = \mathcal{N}(id - P)$ abgeschlossen. □

Satz 1.4.21: *Es seien $\mathcal{X}$ ein Banachraum und M_1, M_2 abgeschlossene Teilräume mit $\mathcal{X} = M_1 \oplus M_2$. Dann ist der Projektor P von $\mathcal{X}$ auf M_1 stetig.*

Beweis: Es ist $\mathcal{N}(P) = M_2$. Wir zeigen nur die Abgeschlossenheit des Graphen von P. Die Behauptung folgt dann aus dem Satz vom abgeschlossenen Graphen, der in §5.1 bewiesen wird. Es seien also x und y definiert durch $x_n \to x$ und $Px_n \to y$. Dann gilt

$$x_n - Px_n \to x - y.$$

Aus $Px_n \in M_1$ und $x_n - Px_n \in M_2$ folgt daher $y \in M_1$ und $x - y \in M_2 = \mathcal{N}(P)$, also $Py = y$ und $Px = Py$. Deshalb ist $Px = y$. □

Der Nachweis der Existenz eines stetigen Projektors auf einen Unterraum $\mathcal{Y} \subset \mathcal{X}$ ist nicht trivial. Satz 1.4.20 besagt, daß dazu die Abgeschlossenheit von $\mathcal{Y}$ notwendig ist. Diese Bedingung reicht jedoch im allgemeinen nicht hin. Dafür gibt es Beispiele in ℓ^p- oder $\mathcal{L}^p$-Räumen. Man vergleiche die Zitate in *Taylor & Lay [1980], S. 247.*

Einfacher wird die Problematik in Hilberträumen.

Definition 1.4.22: *Es sei $\mathcal{X}$ ein Prähilbertraum. Dann heißt ein Projektor P* orthogonal, *wenn er symmetrisch ist.*

Für alle $x, y \in \mathcal{X}$ soll also

$$(Px, y) = (x, Py)$$

gelten.

Satz 1.4.23: *Es seien $\mathcal{X}$ ein Prähilbertraum und $P \in \mathcal{L}(\mathcal{X}, \mathcal{X})$. Dann sind äquivalent:*

(1) P ist ein orthogonaler Projektor auf $\mathcal{R}(P)$.

(2) P ist ein Projektor mit $\mathcal{R}(P)^\perp = \mathcal{N}(P)$.

(3) P ist ein stetiger Projektor mit $\|P\| \leq 1$.

(4) $\forall\, x \in \mathcal{X}$ ist $(id - P)x \in \mathcal{R}(P)^\perp$.

(5) P ist ein Projektor mit

$$\forall\, x \in \mathcal{X} \qquad \|x - Px\| = \operatorname{dist}\big(x, \mathcal{R}(P)\big).$$

Beweis:

(1) ⇒ (2): Aus $y \in \mathcal{R}(P)^\perp$ folgt

$$\forall\, x \in \mathcal{X} \qquad 0 = (Px, y) = (x, Py),$$

also $y \in \mathcal{N}(P)$.

(2) $\Rightarrow$ (3): Es ist

$$\mathcal{X} = \mathcal{R}(P) \oplus \mathcal{N}(P) = \mathcal{R}(P) \oplus \mathcal{R}(P)^{\perp}.$$

Mit $x = Px + (id - P)x$ folgt daraus

$$\|x\|^2 = \|Px\|^2 + \|(id - P)x\|^2.$$

(3) $\Rightarrow$ (4): Es seien $x \in \mathcal{X}$, $y \in \mathcal{R}(P)$, $\varepsilon > 0$, $\alpha \in \mathbb{C}$ mit $|\alpha| = 1$ und $z := x - Px \in \mathcal{N}(P)$. Dann ist

$$\|y\|^2 = \|P(\varepsilon z + \alpha y)\|^2 \leq \|\varepsilon z + \alpha y\|^2 = \varepsilon^2 \|z\|^2 + 2\varepsilon \operatorname{Re}(z, \alpha y) + \|y\|^2$$

oder

$$0 \leq \varepsilon \|z\|^2 + 2 \operatorname{Re}(z, \alpha y),$$

und es folgt für $\varepsilon \to 0$

$$\forall\, \alpha \in \mathbb{C},\ |\alpha| = 1, \qquad 0 \leq \operatorname{Re}(z, \alpha y).$$

Wegen $\operatorname{Im}(z, y) = \operatorname{Re}(z, \mathrm{i}y)$ ist dann $(z, y) = 0$.

(4) $\Rightarrow$ (1): Für alle $x, y \in \mathcal{X}$ ist

$$(x, Py) - (Px, y) = (x - Px, Py) + (Px, Py - y) = 0.$$

Mithin ist P symmetrisch. P ist ein Projektor wegen

$$\forall\, x, y \in \mathcal{X} \qquad 0 = (Py, x - Px) = (y, Px - P^2 x).$$

(4) $\Rightarrow$ (5): Es seien $x, y \in \mathcal{X}$ und $z := y - x$. Dann ist

$$\|x - Py\|^2 = \|x - P(x + z)\|^2 = \|x - Px\|^2 - 2 \operatorname{Re} \underbrace{(x - Px, Pz)}_{=0} + \|Pz\|^2.$$

Daraus folgt

$$\operatorname{dist}\big(x, \mathcal{R}(P)\big) = \|x - Px\|.$$

(5) $\Rightarrow$ (4): Aus dem Beweis des Approximationssatzes folgt

$$\mathcal{X} = \mathcal{R}(P) \oplus \mathcal{R}(P)^{\perp}.$$ □

Bemerkung 1.4.24:

1. id und O sind orthogonale Projektoren. Es sei M ein vollständiger Unterraum des Prähilbertraumes $\mathcal{X}$. Dann gibt es einen orthogonalen Projektor auf M.

2. Für stetige Projektoren gilt nach Satz 1.4.20

$$P = O \quad \textit{oder} \quad \|P\| \geq 1.$$

Für orthogonale Projektoren folgt daraus nach (3)

$$P = O \quad \textit{oder} \quad \|P\| = 1.$$

1.5 Dualräume und beschränkt lineare Funktionale

In Definition 1.1.5 haben wir bereits den algebraisch konjugierten Raum $\mathcal{X}^f$ zu einem Vektorraum $\mathcal{X}$ eingeführt. Es war

$$\mathcal{X}^f := \mathcal{L}(\mathcal{X}, \mathbb{K})$$

die Menge der linearen Funktionale auf $\mathcal{X}$. Inzwischen stehen uns auch topologische Strukturen zur Verfügung; es sei also $\mathcal{X}$ ein normierter Raum. Dann bezeichnet man entsprechend

Definition 1.5.1: *Es sei $\mathcal{X}$ ein normierter Vektorraum. Dann heißt der Raum der stetigen linearen Funktionale auf $\mathcal{X}$,*

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

der topologisch konjugierte Raum *von $\mathcal{X}$. Man nennt ihn auch* Dualraum *oder* adjungierten Raum *von $\mathcal{X}$.*

$\mathcal{X}'$ ist ein Banachraum (Folgerung 1.3.11.3). Mit $x \in \mathcal{X}$ und $x' \in \mathcal{X}'$ schreiben wir wieder

$$x'(x) =: x'x.$$

Nach Folgerung 1.3.18.2 ist $\mathcal{X}' = \mathcal{X}^f$, wenn $\dim \mathcal{X} < \infty$ ist. Im allgemeinen ist $\mathcal{X}'$ jedoch kleiner als $\mathcal{X}^f$. Das ist für viele Anwendungen vorteilhaft, wirft aber zunächst die wichtige Frage auf, ob der Fortsetzungssatz aus §1.1.2 auch in dieser Situation gilt. Diese Frage wurde von Hans Hahn, 1879–1934, und Stefan Banach geklärt. Nach ihnen ist der entsprechende Fortsetzungssatz benannt.

1.5.1 Der Satz von Hahn-Banach

Satz von Hahn-Banach: *Es seien $\mathcal{X}$ ein normierter Raum, $\mathcal{M} \subset \mathcal{X}$ ein Unterraum und $m' \in \mathcal{M}'$. Dann gilt*

$$\exists\, x' \in \mathcal{X}' \quad \textit{mit } x' \,|\, \mathcal{M} = m' \textit{ und } \|x'\| = \|m'\|.$$

Beweis:

1. Es sei $\mathbb{K} = \mathbb{R}$. Dann ist

$$\begin{aligned} p \;:\; \mathcal{X} &\longrightarrow \mathbb{R}, \\ x &\longmapsto \|m'\| \cdot \|x\| \end{aligned}$$

ein sublineares Funktional, man vergleiche Definition 1.1.16. Außerdem gilt

$$\forall\, \alpha \in \mathbb{R} \quad p(\alpha x) = |\alpha| p(x) \qquad \text{und} \qquad \forall\, m \in \mathcal{M} \quad m'm \le |m'm| \le p(m).$$

Nach Satz 1.1.17 besitzt m' deshalb eine algebraische Fortsetzung $F \in \mathcal{L}(\mathcal{X}, \mathbb{R})$ mit $F|\mathcal{M} = m'$ und

$$\forall\, x \in \mathcal{X} \qquad Fx \le p(x) = \|m'\| \cdot \|x\|.$$

Nun ist

$$-Fx = F(-x) \le p(-x) = p(x),$$

also gilt

$$\forall\, x \in \mathcal{X} \qquad |Fx| \le p(x) = \|m'\| \cdot \|x\|.$$

Mithin ist $F \in \mathcal{CL}(\mathcal{X}, \mathbb{R})$ und $\|F\| \le \|m'\|$. Weil F eine Fortsetzung von m' ist, gilt sogar $\|F\| = \|m'\|$. Man wähle also $x' := F$.

2. Es sei $\mathbb{K} = \mathbb{C}$. Diesen Fall führt man mit etwas Rechnung auf den reellen zurück. Es sei $\mathcal{Y} \subseteq \mathcal{X}$ ein Unterraum von $\mathcal{X}$ und $f \in \mathcal{L}(\mathcal{Y}, \mathbb{K})$. Dann zerlegen wir f

$$f(y) = f_1(y) + \mathrm{i} f_2(y) \quad \text{mit } f_j \in \mathcal{L}(\mathcal{Y}, \mathbb{R}),\ j = 1, 2.$$

Aus $f(\mathrm{i}\, y) = \mathrm{i}\, f(y)$ folgt

$$f_1(\mathrm{i}\, y) + \mathrm{i}\, f_2(\mathrm{i}\, y) = \mathrm{i}\, f_1(y) - f_2(y)$$

also

$$f_2(y) = -f_1(\mathrm{i}\, y).$$

Das wird im folgenden ausgenutzt.

Wir zerlegen $m' \in \mathcal{CL}(\mathcal{M}, \mathbb{K})$ entsprechend in

$$m'x = m'_1 x + \mathrm{i}\, m'_2 x \quad \text{mit } m'_j \in \mathcal{CL}(\mathcal{M}, \mathbb{R})$$

und fassen nun für den Augenblick $\mathcal{M}$ als Vektorraum über $\mathbb{R}$ auf. Dann existiert im Sinne des ersten Beweisschrittes eine Fortsetzung $x'_1 \in \mathcal{CL}(\mathcal{X}, \mathbb{R})$ von m'_1. Es sei

$$x'x = x'_1 x + \mathrm{i}\, x'_2 x := x'_1 x - \mathrm{i}\, x'_1(\mathrm{i}\, x).$$

Für $m \in \mathcal{M}$ gilt dann $x'm = m'm$, und wegen

$$x'(\mathrm{i}\, x) = x'_1(\mathrm{i}\, x) + \mathrm{i}\, x'_1 x = -x'_2 x + \mathrm{i}\, x'_1 x = \mathrm{i}(x'_1 x + \mathrm{i}\, x'_2 x) = \mathrm{i}\, x'x$$

ist sogar $x' \in \mathcal{CL}(\mathcal{X}, \mathbb{K})$. Es bleibt der Nachweis von $\|x'\| = \|m'\|$: Es sei $x'x = r\, \mathrm{e}^{\mathrm{i}\varphi}$. Dann ist

$$r = \mathrm{e}^{-\mathrm{i}\varphi} x'x = x'(\mathrm{e}^{-\mathrm{i}\varphi} x) = x'_1(\mathrm{e}^{-\mathrm{i}\varphi} x) - \mathrm{i}\, x'_1(\mathrm{i}\, \mathrm{e}^{-\mathrm{i}\varphi} x),$$

und es folgt

$$|x'x| = r = x_1'(\mathrm{e}^{-\mathrm{i}\varphi}x) \le p(\mathrm{e}^{-\mathrm{i}\varphi}x) = p(x) = \|m'\| \cdot \|x\|.$$

Damit ist der Satz von Hahn-Banach bewiesen. □

Ist $\mathcal{X}$ ein Hilbertraum, dann vereinfacht sich die Situation. Nach Satz 1.3.12 läßt sich $m' \in \mathcal{CL}(\mathcal{M}, \mathbb{K})$ stetig zu $\bar{m}' \in \mathcal{CL}(\bar{\mathcal{M}}, \mathbb{K})$ fortsetzen. Es sei $P : \mathcal{X} \to \bar{\mathcal{M}}$ der orthogonale Projektor auf $\bar{\mathcal{M}}$. Durch

$$x'x := \bar{m}'(Px)$$

erhält man daher eine stetige Fortsetzung von m'.

Folgerung 1.5.2: *Es seien $\mathcal{X}$ ein normierter Raum und $S_{\mathcal{X}'} := \{x' \in \mathcal{X}' \mid \|x'\| = 1\}$ die Sphäre in $\mathcal{X}'$. Dann gilt:*

(1) $\forall\, x \in \mathcal{X} \quad \exists\, s' \in S_{\mathcal{X}'} \quad s'x = \|x\|.$

(2) $\forall\, x \in \mathcal{X} \quad \sup_{s' \in S_{\mathcal{X}'}} |s'x| = \|x\|.$

(3) Es sei $x \in \mathcal{X}$ fest. Dann gilt

$$\Big(\forall\, x' \in \mathcal{X}' \quad x'x = 0\Big) \implies x = o.$$

Beweis: Für festes $x \in \mathcal{X}$ seien $\mathcal{M} := \mathrm{span}(x)$ und $m'x := \|x\|$. Dann werde $m' \in \mathcal{M}'$ zu $s' \in S_{\mathcal{X}'}$ fortgesetzt. Das beweist die erste Aussage, und die zweite ergibt sich damit aus

$$|s'(x)| \le \|s'\| \cdot \|x\| = \|x\|.$$

Die dritte Aussage folgt aus der zweiten. □

Es gilt ebenfalls der

Trennungssatz: *Es seien $\mathcal{X}$ ein normierter Raum, $\mathcal{X}_0$ ein Teilraum und $x_1 \in \mathcal{X} \backslash \mathcal{X}_0$ mit $\mathrm{dist}(x_1, \mathcal{X}_0) > 0$. Dann gilt*

$$\exists\, s_1' \in S_{\mathcal{X}'} \quad \textit{mit } s_1'x = \begin{cases} 0 & \textit{für } x \in \mathcal{X}_0 \\ \mathrm{dist}(x, \mathcal{X}_0) & \textit{für } x = x_1. \end{cases}$$

Man sagt s_1' und $d := \mathrm{dist}(x_1, \mathcal{X}_0) > 0$ *trennen* x_1 von $\mathcal{X}_0$.

Zum Beweis sei

$$\mathcal{M} := \mathrm{span}\,(\mathcal{X}_0, x_1)\,.$$

Wie im Beweis des Trennungssatzes in §1.1.2 läßt sich dann

$$m = x_0 + \alpha\, x_1 \in \mathcal{M} \quad \text{mit } x_0 \in \mathcal{X}_0$$

eindeutig darstellen. Durch $m'm := \alpha d$ wird ein $m' \in \mathcal{CL}(\mathcal{M}, \mathbb{K})$ definiert. Für $\alpha \neq 0$ ist

$$\|m\| = |\alpha| \left\| \frac{x_0}{\alpha} + x_1 \right\| \ge |\alpha|\, d,$$

also $|m'm| = |\alpha| d \le \|m\|$, das heißt $\|m'\| \le 1$. Es ist sogar $\|m'\| = 1$: Um das zu zeigen, gehen wir von

$$\forall\, \varepsilon > 0 \quad \exists\, x \in \mathcal{X}_0 \qquad \|x - x_1\| < d + \varepsilon$$

aus und wählen

$$y := \frac{x - x_1}{\|x - x_1\|} \in \mathcal{M}.$$

Dann ist $\|y\| = 1$ und

$$|m'y| = \frac{d}{\|x - x_1\|} > \frac{d}{d + \varepsilon},$$

also $\|m'\| > d/(d+\varepsilon) \to 1$ für $\varepsilon \to 0$. Wir wählen nun s_1' als Fortsetzung von m'. □

Schließlich folgern wir noch

Satz 1.5.3: *Es seien $\mathcal{X}$ ein normierter Raum und $\mathcal{X}'$ separabel. Dann ist $\mathcal{X}$ separabel.*

Leider ist die Umkehrung falsch; ℓ^1 ist separabel, $\ell^\infty = (\ell^1)'$ aber nicht.

Zum Beweis des Satzes wählen wir eine dichte Menge $\{s'_n\}_{n\in\mathbb{N}}$ in $S_{\mathcal{X}'}$. Wir wählen ferner $s_n \in S_{\mathcal{X}}$ mit

$$|s'_n s_n| \geq \frac{1}{2}.$$

Das ist möglich wegen

$$1 = \|s'_n\| = \sup_{s\in S_{\mathcal{X}}} |s'_n s|.$$

Es sei nun $\mathcal{M} := \overline{\text{span}\,(s_1, s_2, \ldots)}$. Unser Ziel ist es, $\mathcal{X} = \mathcal{M}$ zu zeigen. Wir gehen indirekt vor. Es sei $\mathcal{M} \neq \mathcal{X}$ und $x_0 \in \mathcal{X}\backslash\mathcal{M}$. Aufgrund des Trennungssatzes gibt es dann ein $s' \in S_{\mathcal{X}'}$ mit

$$s'x_0 \neq 0 \quad \text{und} \quad \forall\, m \in \mathcal{M} \quad s'm = 0.$$

Es folgt

$$\frac{1}{2} \leq |s'_n s_n| = |s'_n s_n - s's_n| \leq \|s'_n - s'\| \to 0$$

für eine geeignete Teilfolge von (s'_n). Das ist ein Widerspruch. □

1.5.2 Konjugierte Abbildungen

Es seien nun $\mathcal{X}, \mathcal{Y}$ normierte Räume und $A \in \mathcal{CL}(\mathcal{X}, \mathcal{Y})$. Dann definieren wir analog zur transponierten Abbildung A^{tr}

$$\begin{aligned} A' \;:\; & \mathcal{Y}' \longrightarrow \mathcal{X}', \\ & y' \longmapsto x' := y' \circ A, \end{aligned}$$

also

$$\forall\, x \in \mathcal{X} \qquad x'x = y'(Ax).$$

A' ist wohldefiniert und heißt die zu A *konjugierte Abbildung*. Es gilt also

$$\forall\, y' \in \mathcal{Y}' \quad \forall\, x \in \mathcal{X} \qquad (A'y')x = y'(Ax).$$

Satz 1.5.4: *Es seien $\mathcal{X}, \mathcal{Y}$ normierte Räume und $A \in \mathcal{CL}(\mathcal{X}, \mathcal{Y})$. Dann ist $A' \in \mathcal{CL}(\mathcal{Y}', \mathcal{X}')$, und es gilt $\|A'\| = \|A\|$.*

Beweis: Es ist

$$\|A'\| = \sup_{y'\in S_{\mathcal{Y}'}} \|A'y'\| = \sup_{y'\in S_{\mathcal{Y}'}} \sup_{x\in S_{\mathcal{X}}} |(A'y')x| = \sup_{y'\in S_{\mathcal{Y}'}} \sup_{x\in S_{\mathcal{X}}} |y'(Ax)| =$$
$$= \sup_{x\in S_{\mathcal{X}}} \sup_{y'\in S_{\mathcal{Y}'}} |y'(Ax)| = \sup_{x\in S_{\mathcal{X}}} \|Ax\| = \|A\|.$$

Die vorletzte Gleichheit erhält man aus Folgerung 1.5.2. □

Wir betrachten nun analog zu $\mathcal{X}^{\prime\prime}$

$$\mathcal{X}'' := (\mathcal{X}')'.$$

Es sei

$$\begin{aligned} J_{\mathcal{X}} \;:\; & \mathcal{X} \longrightarrow \mathcal{X}'', \\ & x \longmapsto x'' \end{aligned}$$

vermöge

$$\forall\, x \in \mathcal{X} \quad \overset{1}{\exists}\, x'' \in \mathcal{X}'' \quad \forall\, x' \in \mathcal{X}' \qquad x''x' = x'x$$

die kanonische Abbildung von $\mathcal{X}$ in $\mathcal{X}''$. Sie ist wohldefiniert, und aus Folgerung 1.5.2 erhalten wir wieder

$$\|J_{\mathcal{X}} x\| = \sup_{x'\in S_{\mathcal{X}'}} |(J_{\mathcal{X}} x)x'| = \sup_{x'\in S_{\mathcal{X}'}} |x'x| = \|x\|,$$

also $\|J\| = 1$.

Definition 1.5.5: *Der normierte Raum $\mathcal{X}$ heißt* (topologisch) reflexiv, *wenn $\mathcal{R}(J_{\mathcal{X}}) = \mathcal{X}''$ ist.*

Die kanonische Abbildung J ist dann eine Kongruenz, und wir können in diesem Sinne $\mathcal{X}$ mit $\mathcal{X}''$ identifizieren.

Lemma 1.5.6: *Es seien $\mathcal{X}, \mathcal{Y}$ normierte Räume und $A \in \mathcal{CL}(\mathcal{X}, \mathcal{Y})$. Dann kommutiert das Diagramm*

$$\begin{array}{ccc} \mathcal{X} & \xrightarrow{A} & \mathcal{Y} \\ {\scriptstyle J_{\mathcal{X}}}\downarrow & & \downarrow{\scriptstyle J_{\mathcal{Y}}} \\ \mathcal{X}'' & \xrightarrow{A''} & \mathcal{Y}''. \end{array}$$

Es gilt also

$$A'' \circ J_{\mathcal{X}} \;=\; J_{\mathcal{Y}} \circ A.$$

Für alle x, y' ist nämlich

$$(J_{\mathcal{Y}} Ax)y' = y'(Ax) = (A'y')x = (J_{\mathcal{X}} x)A'y' = (A'' J_{\mathcal{X}} x)y'. \qquad \square$$

Lemma 1.5.7: *Es sei $\mathcal{X}$ ein normierter Raum. Dann gilt in $\mathcal{X}'$*

$$J'_{\mathcal{X}} \circ J_{\mathcal{X}'} = id.$$

Beweis: Es ist $J'_{\mathcal{X}} : \mathcal{X}''' \to \mathcal{X}'$ und $J_{\mathcal{X}'} : \mathcal{X}' \to \mathcal{X}'''$. Für alle x, x' ist daher

$$\left(J'_{\mathcal{X}} J_{\mathcal{X}'} x'\right) x = \left(J_{\mathcal{X}'} x'\right) \left(J_{\mathcal{X}} x\right) = \left(J_{\mathcal{X}} x\right) x' = x'x. \qquad \square$$

Satz 1.5.8: *$\mathcal{X}$ sei ein reflexiver normierter Raum. Dann ist auch $\mathcal{X}'$ reflexiv.*

Beweis: Es gilt

$$J'_{\mathcal{X}} x''' = o \iff \forall\, x \in \mathcal{X}\ \ (J'_{\mathcal{X}} x''')x = 0 \iff \forall\, x \in \mathcal{X}\ \ x'''(J_{\mathcal{X}} x) = 0 \iff x''' = o.$$

Mithin ist $J'_{\mathcal{X}} : \mathcal{X}''' \to \mathcal{X}'$ injektiv, und das letzte Lemma zeigt die Existenz von

$$J_{\mathcal{X}'} = (J'_{\mathcal{X}})^{-1},$$

also

$$\mathcal{R}(J_{\mathcal{X}'}) = \mathcal{R}\left((J'_{\mathcal{X}})^{-1}\right) = \mathcal{X}'''. \qquad \square$$

Für Banachräume gilt auch die Umkehrung

Satz 1.5.9: *Es seien $\mathcal{X}$ ein Banachraum und $\mathcal{X}'$ reflexiv. Dann ist auch $\mathcal{X}$ reflexiv.*

Beweis: $J_{\mathcal{X}'}$ ist surjektiv. Aus Lemma 1.5.7 folgt daraus die Injektivität von $J'_{\mathcal{X}}$. Nach Voraussetzung ist $\mathcal{X}$ vollständig und $J_{\mathcal{X}}$ isometrisch. Mithin ist $\mathcal{R}(J_{\mathcal{X}})$ ein abgeschlossener Teilraum von $\mathcal{X}''$. Es sei $\mathcal{R}(J_{\mathcal{X}}) \neq \mathcal{X}''$. Dann gibt es aufgrund des Trennungssatzes ein $x''' \neq o$ mit

$$\forall\, x \in \mathcal{X} \qquad 0 = x'''\left(J_{\mathcal{X}} x\right) = \left(J'_{\mathcal{X}} x'''\right) x.$$

Wegen der Injektivität von $J'_{\mathcal{X}}$ folgt daraus aber der Widerspruch $x''' = o$. $\square$

Satz 1.5.10: *Reflexive Räume sind Banachräume.*

Beweis: $\mathcal{X}'' = \mathcal{CL}(\mathcal{X}', \mathbb{K})$ ist nach Folgerung 1.3.11.3 ein Banachraum. Nun ist $\mathcal{X}$ zu $\mathcal{X}''$ kongruent. $\square$

Es gibt viele Beispiele für reflexive Räume. Die Räume ℓ^p und $\mathcal{L}^p(\Omega)$ mit $1 < p < \infty$ gehören dazu.

1.5.3 Der Rieszsche Darstellungssatz

Zum Abschluß dieses Kapitels soll noch eine wichtige Darstellung stetiger linearer Funktionale in Hilberträumen angegeben werden. Sie geht auf FRIGYES RIESZ, 1880–1956, zurück.

Rieszscher Darstellungssatz: *Es seien $\mathcal{X}$ ein Hilbertraum und $F \in \mathcal{X}'$ ein stetiges lineares Funktional. Dann gilt*

$$\overset{1}{\exists}\, f \in \mathcal{X} \quad \forall\, x \in \mathcal{X} \qquad Fx = (x, f)$$

und $\|f\|_{\mathcal{X}} = \|F\|_{\mathcal{X}'}$.

Beweis: Es sei $\mathcal{N} := \{x \in \mathcal{X} \mid Fx = 0\}$. $\mathcal{N}$ ist ein abgeschlossener Teilraum von $\mathcal{X}$. Aus $x_n \in \mathcal{N}$, $x_n \to x$ folgt nämlich

$$|Fx| = |Fx_n + F(x - x_n)| \leq \|F\| \cdot |x - x_n| \to 0.$$

Nun sind zwei Fälle möglich:

1. Es sei $\mathcal{N} = \mathcal{X}$. Dann wähle man $f := o$, und dann ist $\|f\| = \|F\|$.

2. Es sei $\mathcal{N} \neq \mathcal{X}$. Dann ist

$$\mathcal{X} = \mathcal{N} \oplus \mathcal{N}^{\perp},$$

und es gibt ein $w \in \mathcal{N}^{\perp}$ mit $w \neq 0$ und $F\,w \neq 0$. Es folgt

$$x - w\frac{Fx}{Fw} \in \mathcal{N}$$

oder

$$(x, w) = \frac{\|w\|^2}{Fw} Fx.$$

Man wähle deshalb

$$f := \frac{\overline{Fw}}{\|w\|^2} w.$$

Dann ist $Fx = (x, f)$. Es sei auch $(x, f') = Fx$. Dann folgt

$$\forall\, x \in \mathcal{X} \quad (x, f - f') = 0,$$

also $\|f - f'\|^2 = 0$, das heißt $f = f'$. Es bleibt $\|f\| = \|F\|$ zu zeigen. Das folgt aus

$$\|f\| = \sup_{\|w\|=1} \frac{|Fw|}{\|w\|} \leq \|F\|$$

und

$$\|F\| = \sup_{\|x\|=1} |Fx| = \sup_{\|x\|=1} |(x, f)| \leq \|f\|.$$

Damit ist der Rieszsche Darstellungssatz bewiesen. □

Der Rieszsche Darstellungssatz liefert also eine konjugiert lineare Isometrie zwischen $\mathcal{X}'$ und $\mathcal{X}$. Es sei

$$\begin{aligned} R \;:\; \mathcal{X} &\longrightarrow \mathcal{X}', \\ x &\longmapsto Rx \end{aligned}$$

mit

$$\forall\, z \in \mathcal{X} \qquad (z, x) =: (Rx)z$$

die *Rieszabbildung*. R ist eine konjugiert lineare Abbildung mit

$$\|Rx\| = \sup_{z \in S_{\mathcal{X}}} |(z, x)| = \|x\|.$$

Der Rieszsche Darstellungssatz lautet dann

Satz 1.5.11: *Die Rieszabbildung $R : \mathcal{X} \to \mathcal{X}'$ ist surjektiv und isometrisch.*

Damit sind $\mathcal{X}, \mathcal{X}'$ und $\mathcal{X}''$ (konjugiert) kongruent. Insbesondere gilt

Satz 1.5.12: *Hilberträume sind reflexiv.*

Beweis: Es sei $x'' \in \mathcal{X}''$ vorgegeben. Dann wählen wir $x' \in \mathcal{X}'$ vermöge

$$\forall\, z \in \mathcal{X} \qquad x'z := \overline{x''(Rz)}.$$

Weil R surjektiv ist, gibt es ein $x \in \mathcal{X}$ mit

$$Rx = x',$$

und es folgt für alle $z \in \mathcal{X}$

$$(J_{\mathcal{X}}\, x)(Rz) = (Rz)x = (x, z) = \overline{(z, x)} = \overline{(Rx)z} = \overline{x'z} = x''(Rz).$$

Wegen der Surjektivität von R ist also $J_{\mathcal{X}}\, x = x''$. □

Es seien $p, q \in (1, \infty)$ mit

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Es sei ferner $\mathcal{X} = \ell^p$ und mit $x = (\xi_1, \xi_2, \ldots) \in \ell^p$ und $y = (\eta_1, \eta_2, \ldots) \in \ell^q$ die Abbildung

$$R \;:\; \ell^p \longrightarrow (\ell^q)'$$

vermöge

$$(Rx)y := \sum_{n=1}^{\infty} \eta_n \bar{\xi}_n$$

definiert. R ist wieder eine Kongruenz ist. Der Raum ℓ^p ist reflexiv.

Die Räume ℓ^1 und ℓ^∞ sind aber nicht reflexiv. Wäre nämlich ℓ^1 reflexiv, dann wäre ℓ^1 zu $(\ell^1)''$ kongruent und wegen $(\ell^1)' \cong \ell^\infty$ wäre ℓ^1 zu $(\ell^\infty)'$ kongruent. Der Raum ℓ^1 ist separabel. Aus Satz 1.5.3 würde daher die Separabilität von ℓ^∞ folgen. Das ist aber falsch.

2 Fixpunktsätze

Für Theorie und Anwendungen ist es außerordentlich wichtig, Gleichungen lösen zu können, etwa

$$A(x) \stackrel{!}{=} f.$$

Eine solche Aufgabe läßt sich sehr einfach als *Fixpunktgleichung* schreiben, zum Beispiel als

$$B(x) \stackrel{!}{=} x$$

mit $B(x) := A(x) + x - f$. In der Praxis wird man B wohl geschickter wählen. Dabei seien f aus einer vorgegebenen Menge $\mathcal{X}$ gewählt und $B : \mathcal{X} \to \mathcal{X}$ eine Abbildung. Wir wollen uns im folgenden solche Fixpunktgleichungen etwas näher ansehen.

2.1 Der Banachsche Fixpunktsatz

Es seien (X, d) ein vollständiger metrischer Raum und $T : X \to X$ eine *kontrahierende Abbildung*. Es sei also

$$\exists\, k \in [0,1) \quad \forall\, x, y \in X \qquad d\big(T(x), T(y)\big) \le k\, d(x,y).$$

Dann gilt:

Der Banachsche Fixpunktsatz: *Es seien (X, d) ein vollständiger metrischer Raum und*

$$T \;:\; X \longrightarrow X$$

eine kontrahierende Abbildung. Dann besitzt die Fixpunktgleichung

$$Tx \;=\; x$$

genau eine Lösung $\hat{x}$. Mit beliebigem $x_0 \in X$ gilt

$$x_n := T\, x_{n-1} \to \hat{x}$$

und

$$d(x_n, \hat{x}) \le \frac{k^n}{1-k} d(x_1, x_0).$$

Der Beweis ist sehr einfach und wird in den meisten Vorlesungen zur Infinitesimalrechnung gebracht (vgl. *Barner & Flohr [1982], S. 146*). Der Satz liefert ein konstruktives Approximationsverfahren mit Fehlerabschätzung, er ist in seiner Art optimal. In einer Vorlesung über Infinitesimalrechnung wird er vielfach angewandt, beispielsweise zum Beweis des Satzes von Picard-Lindelöf, beim Newtonverfahren oder beim Satz von der lokalen Umkehrfunktion.

2.2 Der Brouwersche Fixpunktsatz

Weil sich nicht alle Fixpunktaufgaben mit einer kontrahierenden Abbildung formulieren lassen und weil man auch die Eindeutigkeit gar nicht immer erwartet, wollen wir uns im folgenden nach weiteren Fixpunktsätzen umsehen. Ein solcher ist der Brouwersche, der ebenfalls meist im Rahmen der Vorlesungen über Infinitesimalrechnung bewiesen wird (vgl. *Barner & Flohr [1982], S. 418*). Ist $\bar{B} \subset \mathbb{R}^n$ eine abgeschlossene Kugel in $\mathbb{R}^n$. Dann lautet er

Der Brouwersche Fixpunktsatz: *Jede stetige Abbildung $f : \bar{B} \to \bar{B}$ besitzt einen Fixpunkt. Es sei K ein zu $\bar{B}$ homöomorphes Gebiet. Dann gilt diese Aussage auch für stetige Abbildungen $f : K \to K$.*

In $\mathbb{R}^1$ ist dieser Satz eine unmittelbare Folge des Zwischenwertsatzes. Wir wollen ihn nun weiter verallgemeinern. Zur Abkürzung sagen wir, daß ein topologischer Raum $\mathcal{X}$ die *Fixpunkteigenschaft* hat, wenn jede stetige Abbildung

$$f \;:\; \mathcal{X} \longrightarrow \mathcal{X}$$

einen Fixpunkt besitzt. Eine abgeschlossene Kugel in $\mathbb{R}^n$ besitzt also die Fixpunkteigenschaft.

Allgemeiner Brouwerscher Fixpunktsatz: *Eine kompakte konvexe Teilmenge Q des $\mathbb{R}^n$ besitzt die Fixpunkteigenschaft.*

Zum Beweis holen wir etwas aus. Es sei daran erinnert, daß eine Teilmenge $Q \subset \mathcal{X}$ *konvex* heißt, wenn

$$\forall\, x, y \in Q \quad \forall\, \lambda \in [0,1] \qquad \lambda x + (1-\lambda) y \in Q$$

ist. Eine Teilmenge $A \subset \mathcal{X}$ heißt *Retrakt* , wenn es eine stetige *Retraktion* von $\mathcal{X}$ auf A gibt, wenn es also eine stetige Abbildung

$$\rho \;:\; \mathcal{X} \longrightarrow A$$

mit $\rho \,|\, A = id$ gibt. Es gilt

Lemma 2.2.1: *$\mathcal{X}$ besitze die Fixpunkteigenschaft, und A sei ein Retrakt von $\mathcal{X}$. Dann besitzt auch A die Fixpunkteigenschaft.*

Beweis: Es seien $f : A \to A$ eine stetige Abbildung, $\rho : \mathcal{X} \to A$ eine Retraktion und $i : A \to \mathcal{X}$ die Inklusion. Es sei ferner

$$\begin{aligned} g \;:\; & \mathcal{X} \longrightarrow \mathcal{X}, \\ & x \longmapsto \big(i \circ f \circ \rho\big)(x). \end{aligned}$$

Dann ist g stetig und besitzt einen Fixpunkt $\hat{x} \in \mathcal{X}$. Wegen $\mathcal{R}(g) \subset A$ ist dann $\hat{x} \in A$ und $\rho(\hat{x}) = \hat{x}$, also

$$\hat{x} \;=\; f(\hat{x}). \qquad \square$$

Wir werden im folgenden zeigen, daß eine kompakte konvexe Teilmenge Q des $\mathbb{R}^n$ ein Retrakt einer abgeschlossenen Kugel ist. Damit ist dann der allgemeine Brouwersche Fixpunktsatz bewiesen.

Weil Q kompakt ist, gibt es zu $x \in \mathbb{R}^n$ ein $q(x) \in Q$ mit

$$|x - q(x)| = \operatorname{dist}(x, Q).$$

Dieses $q(x)$ ist eindeutig bestimmt, weil Q konvex ist. Für $x \notin Q$ seien nämlich $q_1(x)$ und $q_2(x)$ zwei verschiedene Punkte kürzesten Abstands. Dann gehört die Verbindungsstrecke $[q_1(x), q_2(x)]$ zu Q und insbesondere auch

$$q_3(x) := \frac{1}{2}\Big(q_1(x) + q_2(x)\Big).$$

Im gleichschenkligen Dreieck $\big(x, q_1(x), q_2(x)\big)$ ist die Höhe $[x, q_3(x)]$ kürzer als die Seiten $[x, q_1(x)]$, $[x, q_2(x)]$, und das ist ein Widerspruch.

Die Abbildung

$$\begin{aligned} \rho \;:\; & \mathbb{R}^n \longrightarrow Q, \\ & x \longmapsto q(x) \end{aligned}$$

ist also wohldefiniert. Wegen $\rho \,|\, Q = id$ ist ρ eine Retraktion. Wir wollen ihre Stetigkeit zeigen. Das geschieht in zwei Schritten:

1. Es gilt

$$\forall\, x \in \mathbb{R}^n \quad \forall\, q \in Q \qquad \big(q - q(x)\big) \cdot \big(x - q(x)\big) \le 0. \tag{$*$}$$

Für $0 \le t \le 1$ sei nämlich

$$\varphi(t) := \Big| x - \underbrace{\big[q(x) + t\big(q - q(x)\big)\big]}_{\in Q} \Big|^2 \ge 0.$$

$\varphi(t)$ wird nach Definition von $q(x)$ in $t = 0$ minimal. Weil φ in $t = 0$ von rechts differenzierbar ist, folgt $\varphi'(0) \ge 0$. Nun ist

$$\varphi'(t) = -2\big(x - [q(x) + t(q - q(x))]\big) \cdot \big(q - q(x)\big),$$

also

$$\varphi'(0) = -2\big(x - q(x)\big) \cdot \big(q - q(x)\big) \ge 0.$$

2. Es gilt

$$\forall\, x, y \in \mathbb{R}^n \qquad |q(x) - q(y)| \le |x - y|. \tag{$**$}$$

Daraus folgt dann die Stetigkeit von q. Aus $(*)$ erhalten wir nämlich für $q = q(y)$

$$\big(q(y) - q(x)\big) \cdot \big(x - q(x)\big) \leq 0$$

und analog

$$\big(q(x) - q(y)\big) \cdot \big(y - q(y)\big) \leq 0,$$

zusammen also

$$\big(q(y) - q(x)\big) \cdot \big([q(y) - q(x)] + [x - y]\big) \leq 0.$$

Mithin ist

$$|q(y) - q(x)|^2 + \big(q(y) - q(x)\big) \cdot (x - y) \leq 0,$$

und es folgt aus der Schwarzschen Ungleichung

$$|q(y) - q(x)|^2 \leq (y - x)\big(q(y) - q(x)\big) \leq |y - x| \cdot |q(y) - q(x)|.$$

Daraus erhalten wir die Behauptung. □

Es sei nun $\bar{B} = \bar{B}(0, r) \subset \mathbb{R}^n$ eine genügend große Kugel mit $Q \subset \bar{B}$. Dann ist $\rho|\bar{B}$ eine stetige Retraktion von $\bar{B}$ auf Q. Damit ist der Allgemeine Brouwersche Fixpunktsatz bewiesen. □

2.3 Der Schaudersche Fixpunktsatz

Es sei $\mathcal{X}$ nun ein reeller normierter Raum. Wir wollen die Brouwerschen Resultate auf solche Räume übertragen. Es sei $F := \{x_1, x_2, \ldots, x_n\}$ eine endliche Teilmenge von $\mathcal{X}$. Dann ist

$$\operatorname{conv}(F) := \left\{ \sum_{i=1}^{n} \lambda_i x_i \;\middle|\; x_i \in F,\ \lambda_i \in [0,1] \ \text{ und } \sum_{i=1}^{n} \lambda_i = 1 \right\}$$

die *konvexe Hülle* von F. Offenbar gilt

$$\operatorname{conv}(F) \subset \operatorname{span}(F),$$

und es ist $\dim \operatorname{span}(F) =: m \leq n$. Daher ist $\operatorname{span}(F)$ zu $\mathbb{R}^m$ kongruent. Es sei noch bemerkt, daß $\operatorname{conv}(F)$ Durchschnitt aller konvexen Teilmengen von $\mathcal{X}$ ist, die F enthalten.

Als erstes zeigen wir einen nach Juliusz Pawel Schauder, 1899–1943, benannten Satz.

Der Schaudersche Projektionssatz: *Es sei $K \subset \mathcal{X}$ eine präkompakte Teilmenge. Dann gibt es zu jedem $\varepsilon > 0$ eine endliche Teilmenge $F \subset \mathcal{X}$ und eine stetige Abbildung*

$$P \ : \ K \longrightarrow \operatorname{conv}(F),$$

die Schauderprojektion, *mit*

$$\forall\, x \in K \qquad \|P(x) - x\| < \varepsilon.$$

Beweis: Zu $\varepsilon > 0$ wählen wir für K das endliche ε-Netz $F = \{x_1, \ldots, x_m\}$. Für $i = 1, \ldots, m$ seien

$$\begin{aligned} \varphi_i \ &: \ K \longrightarrow \mathbb{R}, \\ &x \longmapsto \begin{cases} \varepsilon - \|x - x_i\| & \text{für } x \in B(x_i, \varepsilon) \\ 0 & \text{sonst} \end{cases} \end{aligned}$$

und

$$\begin{aligned} \varphi \ &: \ K \longrightarrow \mathbb{R}, \\ &x \longmapsto \sum\nolimits_{i=1}^{m} \varphi_i(x). \end{aligned}$$

Dann ist

$$\forall\, x \in K \qquad \varphi(x) > 0.$$

Es sei schließlich

$$\begin{aligned} P \ &: \ K \longrightarrow \operatorname{conv}(F), \\ &x \longmapsto \sum\nolimits_{i=1}^{m} \frac{\varphi_i(x)}{\varphi(x)} x_i. \end{aligned}$$

P ist stetig, weil alle φ_i es sind und φ nicht verschwindet. Es folgt

$$\|P(x)-x\| = \Big\| \sum_{i=1}^{m} \frac{\varphi_i(x)}{\varphi(x)}(x_i - x) \Big\| \le \sum_{i=1}^{m} \frac{\varphi_i(x)}{\varphi(x)} \|x_i - x\| < \varepsilon,$$

weil $\varphi_i(x)$ für $\|x_i - x\| \ge \varepsilon$ verschwindet. □

Die Idee ist nun, mit Hilfe dieses Projektionssatzes den Brouwerschen Satz auf den unendlich dimensionalen Fall zu übertragen. Dabei müssen wir allerdings etwas vorsichtig sein. Wir wissen bereits, daß in unendlich dimensionalen Räumen die Sphäre nicht kompakt ist. In der Fixpunkttheorie gilt analog:

Beispiel 2.3.1: *Es gibt einen unendlich dimensionalen normierten Raum $\mathcal{X}$ mit einer abgeschlossenen, beschränkten und konvexen Teilmenge $C \subset \mathcal{X}$ ohne Fixpunkteigenschaft.*

Beweis: Es seien $\mathcal{X} = \ell^2$, $C := \bar{B}(o,1)$, $x = (\xi_1, \xi_2, \dots)$ und

$$f(x) := \Big(\sqrt{1 - \|x\|^2}, \xi_1, \xi_2, \dots \Big).$$

f ist als Komposition stetiger Funktionen selbst stetig. Es ist $\|f(x)\|^2 = 1$, also

$$f \;:\; C \longrightarrow S_{\mathcal{X}} \subset C.$$

f hat aber keinen Fixpunkt. Denn es sei $\hat{x} = f(\hat{x})$. Dann ist $\hat{x} \in S_{\mathcal{X}}$, und

$$(\hat{\xi}_1, \hat{\xi}_2, \dots) = f(\hat{\xi}_1, \hat{\xi}_2, \dots) = (0, \hat{\xi}_1, \hat{\xi}_2, \dots),$$

also $\hat{x} = o$. Das ist ein Widerspruch. □

Um diese Schwierigkeit zu beheben, fordert man, daß das Bild von f relativ kompakt sein soll.

Definition 2.3.2: *Es seien $\mathcal{X}, \mathcal{Y}$ normierte Räume und*

$$f \;:\; \mathcal{X} \longrightarrow \mathcal{Y}$$

eine stetige Abbildung. Dann heißt f vollstetig, *wenn für jede beschränkte Menge $B \subset \mathcal{X}$ das Bild $f(B)$ in $\mathcal{Y}$ relativ kompakt ist.*

Dann gilt

Der Schaudersche Fixpunktsatz: *Es seien C eine beschränkte, abgeschlossene und konvexe Teilmenge des normierten Raumes $\mathcal{X}$ und*

$$f \;:\; C \longrightarrow C$$

vollstetig. Dann besitzt f einen Fixpunkt.

Beweis: Es sei $K := \overline{f(C)}$. Nach Voraussetzung ist $K \Subset C$. Für jedes feste n sei $F_n \subset C$ ein endliches $1/n$-Netz für K und $P_n : K \to \operatorname{conv}(F_n)$ die Schauderprojektion. Dann ist $\operatorname{conv}(F_n) \subset C$. Es seien $f_n := \big(P_n \circ f\big)\big|\operatorname{conv}(F_n)$, also

$$f_n \;:\; \operatorname{conv}(F_n) \longrightarrow \operatorname{conv}(F_n).$$

f_n besitzt einen Fixpunkt y_n, das folgt aus dem Allgemeinen Brouwerschen Fixpunktsatz. Weil f vollstetig ist, enthält $\big(f(y_n)\big)$ eine konvergente Teilfolge. Es sei $\big(f(y_n)\big)$ bereits diese Teilfolge mit

$$f(y_n) \to y \in C.$$

Dann ist y ein Fixpunkt von f. Wegen $f(y_n) \in K$ ist nämlich

$$\Big\| \underbrace{P_n\big(f(y_n)\big)}_{= f_n(y_n)} - f(y_n) \Big\| < \frac{1}{n}.$$

Mithin konvergieren $\big(f_n(y_n)\big) = (y_n)$ und $\big(f(y_n)\big)$ gegen y. Weil f stetig ist, folgt $f(y) = y$. □

Natürlich enthält der Schaudersche Fixpunktsatz den Brouwerschen. Wir können ihn jetzt auch als

Satz 2.3.3: *Eine kompakte konvexe Teilmenge eines normierten Raumes besitzt die Fixpunkteigenschaft*

formulieren.

Beispiel 2.3.4: *Der* Hilbertkubus

$$K := \left\{ x \in \ell^2 \;\middle|\; |\xi_j| \le \frac{1}{j} \quad \textit{für } j = 1, 2, \ldots \right\}$$

besitzt die Fixpunkteigenschaft.

Beweis: Es seien $x = (\xi_1, \xi_2, \ldots)$ und $x_n := (0, \ldots, 0, \xi_n, \xi_{n+1}, \ldots)$. Dann gilt

$$\forall\, \varepsilon > 0 \quad \exists\, n_0 \in \mathbb{N} \quad \forall\, n > n_0 \qquad \|x_n\|^2 = \sum_{j=n}^{\infty} |\xi_j|^2 \le \sum_{j=n}^{\infty} \frac{1}{j^2} < \varepsilon^2 .$$

Zu vorgegebenem $\varepsilon > 0$ wähle man also $n_0 = n_0(\varepsilon)$. Dann überdecken $B(o, \varepsilon)$ und ein endliches ε-Netz für $\ell^2(n_0)$ den Hilbertkubus K. Mithin ist dieser präkompakt und damit kompakt. □

In vielen Anwendungen ist es schwierig, eine geeignete kompakte konvexe Teilmenge eines normierten Raumes zu finden. So sucht man nach anderen Voraussetzungen. Eine solche ist die *Leray-Schaudersche Randbedingung*, benannt nach Jean Leray, *1906, und Juliusz Pawel Schauder. Es sei

$$f \;:\; \mathcal{X} \longrightarrow \mathcal{X}$$

eine stetige Abbildung auf einem normierten Raum $\mathcal{X}$. Dann genügt f dieser Bedingung, wenn

$$\exists\, r > 0 \quad \Big(\|x\| = r \implies \forall\, \lambda > 1 \quad f(x) \neq \lambda x \Big) \tag{$*$}$$

gilt. Beispielsweise sei

$$\exists\, r > 0 \quad \Big(\|x\| = r \implies \|f(x)\| \le r \Big) \tag{$**$}$$

richtig. Dann genügt f der Randbedingung $(*)$. Bedingung $(**)$ ist in manchen Anwendungen erfüllt, wenn $\|f(x)\|$ für große $\|x\|$ weniger stark als $\|x\|$ wächst.

Satz 2.3.5: *Es sei f eine vollstetige Abbildung eines reellen normierten Raumes $\mathcal{X}$, die der Leray-Schauderschen Randbedingung $(*)$ genügt. Dann besitzt f einen Fixpunkt.*

Beweis: Es seien $r > 0$ aus der Randbedingung $(*)$ und $C := \bar{B}(0, r) \subset \mathcal{X}$. Dann ist C abgeschlossen, beschränkt und konvex. $f(C)$ ist relativ kompakt, weil C beschränkt ist; f bildet jedoch im allgemeinen C nicht in sich ab. Wir modifizieren deshalb $f|C$ etwas. Es sei $\rho : \mathcal{X} \to C$ eine stetige Retraktion von $\mathcal{X}$ auf C, nämlich

$$\rho(x) := \begin{cases} x & \text{für } x \in C \\ \frac{r}{\|x\|} x & \text{sonst.} \end{cases}$$

Dann betrachten wir $g := (\rho \circ f)|C : C \to C$. Weil ρ eine stetige Retraktion ist, ist mit f auch g vollstetig, und es gibt ein $x \in C$ mit $g(x) = x$.

Wir zeigen $f(x) \in C$: Es sei nämlich $f(x) \notin C$, also $\|f(x)\| > r$. Dann ist

$$x = g(x) = \rho\big(f(x)\big) = \frac{r}{\|f(x)\|} f(x),$$

also

$$\|x\| = \left\| \frac{r}{\|f(x)\|} f(x) \right\| = r$$

und

$$f(x) = \frac{\|f(x)\|}{r} x = \lambda\, x \quad \text{mit } \lambda > 1.$$

Das steht im Widerspruch zur Randbedingung $(*)$.

Mithin ist $f(x) \in C$ und damit

$$x = g(x) = \rho\big(f(x)\big) = f(x).$$

Die Abbildung f besitzt also den Fixpunkt x. □

3 Der Satz von Baire-Hausdorff und Folgerungen

In diesem Kapitel behandeln wir zunächst den Satz von Baire-Hausdorff und ziehen anschließend wichtige Folgerungen daraus. Zum Beispiel zeigen wir das Prinzip von der gleichmäßigen Beschränktheit und den Satz von der inversen Abbildung. Der Satz wurde 1897 von William Osgood, 1864–1943, für den $\mathbb{R}^1$ und 1899 von René Louis Baire, 1874–1932, für den $\mathbb{R}^n$ bewiesen.

Der Satz von Baire-Hausdorff: *Es seien $\mathcal{X} \neq \emptyset$ ein vollständiger metrischer Raum und $\{A_n\}_{n\in\mathbb{N}}$ eine Familie abgeschlossener Teilmengen $A_n \subset \mathcal{X}$ mit*

$$\mathcal{X} \;=\; \bigcup_{n\in\mathbb{N}} A_n.$$

Dann gibt es ein $m \in \mathbb{N}$ mit $\mathring{A}_m \neq \emptyset$.

Beweis: Für alle $n \in \mathbb{N}$ sei $\mathring{A}_n = \emptyset$, also $A_n = \partial A_n$. Dann ist für alle offenen, nicht leeren $U \subset \mathcal{X}$

$$U \backslash A_n$$

ebenfalls nicht leer und offen. Es gibt also Folgen (x_n) und (ε_n) mit

$$\begin{aligned}
&x_1 \in \mathcal{X} \quad 0 < \varepsilon_1 \leq 1 \quad \bar{B}(x_1, \varepsilon_1) \subset \mathcal{X}\backslash A_1,\\
&x_2 \in \mathcal{X} \quad 0 < \varepsilon_2 \leq \tfrac{1}{2} \quad \bar{B}(x_2, \varepsilon_2) \subset B(x_1, \varepsilon_1)\backslash A_2,\\
&\quad\vdots\\
&x_n \in \mathcal{X} \quad 0 < \varepsilon_n \leq \tfrac{1}{n} \quad \bar{B}(x_n, \varepsilon_n) \subset B(x_{n-1}, \varepsilon_{n-1})\backslash A_n,\\
&\quad\vdots
\end{aligned}$$

Deshalb gilt

$$\forall\, n \in \mathbb{N} \quad \forall\, j \geq n \qquad x_j \in B(x_n, \varepsilon_n)$$

und

$$\forall\, n \in \mathbb{N} \qquad \bar{B}(x_n, \varepsilon_n) \cap A_n = \emptyset.$$

Mithin ist (x_n) eine Cauchyfolge. Es sei $x \in \mathcal{X}$ ihr Grenzwert. Dann folgt

$$\forall\, n \in \mathbb{N} \qquad x \in \bar{B}(x_n, \varepsilon_n) \quad \text{und} \quad x \notin A_n,$$

also

$$x \notin \bigcup_{n\in\mathbb{N}} A_n = \mathcal{X}.$$

Das ist ein Widerspruch. □

Der Baire-Hausdorffsche Satz wird auch als

Bairescher Kategoriensatz: *Ein vollständiger metrischer Raum ist von zweiter Kategorie*

formuliert. Dabei heißt A *von erster Kategorie*, wenn A eine abzählbare Vereinigung nirgends dichter Mengen ist, andernfalls heißt A *von zweiter Kategorie*. $A \subset \mathcal{X}$ heißt *nirgends dicht*, wenn

$$(\mathring{\bar{A}}) = \emptyset$$

ist, oder äquivalent, wenn $\mathcal{X}\backslash\bar{A}$ in $\mathcal{X}$ dicht ist.

Zum Beispiel ist $\mathbb{Q} \subset \mathbb{R}$ von erster Kategorie, $\mathbb{R}$ selbst ist von zweiter Kategorie.

3.1 Das Prinzip von der gleichmäßigen Beschränktheit

Wir geben nun eine Reihe von Folgerungen des Baireschen Satzes an und beginnen mit dem

Prinzip von der gleichmäßigen Beschränktheit: *Es seien $\mathcal{X}$ ein vollständiger metrischer Raum, $\mathcal{Y}$ ein normierter Raum und $\mathcal{F} \subset \mathcal{C}(\mathcal{X}, \mathcal{Y})$ eine Familie stetiger Funktionen mit*

$$\forall\, x \in \mathcal{X} \quad \exists\, c(x) \geq 0 \qquad \sup_{f \in \mathcal{F}} \|f(x)\| \leq c(x).$$

Dann gilt

$$\exists\, B(x_0, \varepsilon) \subset \mathcal{X} \quad \exists\, k \geq 0 \quad \forall\, x \in \bar{B}(x_0, \varepsilon) \qquad \sup_{f \in \mathcal{F}} \|f(x)\| \leq k.$$

Aus der punktweisen Beschränktheit der Familie $\mathcal{F}$ folgt also ihre gleichmäßige Beschränktheit in einer Kugel.

Beweis: Es seien

$$A_j := \bigcap_{f \in \mathcal{F}} \left\{ x \in \mathcal{X} \;\middle|\; \|f(x)\| \leq j \right\}.$$

Diese Mengen sind abgeschlossen, und nach Voraussetzung ist

$$\bigcup_{j \in \mathbb{N}} A_j \;=\; \mathcal{X}.$$

Mithin gibt es ein $k \in \mathbb{N}$ mit $\mathring{A}_k \neq \emptyset$ und

$$\forall\, x \in A_k \qquad \sup_{f \in \mathcal{F}} \|f(x)\| \leq k.$$

Man wähle $\bar{B}(x_0, \varepsilon) \subset A_k$. □

Bemerkung 3.1.1: *Für $\mathcal{Y} = \mathbb{R}$ folgt aus*

$$\forall\, x \in \mathcal{X} \quad \exists\, c(x) \geq 0 \qquad \sup_{f \in \mathcal{F}} f(x) \leq c(x)$$

analog

$$\exists\, B(x_0, \varepsilon) \subset \mathcal{X} \quad \exists\, k \geq 0 \quad \forall\, x \in \bar{B}(x_0, \varepsilon) \qquad \sup_{f \in \mathcal{F}} f(x) \leq k.$$

Zum Beweis wähle man nur

$$A_j := \bigcap_{f \in \mathcal{F}} \left\{ x \in \mathcal{X} \;\middle|\; f(x) \leq j \right\}.$$ □

Beispiel 3.1.2: *Es seien $\mathcal{X} = \mathcal{H}$ ein Hilbertraum und (φ_n) eine* schwach beschränkte Folge *aus $\mathcal{H}$. Das heißt, es soll*

$$\forall\, x \in \mathcal{H} \quad \exists\, c(x) \geq 0 \quad \forall\, \varphi_n \in \mathcal{H} \qquad \left|(\varphi_n, x)\right| \leq c(x)$$

gelten. Dann ist (φ_n) beschränkt.

Zum Beweis wählen wir

$$\begin{aligned} f_n \;:\; \mathcal{H} &\longrightarrow \mathbb{R}_0^+, \\ x &\longmapsto |(\varphi_n, x)| \end{aligned}$$

und $\mathcal{F} := \{f_1, f_2, \ldots\}$. Dann ist $\mathcal{F} \subset \mathcal{C}(\mathcal{H}, \mathbb{R})$ wegen

$$|f_n(x_1) - f_n(x_2)| \leq |(\varphi_n, x_1) - (\varphi_n, x_2)| \leq \|\varphi_n\| \cdot \|x_1 - x_2\|.$$

Es gibt also eine Kugel $B(x_0, \varepsilon)$ und ein k mit

$$\forall\, x \in \bar{B}(x_0, \varepsilon) \qquad \sup_{n \in \mathbb{N}} |(\varphi_n, x)| \leq k.$$

O.B.d.A. sei $\|\varphi_n\| \neq 0$, andernfalls ist nichts zu beweisen. Man wähle

$$x_n := x_0 + \varepsilon \frac{\varphi_n}{\|\varphi_n\|} \in \bar{B}(x_0, \varepsilon).$$

Dann folgt

$$\big|(\varphi_n, x_0) + \varepsilon \|\varphi_n\|\big| = \big|(\varphi_n, x_n)\big| \le k,$$

also

$$\varepsilon\|\varphi_n\| \le \big|(\varphi_n, x_0)\big| + k \le c(x_0) + k$$

oder

$$\forall\, n \in \mathbb{N} \qquad \|\varphi_n\| \le K := \frac{k + c(x_0)}{\varepsilon}. \qquad \square$$

3.2 Der Satz von Banach-Steinhaus

Der folgende Satz wurde im Jahre 1927 von Stefan Banach, und Hugo Steinhaus, 1887–1972, veröffentlicht. Er enthält die Spezialisierung des Prinzips von der gleichmäßigen Beschränktheit auf *lineare Abbildungen.*

Satz von Banach-Steinhaus: *Es seien $\mathcal{X}$ ein Banachraum, $\mathcal{Y}$ ein normierter Raum und $\mathcal{F} \subset \mathcal{CL}(\mathcal{X}, \mathcal{Y})$ mit*

$$\forall\, x \in \mathcal{X} \qquad \sup_{f \in \mathcal{F}} \|f(x)\|_{\mathcal{Y}} < \infty.$$

Dann gilt

$$\sup_{f \in \mathcal{F}} \|f\|_{\mathcal{CL}(\mathcal{X},\mathcal{Y})} < \infty.$$

Aus der punktweisen Beschränktheit folgt also wieder die gleichmäßige.

Beweis: Es gibt eine Kugel $B(x_0, r)$ und ein k mit

$$\forall\, x \in \bar{B}(x_0, r) \quad \forall\, f \in \mathcal{F} \qquad \|f(x)\| \le k.$$

Es sei nun $x \in \mathcal{X}$ mit $\|x\| = 1$. Dann ist

$$\|f(x)\| = \frac{1}{r}\|f(rx)\| \le \frac{1}{r}\Big\{\|f(rx + x_0)\| + \|f(x_0)\|\Big\} \le \frac{2k}{r},$$

also $\|f\| \le 2k/r$. $\square$

Wir geben einige Varianten dieses Satzes an:

Folgerung 3.2.1: *Es seien $\mathcal{X}$ ein Banachraum und $\mathcal{F} \subset \mathcal{X}' = \mathcal{CL}(\mathcal{X}, \mathbb{K})$. Dann gilt*

$$\Big(\forall\, x \in \mathcal{X} \quad \sup_{f \in \mathcal{F}} |f(x)| < \infty\Big) \implies \sup_{f \in \mathcal{F}} \|f\|_{\mathcal{X}'} < \infty.$$

Folgerung 3.2.2: *Es seien $\mathcal{X}$ ein normierter Raum und $M \subset \mathcal{X}$ eine Teilmenge mit*

$$\forall\, x' \in \mathcal{X}' \qquad \sup_{x \in M} |x'x| < \infty.$$

Dann ist M beschränkt.

Beweis: Nach Satz 1.2.13 ist $\mathcal{X}' = \mathcal{CL}(\mathcal{X}, \mathbb{K})$ vollständig, weil $\mathbb{K}$ es ist. Es sei

$$\mathcal{F} := J_{\mathcal{X}} M \subset \mathcal{X}'' = \mathcal{CL}(\mathcal{X}', \mathbb{K}).$$

Für $f = x'' = J_{\mathcal{X}} x \in \mathcal{F}$ ist dann

$$|fx'| = |(J_{\mathcal{X}} x)x'| = |x'x| \le c(x').$$

Es gibt also ein k mit

$$\forall\, f \in \mathcal{F} \qquad \|f\|_{\mathcal{X}''} \le k$$

oder

$$\forall\, x \in M \qquad \|J_{\mathcal{X}} x\|_{\mathcal{X}''} = \|x\|_{\mathcal{X}} \le k. \qquad \square$$

Folgerung 3.2.3: *Es seien $\mathcal{X}$ ein Banachraum, $\mathcal{Y}$ ein normierter Raum und $\mathcal{F} \subset \mathcal{CL}(\mathcal{X}, \mathcal{Y})$ mit*

$$\forall\, y' \in \mathcal{Y}' \quad \forall\, x \in \mathcal{X} \qquad \sup_{f \in \mathcal{F}} |y' f(x)| < \infty.$$

Dann gilt

$$\exists\, k \geq 0 \quad \forall\, f \in \mathcal{F} \qquad \|f\| \leq k.$$

Beweis:

1. Es seien $x \in \mathcal{X}$ fest und

$$M(x) := \{ f(x) \mid f \in \mathcal{F} \} \subset \mathcal{Y}.$$

Dann gilt

$$\forall\, y' \in \mathcal{Y}' \qquad \sup_{y \in M(x)} |y' y| < \infty,$$

und daraus folgt $\|M(x)\| \leq c(x) < \infty$.

2. Es gilt also

$$\forall\, x \in \mathcal{X} \qquad \sup_{f \in \mathcal{F}} \|f(x)\| < \infty,$$

und daraus folgt die Behauptung. □

Folgerung 3.2.4: *Es seien $\mathcal{X}$ ein Banachraum und (x'_n) eine Folge aus $\mathcal{X}'$ mit*

$$\exists\, x' \in \mathcal{X}' \quad \forall\, x \in \mathcal{X} \qquad (x'_n - x', x) \to 0.$$

Dann ist (x'_n) beschränkt.

Beweis: Es gilt

$$\forall\, x \in \mathcal{X} \qquad |x'_n x| \leq c(x) < \infty,$$

also

$$\sup_{n \in \mathbb{N}} \|x'_n\| \leq k.$$ □

Folgerung 3.2.5: *Es seien $\mathcal{X}$ ein normierter Raum und (x_n) eine Folge aus $\mathcal{X}$ mit*

$$\exists\, x \in \mathcal{X} \quad \forall\, x' \in \mathcal{X}' \qquad (x', x_n - x) \to 0.$$

Dann ist (x_n) beschränkt.

Beweis: Aus

$$\forall\, x' \in \mathcal{X}' \qquad (J_{\mathcal{X}} x_n - J_{\mathcal{X}} x, x') \to 0$$

folgt die Beschränktheit von $(J_{\mathcal{X}} x_n)$ in $\mathcal{X}''$. Damit ist

$$\sup_{n \in \mathbb{N}} \|x_n\|_{\mathcal{X}} = \sup_{n \in \mathbb{N}} \|J_{\mathcal{X}} x_n\|_{\mathcal{X}''} \leq k < \infty.$$ □

3.3 Die Sätze von der offenen und der inversen Abbildung

Auch die nächsten Anwendungen betreffen lineare Abbildungen.

Der Satz von der offenen Abbildung: *Es seien $\mathcal{X}, \mathcal{Y}$ Banachräume und $A \in \mathcal{CL}(\mathcal{X}, \mathcal{Y})$. Dann gilt*

$$A \text{ ist surjektiv} \iff A \text{ ist offen.}$$

Beweis:

$\Leftarrow$: Es sei A offen, das heißt

$$\exists\, \delta > 0 \qquad B(o, \delta) \subset A\, B(o, 1).$$

Das bedeutet

$$\forall\, r > 0 \qquad B(o, r) \subset A\, B(o, \tfrac{r}{\delta}).$$

Damit ist A surjektiv.

$\Rightarrow$: Es sei A surjektiv. Dann gehen wir in zwei Schritten vor:

1. Wir zeigen
$$\exists\, \delta > 0 \qquad B(o,\delta) \subset \overline{AB(o,1)}. \tag{$*$}$$
Aus der Surjektivität von A folgt nämlich
$$\mathcal{Y} = \bigcup_{n\in\mathbb{N}} \overline{AB(o,n)}.$$
Mithin gilt
$$\exists\, k > 0 \quad \exists\, B(y,\varepsilon) \qquad B(y,\varepsilon) \subset \overline{AB(o,k)}.$$
Wir reduzieren das auf den Nullpunkt. Es ist $B(o,\varepsilon) = \big\{ -y + z \,\big|\, z \in B(y,\varepsilon)\big\}$, also mit $y = Ax$
$$B(o,\varepsilon) \subset \Big\{-y+z \;\Big|\; z \in \overline{AB(o,k)}\Big\} = \Big\{-Ax+z \;\Big|\; z \in \overline{AB(o,k)}\Big\} \subset \overline{AB(o,k+\|x\|)}.$$
Daraus folgt für $\delta := \varepsilon/(k+\|x\|)$ Aussage $(*)$.

2. Auf der rechten Seite von $(*)$ stört der Abschluß. Deshalb verkleinern wir δ und approximieren im Wertebereich. Aussage $(*)$ impliziert
$$\forall\, y \in \mathcal{Y}, \|y\| < \delta, \quad \exists\, x \in \mathcal{X}, \|x\| < 1, \qquad \|y - Ax\| < \frac{\delta}{2},$$
also $2(y-Ax) \in B(o,\delta)$. Wir bilden nun zwei Folgen (y_n) in $B(o,\delta) \subset \mathcal{Y}$ und (x_n) in $B(o,1) \subset \mathcal{X}$ mit folgenden Eigenschaften: Es sei $y_0 := y \in B(o,\delta)$ beliebig gewählt und $x_0 \in B(o,1)$ mit $y_0 - Ax_0 \in B(o,\frac{\delta}{2})$. Es seien ferner
$$y_{n+1} := 2(y_n - Ax_n) \in B(o,\delta)$$
und $x_{n+1} \in B(o,1)$ mit $y_{n+1} - Ax_{n+1} \in B(o,\frac{\delta}{2})$. Dann folgt
$$2^{-(n+1)} y_{n+1} = 2^{-n} y_n - A(2^{-n} x_n),$$
also
$$A\sum_{j=0}^{n} 2^{-j} x_j = \sum_{j=0}^{n} \left(2^{-j} y_j - 2^{-(j+1)} y_{j+1}\right) = y - 2^{-(n+1)} y_{n+1} \underset{n\to\infty}{\longrightarrow} y.$$
Wegen $x_n \in B(o,1)$ ist
$$\left(\sum_{j=0}^{n} 2^{-j} x_j\right)_{n\in\mathbb{N}}$$
eine Cauchyfolge. Es sei x ihr Grenzwert. Dann ist $\|x\| \le 2$, und aus der Stetigkeit von A folgt
$$A\,x = \lim_{n\to\infty} A \sum_{j=0}^{n} 2^{-j} x_j = y.$$
Mithin ist
$$B(o,\delta) \subset A\,\bar{B}(o,2) \subset A\,B(o,3)$$
oder
$$B\big(o,\tfrac{\delta}{3}\big) \subset A\,B(o,1).$$
Das war zu zeigen. □

Als letztes beweisen wir:

Der Satz von der inversen Abbildung: *Es seien $\mathcal{X}, \mathcal{Y}$ Banachräume und $A \in \mathcal{CL}(\mathcal{X},\mathcal{Y})$ eine bijektive Abbildung. Dann ist $A^{-1} \in \mathcal{CL}(\mathcal{Y},\mathcal{X})$.*

Dieses Resultat folgt aus dem Satz von der offenen Abbildung. Als surjektive Abbildung ist A offen, mithin A^{-1} stetig. □

4 Schwache Konvergenz

In den Folgerungen 3.2.4,5 ist uns bereits die schwache Konvergenz begegnet.

Definition 4.0.1: *Es sei $\mathcal{X}$ ein normierter Raum.*

(1) Eine Folge (x_n) in $\mathcal{X}$ heißt gegen $x \in \mathcal{X}$ schwach konvergent,

$$\text{w-}\lim_{n\to\infty} x_n = x,$$

wenn

$$\forall\, x' \in \mathcal{X}' \qquad \lim_{n\to\infty} x' x_n = x' x$$

gilt.

(2) Eine Folge (x'_n) in $\mathcal{X}'$ heißt gegen $x' \in \mathcal{X}'$ schwach-stern konvergent,

$$\text{w}^*\text{-}\lim_{n\to\infty} x'_n = x',$$

wenn

$$\forall\, x \in \mathcal{X} \qquad \lim_{n\to\infty} x'_n x = x' x$$

gilt.

In $\mathcal{X}'$ gibt es natürlich auch die schwache Konvergenz

$$\forall\, x'' \in \mathcal{X}'' \qquad \lim_{n\to\infty} x'' x'_n = x'' x'.$$

Um deutlicher zu unterscheiden, nennt man die übliche Konvergenz einer Folge (x_n) in $\mathcal{X}$ dann auch *starke Konvergenz* oder *Norm-Konvergenz*. Folgerungen 3.2.4,5 besagen also, daß schwach bzw. schwach-stern konvergente Folgen (stark) beschränkt sind, also

Satz 4.0.2:

(1) Es sei $\mathcal{X}$ ein normierter Raum. Dann ist eine schwach konvergente Folge in $\mathcal{X}$ beschränkt.

(2) Es sei $\mathcal{X}$ ein Banachraum. Dann ist eine schwach-stern konvergente Folge in $\mathcal{X}'$ beschränkt.

Ebenfalls gilt

Satz 4.0.3: *Schwache Grenzwerte sind eindeutig bestimmt.*

Beweis:

1. Es seien x_1, x_2 zwei schwache Grenzwerte der Folge (x_n). Dann ist

$$\forall\, x' \in \mathcal{X}' \qquad x'(x_1 - x_2) = 0,$$

und aus dem Satz von Hahn-Banach (Folgerung 1.5.2) folgt $x_1 = x_2$.

2. Es seien x'_1, x'_2 zwei schwach-stern Grenzwerte der Folge (x'_n). Dann ist

$$\forall\, x \in \mathcal{X} \qquad (x'_1 - x'_2)x = 0,$$

und deshalb $x'_1 = x'_2$. □

Bemerkung 4.0.4:

1. In $\mathcal{X}$ impliziert die starke Konvergenz die schwache.

2. In $\mathcal{X}'$ impliziert die starke Konvergenz die schwache und diese die schwach-stern Konvergenz. Ist $\mathcal{X}$ reflexiv, dann entsprechen schwache und schwach-stern Konvergenz einander.

Es sei nämlich w-lim $x'_n = x'$, also

$$\forall\, x'' \in \mathcal{X}'' \qquad x'' x'_n \to x'' x'.$$

Dann gilt erst recht

$$\forall\, x \in \mathcal{X} \qquad (J_{\mathcal{X}} x) x'_n \to (J_{\mathcal{X}} x) x'$$

oder

$$\forall\, x \in \mathcal{X} \qquad x'_n x \to x' x,$$

also w*-lim $x'_n = x'$. Für reflexive $\mathcal{X}$ gilt auch der umgekehrte Schluß. □

Beispiel 4.0.5: *Es sei $\mathcal{X}$ ein Hilbertraum und (u_n) ein Orthonormalsystem in $\mathcal{X}$. Dann ist*

$$\text{w-}\lim_{n\to\infty} u_n = o.$$

Aufgrund des Rieszschen Darstellungssatzes gibt es nämlich zu $x' \in \mathcal{X}'$ ein $f \in \mathcal{X}$ mit

$$x'u_n = (u_n, f)$$

und letzteres konvergiert gegen Null, das besagt die Besselsche Ungleichung. □

Satz 4.0.6: *Es sei $\mathcal{X}$ ein normierter Raum. Dann gilt:*

(1) w-lim $x_n = x$ *impliziert* $\|x\| \le \liminf_{n\to\infty} \|x_n\| < \infty$.

(2) w*-lim $x'_n = x'$ *impliziert* $\|x'\| \le \liminf_{n\to\infty} \|x'_n\|$.

Beweis:

Zu (1): Die Folge $(\|x_n\|)$ ist beschränkt. Es sei $s' \in S' \subset \mathcal{X}'$. Dann folgt aus

$$|s'x| \leftarrow |s'x_n| \le \|s'\| \cdot \|x_n\| = \|x_n\|$$

die Abschätzung

$$|s'x| \le \liminf_{n\to\infty} \|x_n\|,$$

und der Satz von Hahn-Banach (Folgerung 1.5.2) impliziert

$$\|x\| = \sup_{s'\in S'} |s'x| \le \liminf_{n\to\infty} \|x_n\|.$$

Zu (2): $\liminf_{n\to\infty} \|x'_n\|$ existiere (in Banachräumen ist $(\|x'_n\|)$ beschränkt). Es sei $s \in S \subset \mathcal{X}$. Dann ist

$$|x's| \le \liminf_{n\to\infty} \|x'_n\|,$$

und es folgt

$$\|x'\| = \sup_{s\in S} |x's| \le \liminf_{n\to\infty} \|x'_n\|.$$ □

Lemma 4.0.7:

(1) Es sei $\mathcal{X}$ ein normierter Raum. Dann gilt:

$$\left(\text{w-}\lim_{n\to\infty} x_n = x \quad \textit{und} \quad \lim_{n\to\infty} x'_n = x' \right) \implies \lim_{n\to\infty} x'_n x_n = x'x.$$

(2) Es sei $\mathcal{X}$ ein Banachraum. Dann gilt:

$$\left(\text{w*-}\lim_{n\to\infty} x'_n = x' \quad \textit{und} \quad \lim_{n\to\infty} x_n = x \right) \implies \lim_{n\to\infty} x'_n x_n = x'x.$$

Beweis

Zu (1): Die Folge $(\|x_n\|)$ ist beschränkt. Deshalb folgt

$$|x'_n x_n - x'x| \le |x'_n x_n - x'x_n| + |x'x_n - x'x| \to 0.$$

Zu (2): Die Folge $(\|x'_n\|)$ ist beschränkt. Deshalb folgt

$$|x'_n x_n - x'x| \le |x'_n x_n - x'_n x| + |x'_n x - x'x| \to 0.$$ □

4.1 Schwache Kompaktheit

Wir beginnen mit der

Definition 4.1.1: *Es sei $\mathcal{X}$ ein normierter Raum.*

(1) Eine Folge (x_n) in $\mathcal{X}$ heißt schwache Cauchyfolge, *wenn*

$$\forall\, x' \in \mathcal{X}' \qquad (x'x_n)$$

eine Cauchyfolge ist.

(2) $\mathcal{X}$ heißt schwach folgenvollständig, *wenn jede schwache Cauchyfolge gegen ein $x \in \mathcal{X}$ schwach konvergiert.*

(3) Eine Teilmenge $M \subset \mathcal{X}$ heißt schwach folgenkompakt, *wenn jede Folge (x_n) in M eine Teilfolge $(x_{n'})$ besitzt, die schwach gegen ein $x \in M$ konvergiert.*

(4) Entsprechend werden schwach-stern Aussagen erklärt. Insbesondere heißt eine Teilmenge $M' \subset \mathcal{X}'$ schwach-stern folgenkompakt, *wenn jede Folge (x'_n) in M' eine Teilfolge $(x'_{n'})$ besitzt, die schwach-stern gegen ein $x' \in M'$ konvergiert.*

Man beachte, daß wir die Eigenschaften vollständig und kompakt im schwachen Sinne nur für Folgen definiert haben. Am Ende dieses Kapitels wird über den topologischen Hintergrund dazu noch etwas berichtet. Das Hauptresultat dieses Abschnitts ist

Satz 4.1.2: *Es sei $\mathcal{X}$ ein separabler normierter Raum. Dann ist*

$$\bar{B}'(o,1) = \{x' \in \mathcal{X}' \mid \|x'\| \le 1\} \subset \mathcal{X}'$$

schwach-stern folgenkompakt.

Beweis: Es seien $\{x_n\}_{n\in\mathbb{N}}$ in $\mathcal{X}$ dicht und $(x'_n)_{n\in\mathbb{N}}$ eine Folge aus $\bar{B}'(o,1)$. Dann ist $x'_n x_1$ in $\mathbb{K}$ beschränkt, und es gibt eine konvergente Teilfolge $(x'_{n,1}x_1)$. So fahren wir fort. Es sei (x'_n) bereits die Diagonalfolge. Dann konvergiert

$$x'_n x_k \quad \text{für alle } k \in \mathbb{N}.$$

Es sei nun $M := \operatorname{span}(x_1, x_2, \ldots) \subset \mathcal{X}$. Dann existiert

$$\forall\, m \in M \qquad \lim_{n\to\infty} x'_n m =: m'm,$$

und es ist $m' \in \mathcal{L}(M, \mathbb{K})$. Aus

$$|m'm| = \lim_{n\to\infty} |x'_n m| \le \|m\|$$

folgt $\|m'\| \le 1$. Mithin läßt sich m' stetig fortsetzen zu

$$x' \in \mathcal{CL}(\bar{M}, \mathbb{K}) = \mathcal{CL}(\mathcal{X}, \mathbb{K}).$$

Es ist $\|x'\| \le 1$, und für alle $x \in \mathcal{X}$ und $\varepsilon > 0$ gilt

$$|x'x - x'_n x| \le |(x' - x'_n)m| + |(x' - x'_n)(x - m)| \le |(x' - x'_n)m| + 2\|x - m\| < \varepsilon.$$

Dabei wähle man zunächst $m \in M$ mit $\|x - m\| < \varepsilon/4$ und dann $n \in \mathbb{N}$ mit

$$|(x' - x'_n)m| < \frac{\varepsilon}{2}. \qquad \square$$

Dieses Resultat verwendet man oft in der Situation $\mathcal{X} = \mathcal{L}^1(\Omega)$ und $\mathcal{X}' = \mathcal{L}^\infty(\Omega)$. Die vorausgesetzte Separabilität von $\mathcal{X}$ ist wichtig. Hierzu ein Beispiel:

Satz 4.1.3: *$\bar{B}'(o,1) \subset (\ell^\infty)'$ ist nicht schwach-stern folgenkompakt.*

Beweis: Es seien $x = (\xi_1, \xi_2, \ldots) \in \ell^\infty$ und $x'_n \in \bar{B}'(o,1)$ mit

$$x'_n x := \frac{1}{n} \sum_{i=1}^{n} \xi_i.$$

Es möge eine schwach-stern konvergente Teilfolge

$$\left(x'_{n_k}\right)_{k\in\mathbb{N}}$$

geben. O.B.d.A. wählen wir sie so, daß

$$\forall\, k \in \mathbb{N} \qquad n_k \geq 4\, n_{k-1}$$

ist. Dann konvergiert also $x'_{n_k}x$ für jedes $x \in \ell^\infty$. Es sei nun speziell

$$\xi_i := (-1)^k \quad \text{für } n_{k-1} < i \leq n_k$$

mit $n_0 = 0$. Dann folgt

1. für gerades k:

$$\begin{aligned} x'_{n_k}x &= \frac{1}{n_k}\sum_{i=1}^{n_k}\xi_i = \frac{1}{n_k}\Big\{(\xi_1+\cdots+\xi_{n_1})+(\xi_{n_1+1}+\cdots+\xi_{n_2})+\cdots\Big\} \\ &= \frac{1}{n_k}\Big\{-n_1+(n_2-n_1)-(n_3-n_2)\pm\cdots+(n_k-n_{k-1})\Big\} \\ &= \frac{1}{n_k}\Big\{\underbrace{-2n_1+2n_2}_{\geq 0}\ \underbrace{-2n_3+2n_4}_{\geq 0}+\underbrace{\cdots}_{\geq 0}-2n_{k-1}+n_k\Big\} \\ &\geq \frac{1}{n_k}(n_k-2n_{k-1}) = 1-2\frac{n_{k-1}}{n_k} \geq \frac{1}{2}. \end{aligned}$$

2. für ungerades k:

$$\begin{aligned} x'_{n_k}x &= \frac{1}{n_k}\Big\{-n_1+(n_2-n_1)-(n_3-n_2)\pm\cdots-(n_k-n_{k-1})\Big\} \\ &= \frac{1}{n_k}\Big\{-2n_1+\underbrace{2n_2-2n_3}_{\leq 0}+\underbrace{2n_4-2n_5}_{\leq 0}+\underbrace{\cdots}_{\leq 0}+2n_{k-1}-n_k\Big\} \\ &\leq \frac{1}{n_k}(2n_{k-1}-n_k) \;=\; 2\frac{n_{k-1}}{n_k}-1 \leq -\frac{1}{2}. \end{aligned}$$

Mithin konvergiert $(x'_{n_k}x)$ nicht, und das ist ein Widerspruch. □

In reflexiven Räumen sind schwache und schwach-stern Konvergenz äquivalent. Die nächsten Sätze zeigen eine enge Beziehung zwischen den Konzepten der Reflexivität und der schwachen Konvergenz.

Satz 4.1.4: *Es sei $\mathcal{X}$ ein reflexiver normierter Raum. Dann ist $\bar{B}(o,1) \subset \mathcal{X}$ schwach folgenkompakt.*

Beweis: Es seien (x_n) eine Folge aus $\bar{B}(o,1) \subset \mathcal{X}$ und

$$M := \overline{\operatorname{span}(x_1, x_2, \ldots)} \subset \mathcal{X}.$$

M ist ein abgeschlossener separabler Teilraum von $\mathcal{X}$. Der Raum $\mathcal{X}$ ist reflexiv; damit ist auch jeder abgeschlossene Teilraum von $\mathcal{X}$ reflexiv, M ist also reflexiv.

$$J_M \;:\; M \longrightarrow M''$$

ist deshalb eine Kongruenz und M'' daher separabel Dann ist auch M' separabel nach Satz 1.5.3. Wenden wir Satz 4.1.2 auf M' an, dann sehen wir, daß $\bar{B}''(o,1) \subset M''$ schwach-stern folgenkompakt ist. Mithin gibt es eine Teilfolge (x_{n_k}) und ein $x'' \in M''$ mit $\|x''\| \leq 1$ und

$$\forall\, m' \in M' \qquad (J_M x_{n_k})m' \to x''m'.$$

Es sei

$$x := J_M^{-1}x'' \in M \subset \mathcal{X}.$$

Dann gilt $\|x\| \leq 1$ und

$$\forall\, m' \in M' \qquad m'x_{n_k} \to (J_M x)m' = m'x.$$

Wegen $\mathcal{X}' \subset M'$ gilt die letzte Aussage erst recht für alle $x' \in \mathcal{X}'$. □

Bemerkung 4.1.5: *In Banachräumen $\mathcal{X}$ gilt auch die Umkehrung. Es sei $\bar{B}(o,1) \subset \mathcal{X}$ schwach folgenkompakt. Dann ist $\mathcal{X}$ reflexiv.*

Einen Beweis findet man in *Yosida [1965], S. 141.*

Folgerung 4.1.6: *Es sei $\mathcal{X}$ ein reflexiver normierter Raum. Dann ist $\mathcal{X}$ schwach folgenvollständig.*

Beweis: Es sei (x_n) eine schwache Cauchyfolge. Dann folgt aus

$$|x'x_n| \leq |x'(x_n - x_m)| + |x'x_m| \leq c(m)$$

die schwache Beschränktheit von (x_n), und aus dem Prinzip von der gleichmäßigen Beschränktheit (Folgerung 3.2.2) die starke. Der Rest folgt aus Satz 4.1.4. □

4.2 Schwache Konvergenz in speziellen Räumen

Wir wollen uns nun die erhaltenen Resultate in speziellen Räumen ansehen und sie ausbauen. Wir beginnen mit

Satz 4.2.1: *$\mathcal{C}([0,1],\mathbb{R})$ ist nicht schwach folgenvollständig.*

Zum Beweis wählen wir $x_n(t) := t^n$ für $t \in [0,1] =: J$. Es sei

$$\mathcal{NBV}(J) := \left\{ f \in \mathcal{BV}(J) \;\middle|\; f(0) = 0 \quad \text{und} \quad \forall\, t \in (0,1) \quad f(t+0) = f(t) \right\}.$$

Dabei ist $f(t+0) := \lim_{\epsilon \to 0} f(t+\epsilon)$. Nun ist $\mathcal{C}(J)' = \mathcal{NBV}(J)$, und ein $x'x$ läßt sich in der Form

$$x'x = \int_0^1 x(t)\, d\mu(t)$$

mit einem geeigneten $\mu \in \mathcal{NBV}(J)$ darstellen (vgl. *Taylor & Lay [1980], S. 146*). Ein $\mu \in \mathcal{NBV}(J)$ ist Differenz zweier monoton wachsender Abbildungen $g, h : J \to \mathbb{R}$. Unsere Folge $\big(x_n(t)\big)$ ist nicht negativ und fällt monoton. Daraus folgt, daß auch

$$\int_0^1 x_n(t)\, dg(t) \quad \text{und} \quad \int_0^1 x_n(t)\, dh(t)$$

monoton fallen. Weil sie beschränkt sind, konvergieren deshalb beide Integrale. Mithin konvergiert $(x'x_n)$ für jedes $x' \in \mathcal{C}'(J)$.

Wir zeigen nun, daß (x_n) nicht schwach gegen einen Grenzwert konvergiert. Punktweise gilt

$$x_n(t) \to y(t) := \begin{cases} 0 & \text{für } 0 \leq t < 1 \\ 1 & \text{für } t = 1. \end{cases}$$

Offenbar ist $y \notin \mathcal{C}(J)$. Es gebe nun ein $x \in \mathcal{C}(J)$ mit

$$\forall\, x' \in \mathcal{C}(J)' \qquad x'x_n \to x'x.$$

Dann wählen wir zu $t_0 \in (0,1)$

$$\mu(t) := \begin{cases} 0 & \text{für } 0 \leq t < t_0 \\ 1 & \text{für } t_0 \leq t \leq 1 \end{cases}$$

und erhalten

$$0 \leftarrow \int_0^1 \big(x_n(t) - x(t)\big)\, d\mu(t) = x_n(t_0) - x(t_0) \to y(t_0) - x(t_0).$$

Daraus folgt $x = y$, und das ist ein Widerspruch. □

Satz 4.2.2: *Es seien $\mathcal{X}$ ein Prähilbertraum, (x_n) eine Folge in $\mathcal{X}$ und $x \in \mathcal{X}$. Dann sind äquivalent:*

(1) $x_n \to x$.

(2) $\text{w-lim}_{n\to\infty} x_n = x$ *und* $\|x_n\| \to \|x\|$.

Beweis: (1) $\Rightarrow$ (2) ist klar.

(2) $\Rightarrow$ (1): Aus

$$\operatorname*{w\text{-}lim}_{n\to\infty}(x_n - x) = 0$$

folgt

$$\forall\, y \in \mathcal{X} \qquad (x_n - x, y) \to 0$$

und daraus

$$\|x_n - x\|^2 = \|x_n\|^2 - 2\,\mathrm{Re}\,(x_n, x) + \|x\|^2 \to 0. \qquad \square$$

Satz 4.2.3: *Es seien $1 < p < \infty$, (x_n) eine Folge in ℓ^p und $x = (\xi_1, \xi_2, \ldots) \in \ell^p$. Dann sind äquivalent:*

(1) $\operatorname{w-lim}_{n\to\infty} x_n = x$.

(2) *(x_n) ist in ℓ^p beschränkt, und es gilt $\xi_i^{(n)} \to \xi_i$ für alle i.*

Beweis: (1) $\Rightarrow$ (2) ist wieder klar.

Wir zeigen (2) $\Rightarrow$ (1): Analog zur Rieszabbildung in §1.5.3 wählen wir

$$\begin{aligned} R \;:\; \ell^q &\longrightarrow (\ell^p)', \\ x &\longmapsto \sum\nolimits_{i=1}^{\infty} \bar{\xi}_i e_i, \end{aligned}$$

also

$$(Rx)y = \sum_{i=1}^{\infty} \bar{\xi}_i \eta_i \quad \text{für } y = (\eta_1, \eta_2, \ldots) \in \ell^p.$$

Dabei sollen e_i der i-te Einheitsvektor und

$$\frac{1}{p} + \frac{1}{q} = 1$$

sein. Die Abbildung R ist eine konjugiert lineare Kongruenz. Ferner sei

$$M' := \operatorname{span}(Re_1, Re_2, \ldots) \subset (\ell^p)'.$$

Es sei nun (x_n) eine Folge aus ℓ^p, die der Voraussetzung (2) genügt. ℓ^p ist reflexiv, nach Satz 4.1.4 ist $\bar{B}(o,1) \subset \ell^p$ schwach folgenkompakt. Es gibt also ein $x \in \ell^p$ und eine Teilfolge von (x_n), die schwach gegen x konvergiert. Wegen

$$(Re_i)x_n = \xi_i^{(n)} \to \xi_i = (Re_i)x,$$

also

$$\forall\, m' \in M' \qquad m'x_n \to m'x,$$

und weil M' in $(\ell^p)'$ dicht ist, konvergiert jede Teilfolge von (x_n) gegen dasselbe x. Mithin konvergiert (x_n) selbst schwach gegen x. $\square$

Die schwache Konvergenz verallgemeinert also die Konvergenz aller Koordinaten, im Falle endlich dimensionaler Räume sind starke und schwache Konvergenz äquivalent. Überraschend ist aber wohl das folgende Resultat:

Satz 4.2.4: *In ℓ^1 sind starke und schwache Folgenkonvergenz äquivalent.*

Beweis: Wir haben nur eine Richtung zu zeigen. Der Beweis erfolgt über einen Widerspruch. Es sei (x_n) eine schwache Nullfolge aus ℓ^1, die nicht stark konvergiert. Dann folgt aus der Beschränktheit von $(\|x_n\|)$ die Existenz einer Teilfolge mit $\|x_n\| \to p > 0$. O.B.d.A. können wir deshalb von der Situation

$$\operatorname*{w\text{-}lim}_{n\to\infty} x_n = 0 \quad \text{und} \quad \|x_n\| = 1$$

ausgehen. Es sei nun wieder $x_n = (\xi_1^{(n)}, \xi_2^{(n)}, \ldots)$. Dann wählen wir induktiv zwei Zahlenfolgen. Es seien $p_0 := 0$ und

$$1 =: n_1 < n_2 < n_3 < \cdots, \qquad 0 < p_1 < p_2 < p_3 < \cdots$$

mit

$$\sum_{i=1}^{p_{k-1}} \left|\xi_i^{(n_k)}\right| < \frac{1}{4} \quad \text{und} \quad \sum_{i=p_{k-1}+1}^{p_k} \left|\xi_i^{(n_k)}\right| > \frac{3}{4}.$$

Eine solche Wahl ist möglich, denn aus

$$\sum_{i=1}^{\infty} \left|\xi_i^{(1)}\right| = \|x_1\| = 1.$$

folgt für $k = 1$ die Existenz eines $p_1 > 0$ mit

$$\sum_{i=1}^{p_1} \left|\xi_i^{(1)}\right| > \frac{3}{4}.$$

Es seien n_k, p_k für $k \le j$ bereits gewählt. Dann erhalten wir aus

$$\forall\, i \in \mathbb{N} \quad \xi_i^{(n)} \to 0$$

die Existenz eines $n_{j+1} > n_j$ mit

$$\sum_{i=1}^{p_j} \left|\xi_i^{(n_{j+1})}\right| < \frac{1}{4}.$$

Es folgt

$$\sum_{i=p_j+1}^{\infty} \left|\xi_i^{(n_{j+1})}\right| = \sum_{i=1}^{\infty} \cdots - \sum_{i=1}^{p_j} \cdots > 1 - \frac{1}{4} = \frac{3}{4}.$$

Mithin existiert auch ein $p_{j+1} > p_j$ mit

$$\sum_{i=p_j+1}^{p_{j+1}} \left|\xi_i^{(n_{j+1})}\right| > \frac{3}{4}.$$

Damit sind die Folgen (n_j) und (p_j) definiert. Es sei nun

$$i \longmapsto k(i)$$

mit

$$p_{k-1} < i \le p_k.$$

Es sei ferner

$$y' := \sum_{i=1}^{\infty} \eta_i\, e_i \in (\ell^1)' \quad \text{mit } \eta_i := \operatorname{sign} \xi_i^{(n_k)},$$

also

$$y'x = \sum_{i=1}^{\infty} \eta_i\, \xi_i.$$

Dann gilt für alle k

$$\begin{aligned} |y' x_{n_k}| = \left|\sum_{i=1}^{\infty} \eta_i\, \xi_i^{(n_k)}\right| &\ge \left|\sum_{i=p_{k-1}+1}^{p_k} \cdots\right| - \sum_{i=1}^{p_{k-1}} |\cdots| - \sum_{i=p_k+1}^{\infty} |\cdots| \\ &\ge \sum_{i=p_{k-1}+1}^{p_k} \left|\xi_i^{(n_k)}\right| - \sum_{i=1}^{p_{k-1}} \left|\xi_i^{(n_k)}\right| - \sum_{i=p_k+1}^{\infty} \left|\xi_i^{(n_k)}\right| = 2 \sum_{i=p_{k-1}+1}^{p_k} \left|\xi_i^{(n_k)}\right| - \|x_{n_k}\| \;\ge\; \frac{1}{2}. \end{aligned}$$

Dem widerspricht $y' x_{n_k} \to 0$. □

Es gilt auch

Satz 4.2.5: *Es sei $\mathcal{X} = \mathcal{C}([0,1], \mathbb{R})$. Dann sind äquivalent:*

(1) $\text{w-lim}_{n\to\infty} x_n = x$.

(2) *(x_n) ist beschränkt, und punktweise gilt $x_n(t) \to x(t)$.*

4.3 Konvexe Mengen

Für Hilberträume haben wir bereits den wichtigen Approximationssatz bewiesen, der vielfältig angewandt wird. Wir wollen in diesem Abschnitt ein ähnliches Ergebnis für reflexive Räume zeigen, das ebenfalls für Anwendungen, etwa in $\mathcal{L}^p$-Räumen, sehr interessant ist. Dazu müssen wir uns noch eingehender mit konvexen Mengen beschäftigen. Das soll nun geschehen. Zur Erinnerung, $M \subset \mathcal{X}$ heißt *konvex*, wenn

$$\forall\, \lambda \in [0,1] \quad \forall\, x_1, x_2 \in M \qquad \lambda\, x_1 + (1-\lambda) x_2 \in M$$

ist. Es sei $S \subset \mathcal{X}$ eine Teilmenge. Dann ist

$$\operatorname{conv}(S) := \bigcap_{\substack{M \supset S \\ M \text{ konvex}}} M$$

die *konvexe Hülle* von S. Die Menge $\operatorname{conv}(S)$ ist die kleinste S umfassende konvexe Menge in $\mathcal{X}$.

Als erstes zeigen wir den folgenden

Trennungssatz: *Es seien $\mathcal{X}$ ein normierter Raum, $M \subset \mathcal{X}$ eine abgeschlossene und konvexe Teilmenge und $x_0 \in \mathcal{X} \backslash M$. Dann gilt*

$$\exists\, x' \in \mathcal{X}' \quad \exists\, \alpha \in \mathbb{R} \qquad \operatorname{Re} x' x \begin{cases} \leq \alpha & \text{für } x \in M \\ > \alpha & \text{für } x = x_0. \end{cases}$$

Wir sagen: x' und α *trennen* x_0 von M. Den entsprechenden Satz haben wir in §1.5.1 für Unterräume bereits kennengelernt.

Den Beweis führen wir in drei Schritten.

1. Wir machen zwei zusätzliche Annahmen:

(i) $\mathcal{X}$ ist ein reeller Vektorraum.

(ii) $\exists\, \delta > 0$ mit $\bar{B}(o, \delta) \subset M$.

Für alle $x \in \mathcal{X}$ sei dann

$$p(x) := \inf \left\{ r \in (0, \infty) \;\middle|\; \frac{x}{r} \in M \right\}.$$

Die Abbildung $p : \mathcal{X} \to \mathbb{R}$ ist ein Funktional mit

$$\begin{aligned} \forall\, \alpha \geq 0 \quad \forall\, x \in \mathcal{X} \qquad & p(\alpha x) = \alpha\, p(x), \\ \forall\, x \in M \qquad & p(x) \leq 1, \\ \forall\, x \in \mathcal{X} \qquad & p(x) \leq \|x\| / \delta. \end{aligned}$$

Die letzte Behauptung folgt mit $y := x\delta / \|x\|$ aus $p(y) \leq 1$.

Weil M konvex ist, ist p sogar sublinear. Es gilt also außerdem $p(x+y) \leq p(x) + p(y)$. Mit x/r_x, $y/r_y \in M$ ist nämlich

$$\lambda \frac{x}{r_x} + (1-\lambda)\, \frac{y}{r_y} \in M,$$

also mit $\lambda := r_x/(r_x + r_y)$

$$\frac{1}{r_x + r_y}(x+y) \in M.$$

Damit ist $p(x+y) \leq r_x + r_y$ oder

$$p(x+y) \leq p(x) + p(y).$$

Wegen $x_0 \notin M = \bar{M}$ ist $p(x_0) > 1$. Wir definieren $f \in \mathcal{L}\big(\operatorname{span}(x_0), \mathbb{R}\big)$ durch

$$\forall\, \alpha \in \mathbb{R} \qquad f(\alpha x_0) := \alpha\, p(x_0).$$

Dann gilt

$$\forall\, x \in \operatorname{span}(x_0) \qquad f(x) \leq p(x),$$

und es gibt nach Satz 1.1.17 eine Fortsetzung $F \in \mathcal{L}(\mathcal{X}, \mathbb{R})$ mit

$$\forall\, x \in \mathcal{X} \qquad F(x) \le p(x).$$

Es folgt

$$F(x) \le p(x) \le \frac{1}{\delta}\,\|x\|,$$
$$-F(x) = F(-x) \le p(-x) \le \frac{1}{\delta}\,\|x\|,$$

also $F \in \mathcal{CL}(\mathcal{X}, \mathbb{R})$, $F(x_0) = p(x_0) > 1$ und

$$\forall\, x \in M \qquad F(x) \le p(x) \le 1.$$

F leistet also das Gewünschte.

2. Es gelte nur noch (i). Wir wählen ein $a \in M$, schieben a in den Nullpunkt und blasen M etwas auf, so daß (ii) erfüllt wird: Es seien $\tilde{x}_0 := x_0 - a$ und

$$\tilde{M} := \left\{ x \in \mathcal{X} \;\middle|\; \operatorname{dist}(x + a, M) \le \frac{d}{2} \right\} \quad \text{mit } d := \operatorname{dist}(x_0, M) > 0.$$

Dann ist $\tilde{M} \subset \mathcal{X}$ abgeschlossen und konvex. Mit $x_1, x_2 \in \tilde{M}$ sei nämlich

$$x_3 := \lambda\, x_1 + (1 - \lambda) x_2 = \lambda(x_1 + a) + (1 - \lambda)(x_2 + a) - a,$$

also

$$\operatorname{dist}(x_3 + a, M) = \operatorname{dist}\big(\lambda(x_1 + a) + (1 - \lambda)(x_2 + a), M\big) = \inf_{m \in M} \big\|\lambda(x_1 + a) + (1 - \lambda)(x_2 + a) - m\big\| =$$
$$= \inf_{m_1, m_2 \in M} \big\|\lambda(x_1 + a - m_1) + (1 - \lambda)(x_2 - a - m_2)\big\| \le \lambda\, \tfrac{d}{2} + (1 - \lambda)\, \tfrac{d}{2} = \tfrac{d}{2}.$$

Das heißt, es ist $x_3 \in \tilde{M}$. Ferner gilt

$$\tilde{x}_0 \in \mathcal{X} \backslash \tilde{M} \quad \text{und} \quad \bar{B}\big(0, \tfrac{d}{2}\big) \subset \tilde{M}.$$

Nach (1) gibt es also $x' \in \mathcal{X}'$ und $\tilde{\alpha} \in \mathbb{R}$, die $\tilde{x}_0$ von $\tilde{M}$ trennen. Dann trennen x' und $\alpha := \tilde{\alpha} + x'a$ das Element x_0 von M. Es ist nämlich

$$x' x_0 = x' \tilde{x}_0 + x' a > \tilde{\alpha} + x' a = \alpha,$$

und für alle $m \in M$ gilt mit $\tilde{m} := m - a \in \tilde{M}$

$$x' m = x' \tilde{m} + x' a \le \tilde{\alpha} + x' a = \alpha.$$

3. Wir behandeln schließlich den allgemeinen Fall $\mathcal{X}$ über $\mathbb{C}$. Dabei gehen wir wie beim Beweis des Satzes von Hahn-Banach in §1.5.1 vor und betrachten $\mathcal{X}$ als Vektorraum über $\mathbb{R}$. Folglich gibt es $\tilde{x}' \in \mathcal{CL}(\mathcal{X}, \mathbb{R})$ und $\alpha \in \mathbb{R}$, die x_0 von M trennen. Es sei für $x \in \mathcal{X}$

$$x' x := \tilde{x}' x - \mathrm{i}\tilde{x}'(\mathrm{i}x).$$

Dann ist $x' \in \mathcal{CL}(\mathcal{X}, \mathbb{C})$. x' und α trennen x_0 von M. □

Bevor wir weitermachen bringen wir noch

Definition 4.3.1: *Es sei $\mathcal{X}$ ein normierter Raum. Dann heißt eine Teilmenge $M \subset \mathcal{X}$* schwach folgenabgeschlossen, *wenn für jede gegen ein $x \in \mathcal{X}$ schwach konvergente Folge (x_n) aus M*

$$\operatorname*{w\text{-}lim}_{n \to \infty} x_n = x \in M$$

ist.

Satz 4.3.2: *Es seien $\mathcal{X}$ ein normierter Raum und $M \subset \mathcal{X}$ eine konvexe Teilmenge. Dann sind äquivalent:*

(1) M ist abgeschlossen.

(2) M ist schwach folgenabgeschlossen.

Beweis:
(1) $\Rightarrow$ (2): Es seien $x_0 \in \mathcal{X}$ und (x_n) eine Folge aus M mit

$$\text{w-}\lim_{n\to\infty} x_n = x_0.$$

Es sei $x_0 \notin M$. Dann gibt es ein $x' \in \mathcal{X}'$ und ein $\alpha \in \mathbb{R}$, die x_0 von M trennen, also

$$\operatorname{Re} x'x \begin{cases} \leq \alpha & \text{für } x = x_n \\ > \alpha & \text{für } x = x_0. \end{cases}$$

Daraus folgt der Widerspruch

$$\alpha \geq \operatorname{Re} x'x_n \to \operatorname{Re} x'x_0 > \alpha.$$

(2) $\Rightarrow$ (1): Diese Aussage ist klar. □

Nun kommen wir zum Hauptresultat dieses Abschnitts, nämlich zum

Approximationssatz: *Es seien $\mathcal{X}$ reflexiv und $M \subset \mathcal{X}$ eine abgeschlossene, konvexe und nicht leere Teilmenge. Dann gilt*

$$\forall\, x_0 \in \mathcal{X} \quad \exists\, x \in M \qquad \|x - x_0\| = \operatorname{dist}(x_0, M).$$

Einen solchen Satz haben wir bisher nur in Hilberträumen zeigen können. Dort galt auch die Eindeutigkeit. Über diese Frage werden wir anschließend noch Aussagen machen.

Beweis: Wir wählen eine Minimalfolge (x_n) aus M mit

$$\|x_n - x_0\| \to \operatorname{dist}(x_0, M).$$

Aus $\|x_n\| \leq \|x_n - x_0\| + \|x_0\|$ folgt die Beschränktheit von (x_n). Nach Satz 4.1.4 gibt es eine Teilfolge, die schwach gegen ein $x \in \mathcal{X}$ konvergiert. Es sei bereits (x_n) diese Folge. Nach Satz 4.3.2 ist $x \in M$, und aus Satz 4.0.6 folgt

$$\|x - x_0\| \leq \liminf_{n\to\infty} \|x_n - x_0\| = \operatorname{dist}(x_0, M).$$

Wegen $x \in M$ ist $\|x - x_0\| \geq \operatorname{dist}(x_0, M)$, also

$$\|x - x_0\| = \operatorname{dist}(x_0, M).$$ □

Damit haben wir die Existenz eines Elementes bester Approximation zeigen können. Kommen wir nun zur Frage nach der Eindeutigkeit: Es sei

$$E := \left\{ x \in M \;\middle|\; \|x - x_0\| = \operatorname{dist}(x_0, M) \right\}.$$

Mit $d := \operatorname{dist}(x_0, M)$ und $x_1, x_2 \in E$ folgt dann

$$d \leq \| \underbrace{\lambda x_1 + (1-\lambda)x_2}_{\in M} - x_0 \| = \|\lambda(x_1 - x_0) + (1-\lambda)(x_2 - x_0)\| \leq$$

$$\leq \lambda\|x_1 - x_0\| + (1-\lambda)\|x_2 - x_0\| = d.$$

Die Menge E ist also konvex. Für $x \in E$ sei

$$y := \frac{x - x_0}{d}.$$

Dann ist $y \in S_{\mathcal{X}} \subset \mathcal{X}$. Es seien $x_1 \neq x_2$. Dann folgt mit $y_i = y(x_i)$

$$\exists\, \varepsilon > 0 \quad \text{mit} \quad 0 < \varepsilon \leq \|y_1 - y_2\| \leq 2 \quad \text{und} \quad \left\| \frac{y_1 + y_2}{2} \right\| = 1. \qquad (*)$$

Daraus soll $y_1 = y_2$ folgen. Eine hinreichende Bedingung dafür geben wir in der folgenden

Definition 4.3.3: *Es sei $(\mathcal{X}, \|\cdot\|)$ ein normierter Raum. Dann heißt die Norm* gleichmäßig konvex, *wenn*

$$\forall\, \varepsilon > 0 \quad \exists\, \delta > 0 \quad \forall\, x, y \in S_{\mathcal{X}},\ \|x - y\| \geq \varepsilon, \qquad \left\| \frac{x + y}{2} \right\| \leq 1 - \delta(\varepsilon)$$

gilt. Den Raum $\mathcal{X}$ selbst nennt man dann ebenfalls gleichmäßig konvex.

Folgerung 4.3.4: *Es seien die Voraussetzungen des Approximationssatzes erfüllt und außerdem $\mathcal{X}$ ein gleichmäßig konvexer Raum. Dann gibt es genau ein Element bester Approximation.*

Ein Beispiel für gleichmäßig konvexe Räume sind natürlich die Hilberträume $\mathcal{H}$. Für $x, y \in S \subset \mathcal{H}$ gilt die Parallelogrammgleichung

$$\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right) = 4,$$

also

$$\left\|\frac{x+y}{2}\right\| = \sqrt{1 - \left\|\frac{x-y}{2}\right\|^2} \leq 1 - \frac{1}{2}\left\|\frac{x-y}{2}\right\|^2 = 1 - \frac{1}{8}\|x-y\|^2.$$

Zu $\varepsilon > 0$ wähle man also $\delta(\varepsilon) := \varepsilon^2/8$.

Gleichmäßig konvex sind ebenfalls die $\mathcal{L}^p(\Omega)$-Räume für $1 < p < \infty$. Die gleichmäßig konvexen Räume sind spezielle reflexive Räume, man vergleiche *Riesz & Nagy [1956], S. 394*. Es gibt aber reflexive Räume, die nicht gleichmäßig konvex sind, man vergleiche *Day [1941]*.

Wir beschließen diesen Abschnitt mit zwei verwandten Ergebnissen. Das erste wird nach Stanislaw Mazur, 1905–1981, benannt.

Der Satz von Mazur: *Es seien $\mathcal{X}$ ein normierter Raum und (x_n) eine Folge aus $\mathcal{X}$ mit*

$$\text{w-}\lim_{n\to\infty} x_n = x \in \mathcal{X}.$$

Dann gilt

$$\forall\, \varepsilon > 0 \quad \exists\, n \in \mathbb{N} \quad \exists\, y \in \operatorname{conv}(x_1, \ldots, x_n) \qquad \|y - x\| < \varepsilon.$$

Beweis: Es seien

$$M_n := \operatorname{conv}(x_1, \ldots, x_n) \quad \text{und} \quad M := \operatorname{conv}(x_1, x_2, \ldots).$$

Dann ist

$$M = \bigcup_{n\in\mathbb{N}} M_n.$$

$\bar{M}$ ist konvex und abgeschlossen, nach Satz 4.3.2 also schwach folgenabgeschlossen. Mithin ist $x \in \bar{M}$, und es gibt ein $y \in M$ mit $\|y - x\| < \varepsilon$. Dann gilt aber

$$\exists\, n \in \mathbb{N} \qquad y \in M_n. \qquad \square$$

Das zweite Resultat geht auf Banach und Stanislaw Saks, 1897–1942, zurück.

Der Satz von Banach-Saks: *Es seien $\mathcal{X} = \mathcal{L}^p(\Omega)$ mit $1 < p < \infty$, $f \in \mathcal{X}$ und (f_n) eine Folge aus $\mathcal{X}$, die schwach gegen f konvergiert. Dann gibt es eine Teilfolge (f_{n_k}) mit*

$$\lim_{k\to\infty} \left\| \frac{f_{n_1} + f_{n_2} + \cdots + f_{n_k}}{k} - f \right\| = 0.$$

Ein Beweis findet sich in *Riesz & Nagy [1956], S. 72.*

4.4 Schwache Topologien

Es folgen nun noch einige Hinweise zu schwachen Topologien. Es seien $\mathcal{X}$ eine Menge, $(\mathcal{Y}, \mathcal{T}_\mathcal{Y})$ ein topologischer Raum und $\mathcal{F}$ eine Familie von Abbildungen $f : \mathcal{X} \to (\mathcal{Y}, \mathcal{T}_\mathcal{Y})$. Dann nennt man die schwächste Topologie von $\mathcal{X}$, bezüglich der alle $f \in \mathcal{F}$ stetig sind, die $\mathcal{F}$-schwache Topologie von $\mathcal{X}$. Sind $(\mathcal{X}_\lambda, \mathcal{T}_\lambda)$ für $\lambda \in \Lambda$ topologische Räume, ist

$$\pi_\mu \;:\; \prod_{\lambda\in\Lambda} \mathcal{X}_\lambda \longrightarrow \mathcal{X}_\mu$$

die Projektion auf X_μ und $\mathcal{F} = \{\pi_\lambda \mid \lambda \in \Lambda\}$, dann erhält man auf diese Weise die schwache Produkttopologie für das kartesische Produkt $\prod_{\lambda\in\Lambda} \mathcal{X}_\lambda$.

Es sei nun $\mathcal{X}$ ein normierter Raum mit der von der Norm erzeugten Topologie $\mathcal{T}_\mathcal{X}$. Dann ist die *schwache Topologie* $\mathcal{T}(\mathcal{X}, \mathcal{X}')$ von $\mathcal{X}$ die schwächste Topologie, in der alle $x' \in \mathcal{X}'$ stetig sind, also

$$\mathcal{T}(\mathcal{X}, \mathcal{X}') := \bigcap \left\{ t \in T(\mathcal{X}) \;\middle|\; \text{alle } x' \in \mathcal{X}' \text{ sind bzgl. } t \text{ stetig} \right\} \subset \mathcal{T}_\mathcal{X}.$$

Dabei ist $T(\mathcal{X})$ die Menge aller Topologien von $\mathcal{X}$.

Für $x \in \mathcal{X}$, $E \subset \mathcal{X}'$ und $\varepsilon > 0$ sei

$$U(x, E, \varepsilon) := \Big\{ y \in \mathcal{X} \;\Big|\; \forall\, x' \in E \quad |x'y - x'x| < \varepsilon \Big\}.$$

Dann ist

$$\big\{ U(x, E, \varepsilon) \;\big|\; x \in \mathcal{X},\, E \subset \mathcal{X}' \text{ endlich},\, \varepsilon > 0 \big\}$$

eine Basis von x in $\mathcal{T}(\mathcal{X}, \mathcal{X}')$. Denken Sie an Röhren, nicht an Kugeln.

Entsprechend definiert man

$$\mathcal{T}(\mathcal{X}', \mathcal{X}'') := \bigcap \Big\{ t \in T(\mathcal{X}') \;\Big|\; \text{alle } x'' \in \mathcal{X}'' \text{ sind bzgl. } t \text{ stetig} \Big\}$$

und

$$\mathcal{T}(\mathcal{X}', J_{\mathcal{X}}\mathcal{X}) := \mathcal{T}(\mathcal{X}', \mathcal{X}) := \bigcap \Big\{ t \in T(\mathcal{X}') \;\Big|\; \text{alle } x'' \in J_{\mathcal{X}}\mathcal{X} \subset \mathcal{X}'' \text{ sind bzgl. } t \text{ stetig} \Big\}.$$

Es gilt also

$$\mathcal{T}(\mathcal{X}', \mathcal{X}) \subset \mathcal{T}(\mathcal{X}', \mathcal{X}'') \subset \mathcal{T}_{\mathcal{X}'}.$$

$\mathcal{T}(\mathcal{X}', \mathcal{X}'')$ ist die schwache Topologie von $\mathcal{X}'$ und $\mathcal{T}(\mathcal{X}', \mathcal{X})$ die schwach-stern Topologie. Sie ist die schwächste der gerade angegebenen Topologien von $\mathcal{X}'$. In endlich dimensionalen Räumen sind diese Topologien natürlich gleich.

Eine Topologie $\mathcal{T}$ auf $\mathcal{X}$ induziert einen Konvergenzbegriff für Folgen (x_n) aus $\mathcal{X}$.

Definition 4.4.1: *Eine Folge (x_n) aus $\mathcal{X}$ heißt gegen $x \in \mathcal{X}$* konvergent, *wenn*

$$\forall\, O, \textit{ offen mit } x \in O, \quad \exists\, n_0 \in \mathbb{N} \quad \forall\, n \geq n_0 \qquad x_n \in O$$

gilt.

Eine Folge kann mehrere Grenzwerte haben. In der trivialen Topologie $\mathcal{T} = \{\emptyset, \mathcal{X}\}$ ist jedes x_n Grenzwert.

Satz 4.2.4 zeigt, daß die Konvergenzbegriffe für Folgen bezüglich $\mathcal{T}_{\mathcal{X}}$ und $\mathcal{T}(\mathcal{X}, \mathcal{X}')$ übereinstimmen können ($\mathcal{X} = \ell^1$). Die Topologien $\mathcal{T}_{\ell^1}$ und $\mathcal{T}(\ell^1, \ell^\infty)$ sind jedoch verschieden. (Vgl. Bemerkung 4.4.3).

In einem metrischen Raum sind die Eigenschaften (Überdeckungs-) kompakt und folgenkompakt äquivalent (Hauptsatz in §1.2.4) Das gilt in allgemeinen topologischen Räumen aber nicht mehr. Man hat

Satz 4.4.2 (von Alaoglu): *Es sei $\mathcal{X}$ ein normierter Raum. Dann ist die abgeschlossene Kugel $\bar{B}'(o, 1)$ in $\mathcal{X}'$ bezüglich $\mathcal{T}(\mathcal{X}', \mathcal{X})$ kompakt.*

In $\mathcal{X} = \ell^\infty$ ist $\bar{B}'(o, 1) \subset \mathcal{X}'$ nicht schwach-stern folgenkompakt (Satz 4.1.3). $\bar{B}'(o, 1) \subset \mathcal{X}'$ ist jedoch schwach-stern folgenkompakt, wenn $\mathcal{X}$ separabel ist (Satz 4.1.2). Ist $\mathcal{X}$ reflexiv, dann ist $\bar{B}(o, 1) \subset \mathcal{X}$ schwach folgenkompakt (Satz 4.1.4). Einen Beweis von Satz 4.4.2 findet man in *Hirzebruch & Scharlau [1971], S. 62* oder in *Taylor & Lay [1980], S. 174*.

Es folgen noch vier Bemerkungen.

Bemerkung 4.4.3: *In $\mathcal{X} = \ell^1$ sind $\mathcal{T}_{\mathcal{X}}$ und $\mathcal{T}(\mathcal{X}, \mathcal{X}')$ verschieden.*

Beweisidee: $B(o, 1) \subset \mathcal{X}$ und $U(o, E, \varepsilon)$, E endlich, sind bezüglich $\mathcal{T}_{\mathcal{X}}$ offen, es ist aber $B(o, 1) \notin \mathcal{T}(\mathcal{X}, \mathcal{X}')$. Es gibt nämlich kein

$$U(o, E, \varepsilon) \subset B(o, 1). \tag{$*$}$$

Das zeigen wir indirekt. Es sei $(*)$ für ein endliches E und $\varepsilon > 0$ erfüllt. Weil E endlich ist, wählen wir etwa $E = \{e_1\}$. Es sei

$$x := \frac{1}{2}\Big(\varepsilon, 1, \frac{1}{2}, \frac{1}{4}, \ldots\Big) \in \ell^1.$$

Dann ist

$$x \in U(o, E, \varepsilon),$$

aber

$$\|x\| = \frac{\varepsilon}{2} + 1,$$

also $x \notin B(o, 1)$. □

Bemerkung 4.4.4: *Es seien $(\mathcal{X}, \mathcal{T}_\mathcal{X})$ ein topologischer Raum und*

$$f \;:\; \mathcal{X} \longrightarrow \mathbb{R}$$

eine Abbildung. Dann heißt f in $x \in \mathcal{X}$ halbstetig nach oben, *wenn*

$$\limsup_{x_n \to x} f(x_n) \leq f(x)$$

gilt, und halbstetig nach unten, *wenn*

$$f(x) \leq \liminf_{x_n \to x} f(x_n)$$

ist.

Hat f beide Eigenschaften, dann folgt

$$\limsup_{n\to\infty} f(x_n) \leq f(x) \leq \liminf_{n\to\infty} f(x_n),$$

also

$$\lim_{n\to\infty} f(x_n) = f(x),$$

und f ist in x stetig.

Es sei $\mathcal{X}$ ein normierter Raum. Dann gilt

Bemerkung 4.4.5: *Die Norm ist in der schwachen Topologie $\mathcal{T}(\mathcal{X}, \mathcal{X}')$ halbstetig nach unten.*

Das folgt aus Satz 4.0.6, nämlich

$$\|x\| \leq \liminf_{n\to\infty} \|x_n\|. \qquad \square$$

Natürlich ist die Norm in der schwachen Topologie nicht stetig, wenn die Topologien $\mathcal{T}_\mathcal{X}$ und $\mathcal{T}(\mathcal{X}, \mathcal{X}')$ verschieden sind. Es sei zum Beispiel $\{u_1, u_2, \ldots\}$ ein Orthonormalsystem in einem Hilbertraum. Dann konvergiert (u_n) schwach gegen o, es gilt aber

$$\limsup_{n\to\infty} \|u_n\| = 1 \not\leq \|o\| = 0.$$

Bemerkung 4.4.6: *Es seien $\mathcal{H}$ ein Hilbertraum und*

$$\begin{aligned} s \;:\; \mathcal{H} \times \mathcal{H} &\longrightarrow \mathbb{K}, \\ \langle f, g\rangle &\longmapsto (f, g). \end{aligned}$$

s ist partiell stetig in der schwachen Topologie, aber nicht bezüglich beider Variablen.

Beweis: Halten wir etwa die zweite Variable g fest und wählen eine Folge (f_n) mit

$$\text{w-}\lim_{n\to\infty} f_n = f,$$

dann folgt $s(f_n, g) \to s(f, g)$. Aus

$$\text{w-}\lim_{n\to\infty} f_n = f \quad \text{und} \quad \text{w-}\lim_{n\to\infty} g_n = g$$

folgt aber nicht $s(f_n, g_n) \to s(f, g)$. Man wähle nur die schwachen Nullfolgen $f_n = g_n = u_n$ mit einem Orthonormalsystem $\{u_1, u_2, \ldots\}$. $\square$

5 Abgeschlossene Operatoren

In diesem Kapitel wollen wir die Behandlung linearer Abbildungen fortsetzen. Für beschränkte lineare Abbildungen haben wir in den vorangegangenen Abschnitten bereits eine Reihe von Ergebnissen hergeleitet. So wollen wir uns jetzt auf unbeschränkte Abbildungen konzentrieren. Gerade solche Abbildungen treten in den Anwendungen häufig auf. Die Differentialoperatoren gehören hierzu. Denken Sie an

$$\begin{aligned} \mathcal{C}_1([0,1]) &\longrightarrow \mathcal{C}([0,1]), \\ x &\longmapsto x'. \end{aligned}$$

Es sei etwa $x_n(t) := t^n$. Dann ist

$$\|x_n\|_\infty = 1 \quad \text{und} \quad \|x_n'\|_\infty = n.$$

Mithin ist $D : \mathcal{C}_1([0,1]) \subset \mathcal{C}([0,1]) \to \mathcal{C}([0,1])$ ein unbeschränkter linearer Operator. Wählt man für $\mathcal{C}_1([0,1])$ die Norm

$$\|x\|_{\mathcal{C}_1} := \|x\|_\infty + \|x'\|_\infty,$$

dann ist natürlich

$$\|x'\|_\infty \leq \|x\|_{\mathcal{C}_1}$$

und $D : \mathcal{C}_1([0,1]) \to \mathcal{C}([0,1])$ deshalb eine beschränkte Abbildung mit $\|D\| \leq 1$.

Unser Beispiel zeigt bereits, daß man sehr sorgfältig darauf achten muß, von woher nach wohin der Operator abbildet und welche Norm gemeint ist. Es sei deshalb noch einmal an die zugrundeliegenden Begriffsbildungen erinnert. Die Räume $\mathcal{X}, \mathcal{Y}$ seien normiert und

$$A \;:\; \mathcal{D}(A) \subset \mathcal{X} \longrightarrow \mathcal{Y}$$

eine lineare Abbildung mit dem Definitionsbereich $\mathcal{D}(A) \subset \mathcal{X}$.

$$G(A) := \big\{ \langle x, Ax \rangle \;\big|\; x \in \mathcal{D}(A) \big\} \subset \mathcal{X} \times \mathcal{Y}$$

ist der *Graph* von A. Er ist also ein Unterraum von $\mathcal{X} \times \mathcal{Y}$. Eine Abbildung

$$B \;:\; \mathcal{D}(B) \subset \mathcal{X} \longrightarrow \mathcal{Y}$$

ist eine *Fortsetzung* von A, kurz $A \subset B$, wenn

$$G(A) \subset G(B)$$

gilt. Es ist dann also

$$\mathcal{D}(A) \subset \mathcal{D}(B) \quad \text{und} \quad B|\mathcal{D}(A) = A.$$

Folgendes sei noch bemerkt:

1. Ein Unterraum $M \subset \mathcal{X} \times \mathcal{Y}$ ist genau dann der Graph eines linearen Operators, wenn

$$\Big(\langle o, y \rangle \in M \Big) \implies \Big(y = o \Big)$$

gilt. Dann ist nämlich für alle $\langle x, y \rangle \in M$

$$Ax := y$$

wohldefiniert. Mit $\langle x, y \rangle$ ist auch $\langle \alpha x, \alpha y \rangle \in M$, und es ist

$$A(\alpha x) = \alpha y = \alpha A x.$$

A ist also ein linearer Operator.

2. Der Definitionsbereich $\mathcal{D}(A)$ einer linearen Abbildung A läßt sich durch

$$\begin{aligned} \|\cdot\|_A \;:\; \mathcal{D}(A) &\longrightarrow \mathbb{R}_0^+, \\ x &\longmapsto \|x\|_A := \sqrt{\|x\|^2 + \|Ax\|^2}. \end{aligned}$$

normieren. Es sei $G(A) \subset \mathcal{X} \times \mathcal{Y}$ der Graph der linearen Abbildung A. Auf $G(A)$ erklären wir die *Graphennorm* vermöge

$$\|\langle x, y\rangle\|_{G(A)} \;:=\; \sqrt{\|x\|^2 + \|y\|^2}.$$

Es sei

$$\begin{aligned} K \;:\; & G(A) \longrightarrow \mathcal{D}(A), \\ & \langle x, Ax\rangle \longmapsto x. \end{aligned}$$

Dann ist K eine Kongruenz zwischen $G(A)$ und $\{\mathcal{D}(A); \|\cdot\|_A\}$. Deshalb nennt man $\|\cdot\|_A$ auch Graphennorm von A.

Wir betrachten nun eine wichtige Klasse (unbeschränkt) linearer Abbildungen, nämlich die abgeschlossenen. Die Differentialoperatoren gehören in diese Klasse.

Definition 5.0.1: *Eine lineare Abbildung*

$$A \;:\; \mathcal{D}(A) \subset \mathcal{X} \longrightarrow \mathcal{Y}$$

heißt abgeschlossen, *wenn der Graph $G(A)$ in $\mathcal{X} \times \mathcal{Y}$ abgeschlossen ist.*

Wenn $G(A)$ abgeschlossen ist, besitzt jede konvergente Folge aus $G(A)$ einen Grenzwert in $G(A)$. Es gilt also

$$\Big(x_n \in \mathcal{D}(A), \quad x_n \to x \in \mathcal{X} \quad \text{und} \quad Ax_n \to y \in \mathcal{Y}\Big) \implies \Big(x \in \mathcal{D}(A) \quad \text{und} \quad y = Ax\Big). \qquad (*)$$

Umgekehrt impliziert $(*)$ die Abgeschlossenheit von $G(A)$.

Lemma 5.0.2: *A sei abgeschlossen und injektiv. Dann ist auch*

$$A^{-1} \;:\; \mathcal{R}(A) \longrightarrow \mathcal{D}(A)$$

abgeschlossen.

Beweis:

$$\begin{aligned} V \;:\; & \mathcal{X} \times \mathcal{Y} \longrightarrow \mathcal{Y} \times \mathcal{X}, \\ & \langle x, y = Ax\rangle \longmapsto \langle y, x = A^{-1}y\rangle \end{aligned}$$

ist eine Kongruenz. □

Lemma 5.0.3: *Es seien $A \in \mathcal{CL}(\mathcal{D}, \mathcal{Y})$. Dann ist A abgeschlossen, wenn $\mathcal{D} \subset \mathcal{X}$ abgeschlossen ist. Ist $\mathcal{Y}$ ein Banachraum, dann gilt auch die Umkehrung.*

Beweis:
$\Rightarrow$: Es sei $\mathcal{D}$ abgeschlossen. Aus $x_n \in \mathcal{D}$, $x_n \to x \in \mathcal{X}$ und $Ax_n \to y \in \mathcal{Y}$ folgt dann $x \in \mathcal{D}$ und

$$\|y - Ax\| \le \|y - Ax_n\| + \|A(x_n - x)\| \to 0.$$

$\Leftarrow$: Aus $x_n \to x$ folgt die Existenz eines $y \in \mathcal{Y}$ mit $Ax_n \to y$. Dann ist $Ax = y$ und damit $x \in \mathcal{D}$. □

Wir fügen einige Bemerkungen an:

1. Es seien $\mathcal{X} = \mathcal{Y}$ und $\mathcal{D} \subset \mathcal{X}$ nicht abgeschlossen. Dann ist

$$id \;:\; \mathcal{D} \subset \mathcal{X} \longrightarrow \mathcal{X}$$

stetig, aber nicht abgeschlossen. Natürlich fragt man, ob ein Operator A in einem solchen Falle *abschließbar* ist, ob es also eine abgeschlossene Fortsetzung von A gibt. Auf diese wichtige Frage werden wir in §5.2 näher eingehen.

2. Wir geben ein Beispiel für einen unstetigen abgeschlossenen Operator: Es seien $\mathcal{X} = \mathcal{Y} = \ell^1$, $x = (\xi_1, \xi_2, \ldots)$,

$$\mathcal{D}(A) := \Big\{x \in \mathcal{X} \;\Big|\; \sum_{i=1}^{\infty} i\,|\xi_i| < \infty\Big\}$$

und

$$A \;:\; \mathcal{D}(A) \subset \mathcal{X} \longrightarrow \mathcal{X},$$
$$x \longmapsto (\xi_1, 2\xi_2, 3\xi_3, \ldots).$$

Wegen $\|Ae_n\| = n$ ist A unstetig. Es gilt aber

$$\|x\| \leq \|A\,x\|.$$

Man sagt dafür, A sei *koerzitiv*. A ist dann injektiv. A ist auch surjektiv. Mit $y = (\eta_1, \eta_2, \ldots) \in \mathcal{X}$ ist nämlich

$$A^{-1}y := \left(\eta_1, \frac{\eta_2}{2}, \frac{\eta_3}{3}, \ldots\right) \in \mathcal{D}(A).$$

Daraus folgt die Abgeschlossenheit von A, denn es seien

$$x_n \in \mathcal{D}(\mathcal{A}), \quad x_n \to x \in \mathcal{X} \quad \text{und} \quad A\,x_n \to y \in \mathcal{X}.$$

Dann folgt mit $z := A^{-1}y \in \mathcal{D}(A)$ aus der Koerzitivität

$$\|x_n - z\| \leq \|A(x_n - z)\| = \|Ax_n - y\| \to 0,$$

also $x = z \in \mathcal{D}(A)$ und $Ax = y$.

3. Es seien $\mathcal{X} = \mathcal{Y} = \mathcal{C}([0,1])$ und

$$A \;:\; \mathcal{X} \longrightarrow \mathcal{X}$$

mit

$$(Ax)(t) \;:=\; \int_0^t x(s)\,ds.$$

A ist stetig, folglich abgeschlossen. Es ist aber

$$\mathcal{R}(A) = \left\{ y \in \mathcal{C}_1([0,1]) \;\middle|\; y(0) = 0 \right\} \subset \mathcal{X}$$

nicht abgeschlossen in $\mathcal{X}$. Weil A injektiv ist, existiert

$$A^{-1} \;:\; \mathcal{R}(A) \subset \mathcal{X} \longrightarrow \mathcal{X},$$
$$y \longrightarrow y'.$$

A^{-1} ist abgeschlossen, aber unbeschränkt.

5.1 Der Satz vom abgeschlossenen Graphen

Wir wollen nun aus dem Satz von der inversen Abbildung in §3.3 eine wichtige Folgerung für abgeschlossene Operatoren herleiten. Vorher zeigen wir noch zwei Hilfssätze.

Lemma 5.1.1: *Es seien $\mathcal{X}$ ein Banachraum und $\mathcal{Y}$ ein normierter Raum. Die lineare Abbildung*

$$A \;:\; \mathcal{D}(A) \subset \mathcal{X} \longrightarrow \mathcal{Y}$$

besitze eine stetige Inverse

$$A^{-1} \;:\; \mathcal{R}(A) \subset \mathcal{Y} \longrightarrow \mathcal{X}.$$

Dann gilt

$$A \text{ ist abgeschlossen} \iff \mathcal{R}(A) \subset \mathcal{Y} \text{ ist abgeschlossen.}$$

Beweis:

$\Rightarrow$: Es sei $y_n = Ax_n \to y \in \mathcal{Y}$. Dann ist $x_n = A^{-1}y_n \in \mathcal{D}(A)$. Wegen der Beschränktheit von A^{-1} ist (x_n) eine Cauchyfolge, es gibt also ein $x \in \mathcal{X}$ mit $x_n \to x$. Weil A abgeschlossen ist, folgt $x \in \mathcal{D}(A)$ und $Ax = y$, also $y \in \mathcal{R}(A)$.

$\Leftarrow$: Es seien $x_n \in \mathcal{D}(A)$, $x_n \to x \in \mathcal{X}$ und $Ax_n \to y \in \mathcal{Y}$. Dann ist $y \in \mathcal{R}(A)$, weil $\mathcal{R}(A)$ abgeschlossen ist. Mithin gibt es ein $\hat{x} \in \mathcal{D}(A)$ mit $A\hat{x} = y$, und es folgt

$$\|x - \hat{x}\| \leq \|x - x_n\| + \|x_n - \hat{x}\| \leq \|x - x_n\| + \|A^{-1}(Ax_n - y)\| \to 0.$$

Es ist also $x = \hat{x} \in \mathcal{D}(A)$ und $Ax = A\hat{x} = y$. A ist deshalb abgeschlossen. □

Lemma 5.1.2: *Es seien $\mathcal{X}, \mathcal{Y}$ Banachräume und*

$$A \;:\; \mathcal{D}(A) \subset \mathcal{X} \longrightarrow \mathcal{Y}$$

eine lineare Abbildung. Dann gilt

$$A \text{ ist abgeschlossen} \iff \{\mathcal{D}(A);\, \|\cdot\|_A\} \text{ ist vollständig.}$$

Beweis: Zu Beginn dieses Kapitels haben wir bereits die Kongruenz

$$G(A) \cong \left\{\mathcal{D}(A);\, \|\cdot\|_A\right\}$$

bemerkt. Ist $G(A) \subset \mathcal{X} \times \mathcal{Y}$ abgeschlossen, dann ist $G(A)$ vollständig und umgekehrt. □

Nun formulieren wir unser Hauptresultat.

Satz vom abgeschlossenen Graphen: *Es seien $\mathcal{X}, \mathcal{Y}$ Banachräume und $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$. Dann gilt*

$$A \text{ ist abgeschlossen} \iff A \text{ ist stetig.}$$

Beweis:

$\Leftarrow$: Das wurde in Lemma 5.0.3 gezeigt.

$\Rightarrow$: Als abgeschlossener Teilraum von $\mathcal{X} \times \mathcal{Y}$ ist $G(A)$ ein Banachraum. Es seien

$$\begin{aligned} P_{\mathcal{X}} &: G(A) \longrightarrow \mathcal{X}, \quad \langle x, y\rangle \longmapsto x, \\ P_{\mathcal{Y}} &: G(A) \longrightarrow \mathcal{Y}, \quad \langle x, y\rangle \longmapsto y. \end{aligned}$$

$P_{\mathcal{X}}$ und $P_{\mathcal{Y}}$ sind stetige lineare Abbildungen, beispielsweise wegen

$$\|P_{\mathcal{X}}\langle x, y\rangle\|_{\mathcal{X}} = \|x\|_{\mathcal{X}} \leq \sqrt{\|x\|_{\mathcal{X}}^2 + \|y\|_{\mathcal{Y}}^2} = \|\langle x, y\rangle\|.$$

$P_{\mathcal{X}}$ ist eine bijektive Abbildung. Aus dem Satz von der inversen Abbildung in §3.3 folgt deshalb

$$P_{\mathcal{X}}^{-1} \in \mathcal{CL}\big(\mathcal{X}, G(A)\big).$$

Dann ist auch

$$A = P_{\mathcal{Y}} \circ P_{\mathcal{X}}^{-1}$$

stetig. □

5.2 Abschließbare Abbildungen

Nun wollen wir uns der wichtigen Frage zuwenden, ob eine lineare Abbildung zu einer abgeschlossenen fortgesetzt werden kann.

Definition 5.2.1: *Eine lineare Abbildung*

$$A \;:\; \mathcal{D}(A) \subset \mathcal{X} \longrightarrow \mathcal{Y}$$

heißt abschließbar, *wenn $\overline{G(A)}$ der Graph einer linearen Abbildung $\bar{A}$ ist. Diese (eindeutig bestimmte) Abbildung $\bar{A}$ heißt dann* Abschluß *von A.*

Es seien

$$\begin{aligned} A &: \mathcal{D}(A) \subset \mathcal{X} \longrightarrow \mathcal{Y} \text{ abgeschlossen}, \\ B &: \mathcal{D}(B) \subset \mathcal{X} \longrightarrow \mathcal{Y} \text{ abschließbar} \end{aligned}$$

mit $\bar{B} = A$. Dann heißt $\mathcal{D}(B)$ Definitionskern *(engl. core) von A.*

In den Anwendungen wählt man als Definitonskern für Differentialoperatoren oft möglichst glatte Funktionen, zum Beispiel die Testfunktionen.

Bemerkung 5.2.2:

(1) $\overline{G(A)}$ *ist genau dann Graph einer linearen Abbildung, wenn*

$$\langle o, y\rangle \in \overline{G(A)} \implies y = o$$

gilt. Das heißt, aus $x_n \in \mathcal{D}(A)$, $x_n \to o$ *und* $Ax_n \to y$ *soll* $y = o$ *folgen.*

(2) Es sei A *abschließbar. Dann ist* $A \subset \bar{A}$, *und in Banachräumen* $\mathcal{X}, \mathcal{Y}$ *gilt*

$$\mathcal{D}(\bar{A}) \cong \left\{\mathcal{D}(A);\ \|\cdot\|_A\right\}^{\sim}.$$

Ferner ist

$$\mathcal{D}(A) \subset \mathcal{D}(\bar{A}) \subset \overline{\mathcal{D}(A)} \quad \textit{und} \quad \mathcal{R}(A) \subset \mathcal{R}(\bar{A}) \subset \overline{\mathcal{R}(A)}.$$

(3) Aus $A \subset B$ *und* $B = \bar{B}$ *folgt die Existenz von* $\bar{A}$ *mit* $\bar{A} \subset B$.

(4) Eine stetige Abbildung ist abschließbar.

(5) Es seien $\mathcal{X}$ *ein normierter Raum und*

$$A \ : \ \mathcal{D} \subset \mathcal{X} \longrightarrow \mathbb{K}^{\mathrm{n}}$$

eine lineare Abbildung. Dann gilt

$$A \textit{ ist abschließbar} \iff A \textit{ ist stetig.}$$

Es folgen noch einige Beispiele:

1. Es seien $\mathcal{X} = \mathcal{Y} = \ell^1$,

$$\mathcal{D}(B) := \left\{x = (\xi_1, \xi_2, \ldots) \in \mathcal{X} \ \middle|\ \exists\, n(x) \in \mathbb{N} \quad \forall\, i > n \quad \xi_i = 0\right\}$$

und

$$\begin{aligned} B \ : \ \mathcal{D}(B) \subset \mathcal{X} &\longrightarrow \mathcal{X}, \\ x &\longmapsto (\xi_1, 2\xi_2, 3\xi_2, \ldots). \end{aligned}$$

Man vergleiche das analoge Beispiel in der zweiten Bemerkung im Anschluß an Lemma 5.0.3. Dort war

$$\mathcal{D}(A) := \left\{x \in \mathcal{X} \ \middle|\ \sum_{i=1}^{\infty} i\,|\xi_i| < \infty\right\}$$

mit derselben Abbildungsvorschrift. Offenbar ist $B \subset A = \bar{A}$. Mithin ist B abschließbar. $\mathcal{D}(B)$ ist in $\mathcal{D}(A)$ bzgl. der $\|\cdot\|_A$-Norm dicht. Deshalb ist $\bar{B} = A$, und $\mathcal{D}(B)$ ist ein Definitionskern von A.

2. Es seien $\mathcal{X} = \ell^2$, $\mathcal{D} := \operatorname{span}(e_1, e_2, \ldots)$ und

$$A \ : \ \mathcal{D} \subset \mathcal{X} \longrightarrow \mathbb{R}$$

vermöge $Ae_n := n$. Dann ist A nicht abschließbar, denn andernfalls wäre A nach Bemerkung 5.2.2.5 stetig.

3. Es seien $I = (-1, 1)$, $\mathcal{X} = \mathcal{L}^1(I)$, $\mathcal{Y} = \mathbb{R}$, $\mathcal{D}(A) := \mathcal{C}(I) \cap \mathcal{L}^1(I)$ und

$$Ax := x(0).$$

Es seien ferner

$$x_n(t) := \begin{cases} 0 & \text{für } -1 \le t < -\frac{1}{n} \\ n(t + \frac{1}{n}) & \text{für } -\frac{1}{n} \le t < 0 \\ -n(t - \frac{1}{n}) & \text{für } 0 \le t < \frac{1}{n} \\ 0 & \text{für } \frac{1}{n} \le t \le 1. \end{cases}$$

Dann ist $Ax_n = 1$ und

$$\|x_n\| = \int_{-1}^{1} |x_n(t)|\, dt = \frac{1}{n} \to 0.$$

Mithin ist A nicht abschließbar.

4. Es seien $\mathcal{X} = \mathcal{L}^\infty(\mathbb{R})$, $\mathcal{Y} = \mathbb{R}$, $\mathcal{D}(A) := \mathcal{L}^1(\mathbb{R}) \cap \mathcal{L}^\infty(\mathbb{R})$ und

$$Ax := \int_{-\infty}^{\infty} x(t)\,dt.$$

Es seien ferner

$$x_n(t) := \begin{cases} 0 & \text{für } -\infty < t < -n \\ \frac{1}{n} & \text{für } -n \le t \le n \\ 0 & \text{für } n < t < \infty. \end{cases}$$

Dann ist $Ax_n = 2$ und

$$\|x_n\|_\infty = \frac{1}{n} \to 0.$$

Mithin ist A nicht abschließbar.

5. Es seien $I = (0,1)$, $\mathcal{X} = \mathcal{Y} = \mathcal{L}^2(I)$,

$$\mathcal{W}_1(I) := \left\{ u \in \mathcal{X} \;\middle|\; \exists\, f \in \mathcal{X} \quad \forall\, \varphi \in \mathring{\mathcal{C}}_\infty(I) \quad (u, \varphi') = -(f, \varphi) \right\}$$

und

$$\begin{aligned} W \;:\; \mathcal{W}_1(I) \subset \mathcal{X} &\longrightarrow \mathcal{X}, \\ u &\longmapsto Wu := u' := f. \end{aligned}$$

Die Abbildung W ist wohldefiniert; man nennt Wu *schwache Ableitung* von u. Offenbar stimmen für $u \in \mathcal{C}_1(I)$ die übliche und die schwache Ableitung überein. W ist abgeschlossen: Es seien nämlich $u_n \in \mathcal{W}_1(I)$, $u_n \to u$ und $u_n' \to y$. Dann folgt für alle $\varphi \in \mathring{\mathcal{C}}_\infty(I)$

$$(u, \varphi') \leftarrow (u_n, \varphi') = -(u_n', \varphi) \to -(y, \varphi),$$

also $u \in \mathcal{W}_1(I)$ und $Wu = y$. Mithin ist $\mathcal{W}_1(I)$ bezüglich

$$\|u\|_{\mathcal{W}_1} := \|u\|_W = \sqrt{\|u\|_{\mathcal{X}}^2 + \|u'\|_{\mathcal{X}}^2}$$

vollständig.

Für ein festes k mit $1 \le k \le \infty$ sei nun

$$\mathcal{S} := \mathcal{C}_k(I) \cap \mathcal{W}_1(I)$$

und

$$\begin{aligned} S \;:\; \mathcal{S} \subset \mathcal{X} &\longrightarrow \mathcal{X}, \\ u &\longmapsto Su := u'. \end{aligned}$$

Offenbar ist $S \subset W = \overline{W}$. Damit ist S abschließbar. Es gilt $\bar{S} \subset W$, und

$$\bar{\mathcal{S}} \subset \mathcal{W}_1(I)$$

ist der Abschluß von $\mathcal{S}$ in $\mathcal{W}_1(I)$ bezüglich der $\|\cdot\|_W$-Norm. Ein Dichteargument zeigt dann schließlich $\bar{S} = W$. Die Menge $\mathcal{S}$ ist also Definitionskern von W; insbesondere ist es $\mathcal{C}_\infty(I) \cap \mathcal{W}_1(I)$.

Neben den schwachen definiert man auch *starke Ableitungen*. Während die ersteren durch *Testen* gewonnen werden, erhält man die letzteren durch *Approximieren*. Es sei nämlich

$$\mathcal{H}_1(I) := \left\{ \mathcal{C}_1(I) \cap \mathcal{W}_1(I);\ \|\cdot\|_W \right\}^\sim \subset \mathcal{W}_1(I).$$

Zu $v \in \mathcal{H}$ existiert dann eine Folge (v_n), $v_n \in \mathcal{C}_1(I) \cap \mathcal{W}_1(I)$, mit

$$\|v_n - v\|_W \to 0.$$

Es ist

$$v' := \mathcal{L}^2(I)\text{-}\lim_{n\to\infty} v_n'$$

die starke Ableitung von v, und es folgt $\mathcal{H}_1(I) = \mathcal{W}_1(I)$, man vergleiche die Einführung oder §6.6.

5.3 Adjungierte Abbildungen

Die zu $A \in \mathcal{CL}(\mathcal{X}, \mathcal{Y})$ *konjugierte Abbildung* A' haben wir bereits in §1.5.2 als

$$\begin{aligned} A' \;:\; & \mathcal{Y}' \longrightarrow \mathcal{X}', \\ & y' \longmapsto y' \circ A \end{aligned}$$

definiert. Ist nun

$$A \;:\; \mathcal{D}(A) \subset \mathcal{X} \longrightarrow \mathcal{Y}$$

eine lineare Abbildung mit dichtem Definitionsbereich, dann kann man analog vorgehen. Es ist $y' \circ A \in \mathcal{L}(\mathcal{D}(A), \mathbb{K})$. Falls für ein $y' \in \mathcal{Y}'$

$$y' \circ A \in \mathcal{CL}(\mathcal{D}(A), \mathbb{K})$$

ist, dann kann man $y' \circ A$ zu $y' \circ A \in \mathcal{CL}(\mathcal{X}, \mathbb{K})$ stetig ergänzen. Man nimmt deshalb

$$\mathcal{D}(A') := \left\{ y' \in \mathcal{Y}' \;\middle|\; y' \circ A \in \mathcal{CL}(\mathcal{D}(A), \mathbb{K}) \right\}$$

und nennt

$$\begin{aligned} A' \;:\; & \mathcal{D}(A') \subset \mathcal{Y}' \longrightarrow \mathcal{X}', \\ & y' \longmapsto y' \circ A \end{aligned}$$

die zu $A \in \mathcal{L}(\mathcal{D}(A), \mathcal{Y})$ *konjugierte Abbildung*. Dabei wird also $\overline{\mathcal{D}(A)} = \mathcal{X}$ vorausgesetzt.

Der Definitionsbereich von A' kann sehr klein sein, dafür geben wir weiter unten ein Beispiel. Es gilt aber für dicht definierte A

Satz 5.3.1: *Es sei A dicht definiert. Dann ist A' abgeschlossen.*

Beweis: Es sei $y'_n \in \mathcal{D}(A')$ mit $y'_n \to y' \in \mathcal{Y}'$ und $A'y'_n \to x' \in \mathcal{X}'$. Dann folgt

$$\forall\, x \in \mathcal{D}(A) \qquad y'_n(Ax) \to y'(Ax) \quad \text{und} \quad y'_n(Ax) = (A'y'_n)x \to x'x,$$

also $y' \circ A = x' \in \mathcal{CL}(\mathcal{X}, \mathbb{K})$. Das bedeutet $y' \in \mathcal{D}(A')$ und $A'y' = y' \circ A = x'$. □

Wenn A selbst abgeschlossen ist, kann man mehr sagen, nämlich

Satz 5.3.2: *Es sei A dicht definiert und abgeschlossen. Dann ist $\mathcal{D}(A')$ eine* totale Teilmenge *von $\mathcal{Y}'$.*

Dabei heißt eine Teilmenge $F \subset \mathcal{X}'$ *total*, wenn folgendes gilt:

$$\Big(\forall\, x' \in F \quad x'x = 0 \Big) \implies x = o.$$

Beweis: Wir haben zu zeigen, daß es zu vorgegebenem $y_0 \neq o$ ein $y' \in \mathcal{D}(A')$ mit $y'y_0 \neq 0$ gibt. Es sei $G(A)$ der Graph von A. Dann ist $\langle o, y_0 \rangle \notin G(A)$. Weil $G(A)$ abgeschlossen ist, folgt aus dem Trennungssatz die Existenz eines $z' \in (\mathcal{X} \times \mathcal{Y})'$ mit

$$z'\langle o, y_0 \rangle \neq 0 \quad \text{und} \quad \forall\, x \in \mathcal{D}(A) \quad z'\langle x, Ax \rangle = 0.$$

Es werde $y' \in \mathcal{Y}'$ durch $y'y := z'\langle o, y \rangle$ und $x' \in \mathcal{X}'$ durch $x'x := z'\langle x, o \rangle$ definiert. Dann ist $y'y_0 \neq 0$ und

$$\forall\, x \in \mathcal{D}(A) \qquad 0 = z'\langle x, Ax \rangle = x'x + y'Ax,$$

also $y' \in \mathcal{D}(A')$. □

Die Behauptung in Satz 5.3.2 formuliert man auch als: $D(A')$ ist *schwach-stern dicht* in $\mathcal{Y}'$. Im allgemeinen ist $\mathcal{D}(A')$ aber nicht in $\mathcal{Y}'$ dicht, auch wenn A abgeschlossen ist. Für abgeschlossene A und reflexive $\mathcal{Y}$ folgt jedoch $\overline{\mathcal{D}(A')} = \mathcal{Y}'$.

Satz 5.3.3: *Es gilt $\mathcal{D}(A') = \mathcal{Y}'$ genau dann, wenn A stetig ist. In diesem Falle ist A' ebenfalls stetig, und es gilt $\|A'\| = \|A\|$.*

Beweis: Die zweite Aussage ist in Satz 1.5.4 enthalten. Es sei also $\mathcal{D}(A') = \mathcal{Y}'$. Dann gilt mit $S := \{x \in \mathcal{D}(A) \mid \|x\| = 1\}$

$$\forall\, y' \in \mathcal{Y}' \qquad \sup_{x \in S} |y'(Ax)| \le \|A'y'\| \le c(y') < \infty.$$

Aus dem Prinzip von der gleichmäßigen Beschränktheit (Folgerung 3.2.2) erhalten wir daher

$$\sup_{x \in S} \|Ax\| < \infty,$$

also die Beschränktheit von A. □

Im folgenden seien nun $\mathcal{X}, \mathcal{Y}$ Hilberträume. Dann steht das Skalarprodukt zur Verfügung, und vermöge der Rieszabbildung $R : \mathcal{X} \to \mathcal{X}'$ sind $\mathcal{X}$ und $\mathcal{X}'$ konjugiert kongruent. Das kann man ausnutzen, um anstelle der konjugierten Abbildung $A' : \mathcal{D}(A') \subset \mathcal{Y}' \to \mathcal{X}'$ die *adjungierte Abbildung*

$$A^* \;:\; \mathcal{D}(A^*) \subset \mathcal{Y} \longrightarrow \mathcal{X}$$

einzuführen. Dadurch werden manche Rechnungen einfacher.

Erinnert sei zunächst an die konjugiert lineare Rieszabbildung

$$R \;:\; \mathcal{X} \to \mathcal{X}',$$

mit

$$\forall\, z \in \mathcal{X} \qquad (Rx)z := (z, x),$$

und an das folgende Diagramm

$$\begin{array}{ccc} \mathcal{X} & \underset{A^*}{\overset{A}{\rightleftarrows}} & \mathcal{Y} \\ {\scriptstyle R_{\mathcal{X}}}\downarrow & & \downarrow{\scriptstyle R_{\mathcal{Y}}} \\ \mathcal{X}' & \xleftarrow{\;A'\;} & \mathcal{Y}' \end{array}.$$

Es sei also $A : \mathcal{D}(A) \subset \mathcal{X} \to \mathcal{Y}$ eine lineare Abbildung mit dichtem Definitionsbereich. Dann definieren wir die zu A *adjungierte Abbildung* A^* als

$$A^* := R_{\mathcal{X}}^{-1} A' R_{\mathcal{Y}}.$$

Offenbar ist A^* linear, und es gilt für alle $x \in \mathcal{D}(A)$ und $y \in \mathcal{D}(A^*)$

$$(Ax, y)_{\mathcal{Y}} = (R_{\mathcal{Y}} y)Ax = (A' R_{\mathcal{Y}} y)x = (x, R_{\mathcal{X}}^{-1} A' R_{\mathcal{Y}} y)_{\mathcal{X}} = (x, A^* y)_{\mathcal{X}}$$

sowie

$$\begin{aligned} \mathcal{D}(A^*) &= \{y \in \mathcal{Y} \mid R_{\mathcal{Y}} y \in \mathcal{D}(A')\} \\ &= \{y \in \mathcal{Y} \mid (R_{\mathcal{Y}} y)A \in \mathcal{CL}(\mathcal{D}(A), \mathbb{K})\} \\ &= \{y \in \mathcal{Y} \mid \exists\, z \in \mathcal{X} \quad \forall\, x \in \mathcal{D}(A) \quad (x, z)_{\mathcal{X}} = (R_{\mathcal{Y}} y)Ax = (Ax, y)_{\mathcal{Y}}\}. \end{aligned}$$

Die letzte Gleichheit folgt aus dem Rieszschen Darstellungssatz in §1.5.3.

Die zu A Adjungierte A^* ist nach Satz 5.3.1 abgeschlossen. Für beschränkte Abbildungen zeigt man leicht

1. Es seien $A, B \in \mathcal{CL}(\mathcal{X}, \mathcal{Y})$ und $\alpha \in \mathbb{K}$. Dann ist

$$(A + B)^* = A^* + B^* \quad \text{und} \quad (\alpha A)^* = \bar{\alpha} A^*.$$

2. Es seien $A \in \mathcal{CL}(\mathcal{Y}, \mathcal{Z})$ und $B \in \mathcal{CL}(\mathcal{X}, \mathcal{Y})$. Dann ist

$$(AB)^* = B^* A^*.$$

3. Es sei $A \in \mathcal{CL}(\mathcal{X}, \mathcal{Y})$. Dann ist

$$A^{**} := (A^*)^* = A.$$

4. Es sei $A \in \mathcal{CL}(\mathcal{X}, \mathcal{Y})$. Dann ist $\|A^*\| = \|A\|$.

Die letzte Aussage folgt aus Satz 5.3.3. Ferner gilt

Lemma 5.3.4: *Es sei $A \in \mathcal{CL}(\mathcal{X}, \mathcal{Y})$. Dann ist*

$$\|A^*A\| = \|AA^*\| = \|A^*\| \cdot \|A\| = \|A\|^2.$$

Beweis: Aus

$$\|Ax\|^2 = (Ax, Ax) = (x, A^*Ax) \leq \|A^*A\| \cdot \|x\|^2$$

folgt

$$\|A\|^2 \leq \|A^*A\| \leq \|A^*\| \cdot \|A\| \leq \|A\|^2.$$

Also ist $\|A^*A\| = \|A^*\| \cdot \|A\|$. Ersetzt man nun A durch A^*, dann folgt die Behauptung. □

Algebraische Regeln für unbeschränkte Abbildungen sind etwas schwieriger zu beweisen, weil man jeweils auf den Definitionsbereich zu achten hat. Für dicht definierte abgeschlossene Abbildungen A, B in $\mathcal{X}$ nach $\mathcal{Y}$ kann man jedoch relativ leicht zeigen:

5. Wenn $\mathcal{D}(A+B) := \mathcal{D}(A) \cap \mathcal{D}(B)$ in $\mathcal{X}$ dicht ist, dann ist $(A+B)^*$ eine Fortsetzung von $A^* + B^*$.
6. Mit $\alpha \in \mathbb{K}$ ist

$$(\alpha A)^* = \bar{\alpha} A^*.$$

7. Es sei $\mathcal{X} = \mathcal{Y}$ und $\alpha \in \mathbb{K}$. Dann ist

$$(\alpha\, id + A)^* = \bar{\alpha}\, id + A^*.$$

8. Aus $A \subset B$ folgt $B^* \subset A^*$.
9. Es ist

$$\mathcal{R}(A)^\perp = \mathcal{N}(A^*).$$

Auch $\mathcal{D}(A^*)$ kann sehr klein sein: Es sei etwa $A : \mathcal{D} \subset \ell^2 \to \mathbb{R}$ wie im zweiten Beispiel in §5.2. Dann ist für $y \in \mathcal{D}(A^*)$

$$n\, y = (Ae_n, y)_{\mathbb{R}} = (e_n, A^*y)_{\ell^2}.$$

Mit $z := A^*y \in \ell^2$ folgt deshalb

$$z = y\,(1, 2, 3, \ldots)$$

oder $y = 0$. Es ist also $\mathcal{D}(A^*) = \{o\}$.

An solchen Fällen sind wir natürlich weniger interessiert, der soeben betrachtete Operator ist aber nicht abschließbar. Um weitere Eigenschaften der Adjungierten bequem nachweisen zu können, stellen wir zunächst ein etwas technisches Lemma bereit. Es seien $\mathcal{X}_1$, $\mathcal{X}_2$ zwei Hilberträume. Kanonisch wählen wir für $\mathcal{X}_1 \times \mathcal{X}_2$ das Skalarprodukt

$$\big(\langle x_1, x_2\rangle, \langle y_1, y_2\rangle\big)_{\mathcal{X}_1 \times \mathcal{X}_2} := \big(x_1, y_1\big)_{\mathcal{X}_1} + \big(x_2, y_2\big)_{\mathcal{X}_2}.$$

Es sei

$$\begin{aligned} V \;:\; & \mathcal{X}_1 \times \mathcal{X}_2 \longrightarrow \mathcal{X}_2 \times \mathcal{X}_1, \\ & \langle x_1, x_2\rangle \longmapsto \langle x_2, x_1\rangle. \end{aligned}$$

V ist eine lineare bijektive Abbildung. Es sei $M \subset \mathcal{X}_1 \times \mathcal{X}_2$ ein Unterraum. Dann ist

$$(VM)^\perp \;=\; V\left(M^\perp\right).$$

Es sei ferner

$$A \;:\; \mathcal{D}(A) \subset \mathcal{X}_1 \longrightarrow \mathcal{X}_2$$

eine injektive lineare Abbildung. Dann ist

$$G(A^{-1}) = V\, G(A),$$

und es gilt

Lemma 5.3.5: *Es seien*

$$A \;:\; \mathcal{D}(A) \subset \mathcal{X}_1 \longrightarrow \mathcal{X}_2$$

eine lineare Abbildung und V wie soeben definiert. Dann ist

$$\big(VG(A)\big)^{\perp} \subset \mathcal{X}_2 \times \mathcal{X}_1$$

genau dann der Graph eines linearen Operators, wenn $\mathcal{D}(A)$ in $\mathcal{X}_1$ dicht ist. In diesem Falle gilt

$$\big(VG(A)\big)^{\perp} = G(-A^*).$$

Beweis: Es ist $VG(A) = \langle Ax, x\rangle$, und es sei

$$\big(VG(A)\big)^{\perp} =: \langle y, y^*\rangle.$$

$\Rightarrow$: Es sei $y = 0$. Dann ist

$$\forall\, x \in \mathcal{D}(A) \qquad (x, y^*)_{\mathcal{X}_1} = 0,$$

und es soll $y^* = 0$ folgen. Dazu muß $\mathcal{D}(A)$ in $\mathcal{X}_1$ dicht sein.

$\Leftarrow$: Es ist

$$\forall\, x \in \mathcal{D}(A) \qquad 0 = (Ax, y)_{\mathcal{X}_2} + (x, y^*)_{\mathcal{X}_1},$$

also $y \in \mathcal{D}(A^*)$ und $y^* = -A^* y$. □

Wir zeigen nun $(A^*)^{-1} = (A^{-1})^*$, wenn diese Abbildungen existieren, nämlich

Satz 5.3.6: *Es seien $\mathcal{X}, \mathcal{Y}$ Hilberträume und*

$$A \;:\; \mathcal{D}(A) \subset \mathcal{X} \longrightarrow \mathcal{Y}$$

eine dicht definierte, lineare und injektive Abbildung mit in $\mathcal{Y}$ dichtem Wertebereich $\mathcal{R}(A)$. Dann gilt:

$$A^* \;:\; \mathcal{D}(A^*) \subset \mathcal{Y} \longrightarrow \mathcal{X}$$

ist injektiv und mit $A^{-1} : \mathcal{R}(A) \subset \mathcal{Y} \longrightarrow \mathcal{X}$

$$(A^{-1})^* = (A^*)^{-1}.$$

Beweis: Es ist

$$\mathcal{Y} = \overline{\mathcal{R}(A)} \oplus \mathcal{N}(A^*).$$

Wegen $\overline{\mathcal{R}(A)} = \mathcal{Y}$ folgt daraus die Injektivität von $\mathcal{A}^*$. Aus Lemma 5.3.5 erhalten wir daher

$$\begin{aligned} G\left((A^{-1})^*\right) = \big(V^{-1}G(-A^{-1})\big)^{\perp} = G(-A)^{\perp} = V^{-1}V\left(G(-A)^{\perp}\right) = \\ = V^{-1}\big(VG(-A)\big)^{\perp} = V^{-1}G(A^*) = G\left((A^*)^{-1}\right). \end{aligned}$$

Es gilt auch □

Lemma 5.3.7: *Es sei*

$$A \;:\; \mathcal{D}(A) \subset \mathcal{X} \longrightarrow \mathcal{Y}$$

eine dicht definierte lineare Abbildung. A^* sei ebenfalls dicht definiert. Dann ist

$$A \subset (A^*)^* =: A^{**}.$$

Beweis: Nach Voraussetzung existiert A^{**}. Mit $u \in \mathcal{D}(A)$ und $v \in \mathcal{D}(A^*)$ folgt

$$(Au, v) = (u, A^* v) = (A^{**}u, v),$$

es ist also $A \subset A^{**}$. □

Wir verschärfen das letzte Resultat.

Satz 5.3.8: *Es seien $\mathcal{X}, \mathcal{Y}$ Hilberträume und*

$$A \;:\; \mathcal{D}(A) \subset \mathcal{X} \longrightarrow \mathcal{Y}$$

eine dicht definierte lineare Abbildung. Dann gilt

$$\Big(A \text{ ist abschließbar}\Big) \iff \Big(\overline{\mathcal{D}(A^*)} = \mathcal{Y}\Big),$$

und es ist dann $\bar{A} = A^{**}$.

Beweis:

$\Rightarrow$: Aus Lemma 5.3.5 folgt

$$G(A^*) = \big(VG(-A)\big)^\perp = V\big(G(-A)\big)^\perp,$$

also

$$V^{-1}G(A^*) = G(-A)^\perp,$$

und

$$\big(V^{-1}G(A^*)\big)^\perp = G(-A)^{\perp\perp} = \overline{G(-A)} = G(-\bar{A})$$

ist der Graph des linearen Operators $-\bar{A}$. Mithin ist – wiederum nach Lemma 5.3.5 – die Abbildung A^* dicht definiert, und es gilt

$$G(\bar{A}) = \big(V^{-1}G(-A^*)\big)^\perp = G(A^{**}).$$

$\Leftarrow$: Aus Lemma 5.3.7 folgt $A \subset A^{**}$ und nach Satz 5.3.1 ist A^{**} abgeschlossen. □

Wie in der Matrizenrechnung spielen in der Operatorentheorie – und besonders auch in den Anwendungen – die symmetrischen und selbstadjungierten Abbildungen eine besondere Rolle. Für sie erhält man handliche Darstellungsformeln (Verallgemeinerung der Hauptachsentransformation), die das weitere Arbeiten mit ihnen erleichtern. Wir präzisieren zunächst:

Definition 5.3.9: *Es seien $\mathcal{X}$ ein Hilbertraum und*

$$A \;:\; \mathcal{D}(A) \subset \mathcal{X} \longrightarrow \mathcal{X}$$

eine dicht definierte lineare Abbildung. Dann heißt:

(1) A symmetrisch, *wenn $A \subset A^*$ ist. Dann ist also*

$$\forall\, x, y \in \mathcal{D}(A) \qquad (Ax, y) = (x, Ay).$$

(2) A selbstadjungiert, *wenn $A = A^*$ ist.*

(3) A wesentlich selbstadjungiert, *wenn A abschließbar ist und $\bar{A} = A^*$ gilt.*

Man beachte den Unterschied zwischen symmetrischen und selbstadjungierten Operatoren, der in endlich dimensionalen Räumen noch nicht auftritt. In unendlich dimensionalen Räumen ist es manchmal schwierig herauszufinden, ob ein symmetrischer Operator bereits selbstadjungiert ist oder zu einem solchen fortgesetzt werden kann. Auf diese Fragen werden wir in §8 ausführlicher eingehen.

Satz 5.3.10: *Es sei A eine symmetrische Abbildung. Dann gilt:*

(1) A ist abschließbar.

(2) $\bar{A}$ ist symmetrisch.

(3) $A \subset \bar{A} \subset \bar{A}^* = A^*$.

Beweis:

Zu (1): Es seien $x_n \in \mathcal{D}(A)$ mit $x_n \to o$ und $Ax_n \to y$. Dann folgt für alle $x \in \mathcal{D}(A)$

$$0 \leftarrow (Ax, x_n) = (x, Ax_n) \to (x, y),$$

also $y = o$. Mithin ist A abschließbar, und es gilt

$$A \subset \bar{A} \subset A^*,$$

letzteres weil A^* abgeschlossen ist.

Zu (2): Es seien $x_n, y_n \in \mathcal{D}(A)$ mit $x_n \to x \in \mathcal{D}(\bar{A})$, $Ax_n \to \bar{A}x$, $y_n \to y \in \mathcal{D}(\bar{A})$ und $Ay_n \to \bar{A}y$. Dann ist

$$(Ax_n, y_n) = (x_n, Ay_n),$$

also

$$(\bar{A}x, y) = (x, \bar{A}y).$$

Mithin ist $\bar{A}$ symmetrisch.

Zu (3): Aus der Definition folgt

$$\mathcal{D}(\bar{A}^*) = \left\{ y \in \mathcal{X} \;\middle|\; \exists\, z \in \mathcal{X} \quad \forall\, x \in \mathcal{D}(\bar{A}) \qquad (y, \bar{A}x) = (z, x) \right\} = \mathcal{D}(A^*). \qquad \square$$

Eine symmetrische Abbildung A besitzt im allgemeinen also zwei abgeschlossene Fortsetzungen, nämlich

$$A \subset \bar{A} = A^{**} \quad \text{und} \quad A \subset A^*.$$

$\bar{A} = A^{**}$ ist wieder symmetrisch, A^* im allgemeinen aber nicht. Es gilt

$$A \subset \bar{A} = A^{**} \subset A^* = \bar{A}^*.$$

Ist A^* symmetrisch, gilt also $A^* \subset A^{**}$, dann folgt daraus

$$\bar{A} \subset \bar{A}^* = A^* \subset A^{**} = \bar{A}.$$

A ist dann also wesentlich selbstadjungiert. Es gilt

Folgerung 5.3.11: *Es sei A eine symmetrische Abbildung. Dann ist A^* genau dann symmetrisch, wenn A wesentlich selbstadjungiert ist. A^* ist dann selbstadjungiert.*

Beweis: Es sei A wesentlich selbstadjungiert. Dann ist

$$A^{**} = \bar{A} = A^*,$$

A^* also selbstadjungiert. Die andere Richtung ist bereits gezeigt worden. $\square$

Satz 5.3.12: *Es sei A eine injektive symmetrische Abbildung. Dann gilt:*

(1) $\mathcal{R}(A)$ sei in $\mathcal{X}$ dicht. Dann ist A^{-1} symmetrisch, und es gilt $A^{-1} \subset (A^{-1})^* = (A^*)^{-1}$.

(2) Es sei $A = A^*$. Dann ist A^{-1} selbstadjungiert.

Beweis:

Zu (1): Nach Satz 5.3.6 ist

$$(A^{-1})^* = (A^*)^{-1},$$

und aus $A \subset A^*$ folgt $A^{-1} \subset (A^*)^{-1}$.

Zu (2): Es gilt

$$\mathcal{X} = \overline{\mathcal{R}(A)} \oplus \mathcal{N}(A^*) = \overline{\mathcal{R}(A)} \oplus \mathcal{N}(A) = \overline{\mathcal{R}(A)}.$$

Deshalb ist A^{-1} dicht definiert und nach (1) symmetrisch. Wegen

$$(A^{-1})^* = (A^*)^{-1} = A^{-1}$$

ist A^{-1} selbstadjungiert. $\square$

Satz 5.3.13: *Es sei A eine symmetrische Abbildung mit $\mathcal{D}(A) = \mathcal{X}$ oder $\mathcal{R}(A) = \mathcal{X}$. Dann ist A selbstadjungiert.*

Beweis Im Falle $\mathcal{D}(A) = \mathcal{X}$ ist nichts zu zeigen. Es sei also $\mathcal{R}(A) = \mathcal{X}$. Dann folgt aus

$$\mathcal{X} = \overline{\mathcal{R}(A)} \oplus \mathcal{N}(A^*)$$

$\mathcal{N}(A^*) = \{o\}$, also

$$\mathcal{N}(A) \subset \mathcal{N}(A^*) = \{o\}.$$

Mithin existiert A^{-1} und ist symmetrisch. Wegen $\mathcal{D}(A^{-1}) = \mathcal{X}$ ist A^{-1} sogar selbstadjungiert und damit auch A. $\square$

Als letztes geben wir noch eine Charakterisierung symmetrischer Operatoren in komplexen Hilberträumen.

Satz 5.3.14: *Es seien $\mathcal{X}$ ein Hilbertraum über $\mathbb{C}$ und A eine dicht definierte lineare Abbildung. Dann ist A genau dann symmetrisch, wenn*

$$\forall\, x \in \mathcal{D}(A) \qquad (Ax, x) \in \mathbb{R}$$

gilt.

Beweis:

$\Rightarrow$: A sei symmetrisch. Dann ist

$$\overline{(Ax, x)} = (x, Ax) = (Ax, x).$$

$\Leftarrow$: Für alle $x \in \mathcal{D}(A)$ sei $(Ax, x) \in \mathbb{R}$. Dann folgt mit $x, y \in \mathcal{D}(A)$

$$(Ax, y) + (Ay, x) = \big(A(x+y), x+y\big) - (Ax, x) - (Ay, y) \in \mathbb{R}$$

und

$$\mathrm{i}\,\big((Ay, x) - (Ax, y)\big) = \big(A(x+\mathrm{i}y), x+\mathrm{i}y\big) - (Ax, x) - (Ay, y) \in \mathbb{R},$$

also

$$\mathrm{Im}\,(Ax, y) = -\mathrm{Im}\,(Ay, x)$$

und

$$\mathrm{Re}\,(Ax, y) = \mathrm{Re}\,(Ay, x).$$

Daraus folgt

$$(Ax, y) = \mathrm{Re}\,(Ay, x) - \mathrm{i}\,\mathrm{Im}\,(Ay, x) = \overline{(Ay, x)} = (x, Ay). \qquad \square$$

5.4 Ein Beispiel: Der Δ-Operator

Wir wollen nun die gewonnenen Resultate an einem Beispiel verdeutlichen. Es seien G ein Gebiet in $\mathbb{R}^n$, $\mathcal{H} = \mathcal{L}^2(G)$ ein komplexer Hilbertraum, $\mathcal{D}(A) := \mathring{\mathcal{C}}_\infty(G)$ und

$$\begin{aligned} A \;:\; \mathcal{D}(A) \subset \mathcal{H} &\longrightarrow \mathcal{H}, \\ u &\longmapsto -\Delta u. \end{aligned}$$

Dabei ist Δ der *Laplace-Operator*, also

$$\Delta := \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2}.$$

Dieser Differentialoperator tritt in der Physik häufig auf. Man findet ihn in der Elektrostatik, bei Schwingungsvorgängen oder in der Theorie der Schrödingergleichung. Er beschreibt normalerweise physikalisches Verhalten in homogenen isotropen Medien. Allgemeiner hätte man entsprechend

$$\sum_{i=1}^{n} \frac{\partial}{\partial x_i}\, a_{ik}(x) \frac{\partial}{\partial x_k}$$

mit einer positiv definiten Matrix (a_{ik}) zu diskutieren. Wir wollen uns hier aber auf den einfacheren Fall des Δ-Operators beschränken.

Aus der Definition des Δ-Operators folgt unmittelbar

$$\forall\, u \in \mathcal{D}(A) \qquad (Au, u) = \|\nabla u\|^2 \geq 0.$$

Man sagt deshalb auch, A sei *halbbeschränkt*. Aus der Lebesgueschen Integrationstheorie ist klar, daß $\mathcal{D}(A)$ in $\mathcal{H}$ dicht ist. Nach Satz 5.3.14 ist A deshalb symmetrisch und nach Satz 5.3.10 abschließbar. Deshalb stellt sich als erstes die Frage nach der Charakterisierung von $\bar{A}$.

Für Testfunktionen $\varphi \in \mathring{\mathcal{C}}_\infty(G) = \mathcal{D}(A)$ folgt durch partielles Integrieren

$$\begin{aligned} \|\Delta\varphi\|^2 &= \|\partial_1^2\varphi\|^2 + \cdots + 2(\partial_1^2\varphi, \partial_2^2\varphi) + \cdots \\ &= \|\partial_1^2\varphi\|^2 + \cdots + 2(\partial_1\partial_2\varphi, \partial_1\partial_2\varphi) + \cdots \\ &= |\varphi|_2^2. \end{aligned}$$

Dabei ist

$$|\varphi|_2 := \sqrt{\sum_{|\alpha|=2} \|\partial^\alpha \varphi\|^2}$$

die Halbnorm der zweiten Ableitungen. Allgemein sei $\alpha \in \mathbb{N}_0^n$ ein Multiindex, $\alpha = (\alpha_1, \ldots, \alpha_n)$, mit

$$|\alpha| := \alpha_1 + \alpha_2 + \cdots + \alpha_n$$

und entsprechend

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

mit

$$\partial_i^{\alpha_i} := \left(\frac{\partial}{\partial x_i}\right)^{\alpha_i}.$$

Dann ist für $k > 0$

$$|\varphi|_k := \sqrt{\sum_{|\alpha|=k} \|\partial^\alpha \varphi\|^2}$$

die Halbnorm der k-ten Ableitungen und

$$\|\varphi\|_k := \sqrt{\sum_{0 \le j \le k} |\varphi|_j^2}$$

die *k-Norm*. Dabei sei $|\varphi|_0 := \|\varphi\|$.

Für den Δ-Operator gilt also

$$\forall\, \varphi \in \mathcal{D}(A) \qquad \|A\varphi\| = |\varphi|_2.$$

Es ist auch

$$|\varphi|_1^2 = (\nabla\varphi, \nabla\varphi) = -(\Delta\varphi, \varphi) = (A\varphi, \varphi) \le \|\varphi\| \cdot \|A\varphi\| \le \frac{1}{2}\|\varphi\|_A^2.$$

Zusammen folgt daraus

$$\|\varphi\|_2^2 \le \|\varphi\|^2 + \frac{1}{2}\|\varphi\|_A^2 + \|A\varphi\|^2 = \frac{3}{2}\|\varphi\|_A^2$$

oder

$$\sqrt{\tfrac{2}{3}}\,\|\varphi\|_2 \le \|\varphi\|_A \le \|\varphi\|_2. \qquad (*)$$

Diese Abschätzung ist besonders wichtig. Sie impliziert sofort

$$\mathcal{D}(\bar{A}) = \left\{\mathring{\mathcal{C}}_\infty(G);\ \|\cdot\|_2\right\}^\sim =: \mathring{\mathcal{H}}_2(G).$$

Analog wird $\mathring{\mathcal{H}}_k(G)$ definiert, nämlich

$$\mathring{\mathcal{H}}_k(G) := \left\{\mathring{\mathcal{C}}_\infty(G);\ \|\cdot\|_k\right\}^\sim$$

und auch

$$\mathcal{H}_k(G) := \left\{\mathcal{C}_k^*(G);\ \|\cdot\|_k\right\}^\sim$$

mit

$$\mathcal{C}_k^*(G) := \left\{u \in \mathcal{C}_k(G) \mid \|u\|_k < \infty\right\}.$$

Die Räume $\mathcal{H}_k(G)$ bzw. $\mathring{\mathcal{H}}_k(G)$ nennt man *Sobolevräume*, sie werden im nächsten Kapitel ausführlicher diskutiert. Sie sind Hilberträume mit dem Skalarprodukt

$$(u, v)_k := \sum_{0 \le j \le k} \sum_{|\alpha|=j} (\partial^\alpha u,\, \partial^\alpha v).$$

Es ist $\mathcal{H}_0(G) = \mathcal{L}^2(G)$. $u \in \mathcal{H}_k(G)$ bedeutet, daß die verallgemeinerten Ableitungen von u bis zur k-ten Ordnung existieren, $u \in \mathring{\mathcal{H}}_k(G)$ bedeutet, daß außerdem u und die Ableitungen von u bis zur $(k-1)$-ten Ordnung im verallgemeinerten Sinne am Rande ∂G verschwinden.

Damit haben wir $\bar{A}$ charakterisiert. Es ist $\mathcal{D}(\bar{A}) = \mathring{\mathcal{H}}_2(G)$, und es gilt für $u \in \mathcal{D}(\bar{A})$

$$\forall\, \varphi \in \mathring{\mathcal{C}}_\infty(G) \qquad (\bar{A}u, \varphi) = (u, -\Delta\varphi).$$

Man sagt dafür auch, $\bar{A}u$ sei im schwachen Sinne gleich $-\Delta u$.

Nachdem wir nun A abgeschlossen haben, geht es als nächstes um die Frage nach der Selbstadjungiertheit von $\bar{A}$. Charakterisieren wir dazu A^*: Es ist

$$\begin{aligned}\mathcal{D}(A^*) &= \big\{u \in \mathcal{H} \;\big|\; \exists\, f \in \mathcal{H} \quad \forall\, \varphi \in \mathring{\mathcal{C}}_\infty(G) \quad (u, A\varphi) = (f, \varphi)\big\} \\ &= \big\{u \in \mathcal{H} \;\big|\; \Delta u \in \mathcal{H} \;\text{ (im schwachen Sinne)}\big\}\end{aligned}$$

und

$$A^* u := f = -\Delta u.$$

Es gilt also $\mathcal{H}_2(G) \subset \mathcal{D}(A^*)$ und

$$A \subset \bar{A} = A^{**} \subset A^*.$$

Dabei ist im allgemeinen $\mathcal{H}_2(G) \subsetneq \mathcal{D}(A^*)$. Das sieht man am folgenden Beispiel: Es sei

$$G := \big\{x \in \mathbb{R}^3 \;\big|\; 0 < |x| < 1\big\}$$

und

$$u(x) := \frac{1}{|x|}.$$

Dann ist $u \in \mathcal{H}$ und $\Delta u = 0$, also $u \in \mathcal{D}(A^*)$. Andererseits ist aber

$$\partial_1 u(x) = -\frac{x_1}{|x|^3} \notin \mathcal{H},$$

also $u \notin \mathcal{H}_1(G)$ und damit erst recht $u \notin \mathcal{H}_2(G)$. Damit ist $\bar{A}$ auch sicherlich im allgemeinen nicht selbstadjungiert.

Die Problematik der Existenz selbstadjungierter Fortsetzungen werden wir erst in §8 aufgreifen. Für den Fall des Δ-Operators wollen wir sie hier nur andiskutieren und betrachten zwei Fälle:

1. Es sei $G = \mathbb{R}^n$. Dann kann man zeigen, daß Funktionen mit kompaktem Träger dicht in $\mathcal{H}_2(\mathbb{R}^n)$ sind. Daraus folgt

   $$\mathcal{H}_2(\mathbb{R}^n) = \mathring{\mathcal{H}}_2(\mathbb{R}^n).$$

   Ferner kann man – etwa durch Fouriertransformation – nachrechnen, daß aus

   $$u \in \mathcal{H} \quad \text{und} \quad \Delta u \in \mathcal{H}$$

   bereits $u \in \mathcal{H}_2(\mathbb{R}^n)$ folgt. Damit ist in diesem Fall

   $$\mathcal{D}(\bar{A}) = \mathcal{D}(A^*),$$

   also $\bar{A} = A^*$, und A ist wesentlich selbstadjungiert.

2. Es sei G ein beschränktes Gebiet in $\mathbb{R}^n$, etwa $G = B(0,1)$. Jetzt ist

   $$\mathring{\mathcal{H}}_2(G) \subsetneq \mathcal{H}_2(G),$$

   man nehme nur $u = 1$, und $\bar{A}$ ist nicht selbstadjungiert. $\mathcal{D}(\bar{A}) = \mathring{\mathcal{H}}_2(G)$ ist zu klein und $\mathcal{D}(A^*) \supset \mathcal{H}_2(G)$ ist zu groß dafür. Man möchte also ein $B \supset \bar{A}$ finden mit

   $$A \subset \bar{A} \subset B = B^* \subset A^*.$$

   Ein Beispiel mit $G = (0,1) \subset \mathbb{R}^1$ möge das erläutern. Es ist

   $$\int_0^1 fg'' = \int_0^1 f''g + (fg' - f'g)\Big|_0^1.$$

   Zur Selbstadjungiertheit muß also

   $$(fg' - f'g)\Big|_0^1 = 0$$

sein. Wählen wir $f \in \mathcal{D}(\bar{A}) = \mathring{\mathcal{H}}_2(G)$, dann verschwinden sowohl f als auch f' am Rande, und man benötigt für g gar keine Randbedingung mehr. Das sollte man besser ausbalancieren, und dazu gibt es offenbar viele Möglichkeiten. Eine einfache ist wohl, sowohl für f als auch für g die *Dirichletsche Randbedingung* zu stellen, $f(0) = f(1) = 0$. Möglich wären aber wohl auch die *Neumannsche*, $f'(0) = f'(1) = 0$, und viele Kombinationen beider. Bleiben wir bei der Dirichletschen Randbedingung. Man formuliert sie durch $u \in \mathring{\mathcal{H}}_1(G)$. Es sei also

$$\begin{aligned} \mathcal{D}(B) &:= \left\{ u \in \mathring{\mathcal{H}}_1(G) \;\middle|\; \Delta u \in \mathcal{H} \;\; (\text{im schwachen Sinne}) \right\} \\ Bu &:= -\Delta u \;\; (\text{im schwachen Sinne}). \end{aligned}$$

Dann ist B eine symmetrische Fortsetzung von $\bar{A}$. Es seien nämlich $u, v \in \mathcal{D}(B)$ und $\varphi_n \in \mathring{\mathcal{C}}_\infty(G)$ mit $\|\varphi_n - v\|_1 \to 0$. Dann ist

$$(Bu, v) \leftarrow (Bu, \varphi_n) = -(u, \Delta\varphi_n) = (\nabla u, \nabla\varphi_n) \to (\nabla u, \nabla v).$$

Entsprechend folgt

$$(u, Bv) = (\nabla u, \nabla v).$$

Mithin ist B symmetrisch, also abschließbar, und es folgt mit $k := (\sqrt{2}+1)/2$

$$\forall\, u \in \mathcal{D}(B) \qquad \|u\|_1^2 = \|u\|^2 + (u, Bu) \le k\, \|u\|_B^2.$$

Es seien $u_n \in \mathcal{D}(B)$ und $u \in \mathcal{D}(\bar{B})$ mit

$$\|u_n - u\|_B \to 0.$$

Dann gilt auch $\|u_n - u\|_1 \to 0$. Folglich ist $u \in \mathring{\mathcal{H}}_1(G)$, und es existiert Δu im schwachen Sinne. Es ist also $u \in \mathcal{D}(B)$ und B damit abgeschlossen. Der Nachweis der Selbstadjungiertheit ist etwas schwieriger. In Satz 8.1.2 zeigen wir für abgeschlossene symmetrische Abbildungen, daß $\mathcal{D}(A^*)$ als direkte Summe

$$\mathcal{D}(A^*) = \mathcal{D}(A) \oplus \mathcal{R}(A + \mathrm{i})^\perp \oplus \mathcal{R}(A - \mathrm{i})^\perp$$

geschrieben werden kann. Man beweist deshalb, daß für alle $f \in \mathcal{H}$ die Gleichungen

$$(A \pm \mathrm{i})u \stackrel{!}{=} f$$

eindeutig lösbar sind. Dazu verwendet man eine leichte Verallgemeinerung des Rieszschen Darstellungssatzes, in dem man anstelle des Skalarproduktes in $\mathring{\mathcal{H}}_1(G)$ die Formen

$$B(u, v) := (\nabla u, \nabla v) \pm \mathrm{i}\,(u, v)$$

wählt. Wegen

$$|B(u, u)| = \|u\|_1^2$$

sind diese Formen *koerzitiv* über $\mathring{\mathcal{H}}_1(G)$, also $p\,\|u\|_1^2 \le |B(u,u)|$ mit $p > 0$, und die Gleichungen

$$\forall\, v \in \mathring{\mathcal{H}}_1 \qquad B(u, v) \stackrel{!}{=} (f, v)$$

können dann eindeutig nach $u \in \mathring{\mathcal{H}}_1$ aufgelöst werden. $u \in \mathcal{D}(B)$ folgt unmittelbar. Damit sind $\mathcal{R}(B \pm \mathrm{i})^\perp = \{o\}$, und B ist selbstadjungiert.

6 Spezielle Funktionenräume

In diesem Kapitel werden einige spezielle Funktionenräume eingeführt und wichtige Eigenschaften solcher Funktionen kurz zusammengestellt. Gedacht ist dabei vor allem an klassische Räume stetiger Funktionen, an Sobolevräume und Distributionen.

Räume von Zahlenfolgen sind uns bereits mehrfach begegnet. Erinnert sei an den vollständigen metrischen Raum

$$\mathbb{R}^{\mathbb{N}} := \left\{ x = (\xi_i)_{i\in\mathbb{N}} \;\middle|\; \xi_i \in \mathbb{R} \right\}$$

mit der Metrik

$$d(x,y) := \sum_{i=1}^{\infty} \frac{|\xi_i - \eta_i|}{2^i(1+|\xi_i-\eta_i|)}$$

oder an die vollständigen normierten Räume

$$\ell^p := \left\{ x = (\xi_i)_{i\in\mathbb{N}} \;\middle|\; \|x\|_{\ell^p} < \infty \right\}$$

mit $1 \le p \le \infty$ und

$$\|x\|_{\ell^p} := \begin{cases} \left(\sum_{i=1}^{\infty} |\xi_i|^p\right)^{1/p} & \text{für } 1 \le p < \infty \\ \sup_{i\in\mathbb{N}} |\xi_i| & \text{für } p = \infty. \end{cases}$$

Diese Räume sind für $1 < p < \infty$ reflexiv und separabel. Der Raum ℓ^1 ist ebenfalls separabel, ℓ^∞ ist es nicht. Im folgenden wollen wir uns Räume von Abbildungen ansehen.

6.1 Klassische Funktionenräume

In diesem Abschnitt besprechen wir kurz einige klassische Räume stetiger Funktionen. Es sei grundsätzlich $\Omega \subset \mathbb{R}^n$ eine offene Teilmenge des $\mathbb{R}^n$ und

$$\begin{aligned} \mathcal{C}(\Omega) &:= \left\{ f : \Omega \to \mathbb{K} \;\middle|\; f \text{ ist stetig} \right\}, \\ \mathcal{C}_k(\Omega) &:= \left\{ f : \Omega \to \mathbb{K} \;\middle|\; f \text{ ist k-mal stetig differenzierbar} \right\}, \\ \mathcal{C}_\infty(\Omega) &:= \bigcap_{k\in\mathbb{N}_0} \mathcal{C}_k(\Omega). \end{aligned}$$

Dabei sei stets $\mathcal{C}_0(\Omega) := \mathcal{C}(\Omega)$. Gelegentlich werden wir auch vektorwertige Abbildungen, also

$$f \;:\; \Omega \longrightarrow \mathbb{K}^n$$

zulassen.

Es sei noch einmal an die Notation mit *Multiindizes* in §5.4 erinnert. Es sei

$$\alpha := (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{N}_0^n$$

ein solcher Multiindex. Dann ist

$$\begin{aligned} \alpha \le \beta :\Leftrightarrow \forall\, i = 1,\ldots,n \quad \alpha_i \le \beta_i \\ |\alpha| := \textstyle\sum_{i=1}^n \alpha_i \\ \alpha! := \textstyle\prod_{i=1}^n \alpha_i! \\ \binom{\alpha}{\beta} := \textstyle\prod_{i=1}^n \binom{\alpha_i}{\beta_i} \end{aligned}$$

Für $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ sei

$$x^\alpha := \prod_{i=1}^n x_i^{\alpha_i}.$$

Die Ableitungen schreibt man analog. Es seien

$$\partial_k f := \frac{\partial f}{\partial x_k} \quad \text{und} \quad \partial_k^l f := \frac{\partial^l f}{(\partial x_k)^l}$$

sowie

$$\partial^0 f := f \quad \text{und} \quad \partial^\alpha f := \partial_1^{\alpha_1}\partial_2^{\alpha_2}\cdots\partial_n^{\alpha_n} f.$$

Ferner verwenden wir

Definition 6.1.1: *Für eine offene Teilmenge Ω des $\mathbb{R}^n$ und $f : \Omega \to \mathbb{K}$ seien:*

1. Der Träger *(engl.* support*) von $f \in \mathcal{C}(\Omega)$:$\Longleftrightarrow$*

$$\operatorname{supp} f := \overline{\{x \in \Omega \mid f(x) \neq 0\}}.$$

2. $K \Subset \Omega$:$\Longleftrightarrow$ K ist eine kompakte Teilmenge von Ω.

3. Für $m \in \mathbb{N}_0 \cup \{\infty\}$

$$\mathring{\mathcal{C}}_m(\Omega) := \{f \in \mathcal{C}_m(\Omega) \mid \operatorname{supp} f \Subset \Omega\}.$$

Die Elemente von $\mathring{\mathcal{C}}_\infty(\Omega)$ heißen Testfunktionen.

Der Raum $\mathcal{C}_k(\Omega)$ läßt sich leicht metrisieren. Für $K \Subset \Omega$ und $l \in \mathbb{N}_0$ mit $l \leq k$ seien

$$\begin{aligned} P_{K,l} \;:\; \mathcal{C}_k(\Omega) &\longrightarrow \mathbb{R}_0^+, \\ f &\longmapsto \sup_{|\alpha|=l} \sup_{x \in K} |\partial^\alpha f(x)|. \end{aligned}$$

Die $P_{K,l}$ sind Halbnormen auf $\mathcal{C}_k(\Omega)$. Die offene Menge Ω läßt sich nun durch $K_\nu \Subset \Omega$ ausschöpfen, etwa durch

$$K_\nu := \left\{x \in \Omega \mid |x| \leq \nu \text{ und } \operatorname{dist}(x, \mathbb{R}^n \backslash \Omega) \geq \frac{1}{\nu}\right\}.$$

Dann ist

$$\Omega = \bigcup_{\nu \in \mathbb{N}} K_\nu.$$

Aus den Halbmetriken

$$\begin{aligned} d_\nu \;:\; \mathcal{C}_k(\Omega) \times \mathcal{C}_k(\Omega) &\longrightarrow \mathbb{R}_0^+, \\ \langle f, g\rangle &\longmapsto \sum_{l=0}^{k} \frac{P_{K_\nu,l}(f-g)}{2^l\left(1 + P_{K_\nu,l}(f-g)\right)} \end{aligned}$$

erhält man dann für $\mathcal{C}_k(\Omega)$ die Metrik

$$\begin{aligned} d \;:\; \mathcal{C}_k(\Omega) \times \mathcal{C}_k(\Omega) &\longrightarrow \mathbb{R}_0^+, \\ \langle f, g\rangle &\longmapsto \sum_{\nu=1}^{\infty} \frac{d_\nu(f,g)}{2^\nu\left(1 + d_\nu(f,g)\right)}. \end{aligned}$$

Der Raum $\mathcal{C}_k(\Omega)$ ist mit dieser Metrik vollständig. $\mathring{\mathcal{C}}_k(\Omega)$ ist es jedoch nicht; in §6.5 werden wir diese Räume weiter behandeln.

Wir betrachten nun Teilräume von $\mathcal{C}_k(\Omega)$:

1. Es seien $f, \ldots, f^{(k)} \in \mathcal{C}(\Omega)$ auf $\bar{\Omega}$ stetig differenzierbar fortsetzbar. Dann ist $f \in \mathcal{C}_k(\bar{\Omega})$.

2. Es sei

$$f \in \mathcal{BC}_k(\Omega) := \left\{f \in \mathcal{C}_k(\Omega) \mid \forall\, \alpha,\, |\alpha| \leq k, \quad \sup_{x \in \Omega} |\partial^\alpha f(x)| < \infty\right\}.$$

In beiden Fällen wählen wir die Supremumsnorm

$$\|f\|_{\mathcal{BC}_k(\Omega)} := \sum_{|\alpha| \leq k} \sup_{x \in \Omega} |\partial^\alpha f(x)|.$$

3. Es seien $f : \bar{\Omega} \to \mathbb{K}$ und für ein $\alpha \in (0, 1]$

$$\operatorname{höl}_\alpha(f, \bar{\Omega}) := \sup_{x,y \in \bar{\Omega}} \frac{|f(x) - f(y)|}{|x - y|^\alpha}.$$

Dann ist

$$\mathcal{C}_{k,\alpha}(\bar{\Omega}) := \left\{f \in \mathcal{C}_k(\bar{\Omega}) \mid \operatorname{höl}_\alpha(\partial^\beta f, \bar{\Omega}) < \infty \quad \text{für } |\beta| = k\right\}$$

der Raum der hölderstetig differenzierbaren Funktionen. $\mathcal{C}_{k,\alpha}(\bar{\Omega})$ ist ein Banachraum mit der Norm

$$\|f\|_{\mathcal{C}_{k,\alpha}(\bar{\Omega})} := \|f\|_{\mathcal{BC}_k(\bar{\Omega})} + \sum_{|\beta|=k} \operatorname{höl}_\alpha(\partial^\beta f, \bar{\Omega}).$$

Eine Funktion $f \in \mathcal{C}_{0,\alpha}(\bar{\Omega})$ nennt man *hölderstetig* und im Falle $\alpha = 1$ *lipschitzstetig*. Benannt sind diese Räume nach Otto Hölder, 1859–1937, und Rudolf Lipschitz, 1832–1903.

6.2 Lebesguesche Räume

In dieser Vorlesung setzen wir das Lebesguesche Integral als bekannt voraus. Eine Einführung findet man zum Beispiel in *Alt [1985], S. 29f.* Es sei also M eine bezüglich des Lebesgue-Maßes meßbare Teilmenge des $\mathbb{R}^n$ und

$$f \;:\; M \longrightarrow \mathbb{K}$$

meßbar. Dann ist für $1 \leq p < \infty$

$$\mathcal{L}^p(M) := \left\{ f : M \to \mathbb{K} \;\middle|\; \int_M |f|^p < \infty \right\}$$

und für $p = \infty$

$$\mathcal{L}^\infty(M) := \left\{ f : M \to \mathbb{K} \;\middle|\; \operatorname*{ess\,sup}_{x \in M} |f(x)| < \infty \right\}.$$

Diese Räume werden mit den Normen

$$\|f\|_{\mathcal{L}^p(M)} := \begin{cases} \left(\int_M |f|^p\right)^{1/p} & \text{für } 1 \leq p < \infty \\ \operatorname{ess\,sup}_{x \in M} |f(x)| & \text{für } p = \infty \end{cases}$$

versehen. $\mathcal{L}^2(M)$ ist als Hilbertraum mit dem Skalarprodukt

$$(f, g) := \int_M f(x)\overline{g(x)}\, dx$$

besonders interessant. Entsprechend werden auch die $\mathcal{L}^p_k(\Omega)$-Räume der k-mal differenzierbaren Funktionen definiert mit

$$\|f\|_{\mathcal{L}^p_k(\Omega)} := \sum_{0 \leq |\alpha| \leq k} \left\|\partial^\alpha f\right\|_{\mathcal{L}^p(\Omega)}.$$

Wir verwenden ebenfalls

$$\mathcal{L}^p_{loc}(M) := \left\{ f : M \to \mathbb{K} \;\middle|\; \forall\, K \Subset M \;\; f \in \mathcal{L}^p(K) \right\}.$$

Es folgen einige Bemerkungen:

1. Für $p, q \in (1, \infty)$ mit

$$\frac{1}{p} + \frac{1}{q} = 1$$

gilt die *Höldersche Ungleichung*

$$\int_M |fg| \leq \Big(\int_M |f|^p \Big)^{\frac{1}{p}} \Big(\int_M |g|^q \Big)^{\frac{1}{q}}.$$

2. Es sei das Maß $\mu(M) < \infty$. Dann folgt aus

$$\int_M |f| = \int_M \chi_M \cdot |f| \leq \mu(M)^{1/q}\, \|f\|_{\mathcal{L}^p(M)}$$

für $1 \leq p \leq \infty$

$$\mathcal{L}^p(M) \subset \mathcal{L}^1(M).$$

3. Für $1 \leq p < \infty$ ist $\mathcal{L}^q(M)$ zu $\left(\mathcal{L}^p(M)\right)'$ kongruent.

4. Für $1 \leq p < \infty$ ist $\mathring{\mathcal{C}}_0(\mathbb{R}^n)$ in $\mathcal{L}^p(\mathbb{R}^n)$ dicht.

Satz 6.2.1: *Es seien $1 \leq p < \infty$, $f \in \mathcal{L}^p(\mathbb{R}^n)$ und $h \in \mathbb{R}^n$. Dann gilt*

$$\lim_{h \to o} \left\| f(\cdot + h) - f \right\|_{\mathcal{L}^p(\mathbb{R}^n)} = 0.$$

Beweis: Zu $\varepsilon > 0$ wählen wir ein $\varphi \in \mathring{\mathcal{C}}_0(\mathbb{R}^n)$ mit

$$\|f - \varphi\|_{\mathcal{L}^p(\mathbb{R}^n)} < \frac{\varepsilon}{3}.$$

Es sei $k := \big(2\,\mu(\operatorname{supp}\varphi)\big)^{1/p}$. Dann folgt aus der gleichmäßigen Stetigkeit von φ

$$\exists\,\delta > 0 \quad \forall\, h \in \mathbb{R}^n,\ |h| < \delta, \quad \forall\, x \in \mathbb{R}^n \qquad k\,|\varphi(x+h) - \varphi(x)| < \frac{\varepsilon}{3}.$$

Für $|h| < \delta$ erhalten wir deshalb

$$\|f(\cdot + h) - f\|_{\mathcal{L}^p(\mathbb{R}^n)} \le \|f(\cdot + h) - \varphi(\cdot + h)\| + \|\varphi(\cdot + h) - \varphi\| + \|\varphi - f\| \le$$
$$\le 2\,\|f - \varphi\| + k \sup_{x \in M} |\varphi(x+h) - \varphi(x)| < \varepsilon.$$

□

6.3 Einige Abschätzungen

Es folgen nun einige elementare Abschätzungen, die teilweise nur in $\mathbb{R}^1$ gelten, deren Kenntnis aber das weitere Verständnis erleichtert. Wenn nicht anders gekennzeichnet, seien $I := (a,b) \subset \mathbb{R}^1$, $|I| := b - a$ und

$$\|f\| := \|f\|_{\mathcal{L}^2(I)}.$$

Abschätzung 1: *Es sei $f \in \mathcal{C}_2(I) \cap \mathcal{L}^2(I)$ mit $f'' \in \mathcal{L}^2(I)$. Dann ist auch $f' \in \mathcal{L}^2(I)$, und es gilt*

$$\|f'\|^2 \le 41\Big\{\frac{1}{|I|^2}\|f\|^2 + |I|^2\,\|f''\|^2\Big\}.$$

Wir führen den Beweis in zwei Schritten:

1. Es seien $a = 0$, $b = 1$, $\alpha \in (0, \frac{1}{2})$, $x_1 < \alpha$ und $x_2 > 1 - \alpha$. Dann gibt es ein $\eta \in (x_1, x_2)$ mit

$$f'(\eta) = \frac{f(x_2) - f(x_1)}{x_2 - x_1},$$

also

$$|f'(\eta)| \le \frac{1}{1 - 2\alpha}\,\big(|f(x_2)| + |f(x_1)|\big).$$

Daraus folgt

$$|f'(x)| = \Big|f'(\eta) + \int_\eta^x f''\Big| \le \frac{|f(x_2)| + |f(x_1)|}{1 - 2\alpha} + \int_0^1 |f''|.$$

Integrieren wir diese Abschätzung bezüglich x_1 über $(0,\alpha)$ und bezüglich x_2 über $(1-\alpha, 1)$ dann erhalten wir

$$\alpha^2 |f'(x)| \le \frac{\alpha}{1 - 2\alpha}\Big\{\int_0^\alpha |f| + \int_{1-\alpha}^1 |f|\Big\} + \alpha^2 \int_0^1 |f''|$$

oder

$$|f'(x)|^2 \le 3\Big\{\frac{1}{\alpha(1-2\alpha)^2}\|f\|^2 + \|f''\|^2\Big\},$$

also mit $\alpha = 1/6$

$$|f'(x)|^2 \le 3\Big\{\frac{27}{2}\|f\|^2 + \|f''\|^2\Big\}.$$

2. Es sei nun I ein beliebiges Intervall. Dann transformieren wir für $a \le x \le b$ und $0 \le \xi \le 1$

$$x := \xi(b - a) + a$$

und erhalten mit $f(x) =: g(\xi)$

$$f'(x) = \frac{1}{b-a} g'(\xi), \qquad f''(x) = \frac{1}{(b-a)^2}\, g''(\xi).$$

Deshalb ist

$$\|f\|^2 = (b-a)\,\|g\|^2, \qquad \|f'\|^2 = \frac{1}{b-a}\|g'\|^2, \qquad \|f''\|^2 = \frac{1}{(b-a)^3}\,\|g''\|^2,$$

und wir erhalten schließlich

$$\|f'\|^2 = \frac{1}{|I|}\,\|g'\|^2 \le \frac{81}{2|I|}\,\|g\|^2 + \frac{3}{|I|}\,\|g''\|^2 \le \frac{81}{2|I|^2}\,\|f\|^2 + 3|I|^2\,\|f\|^2.$$

□

Abschätzung 2: *Es sei $f \in \mathcal{C}_1(I)$ mit $f' \in \mathcal{L}^2(I)$. Dann ist $f \in \mathcal{L}^2(I)$, und es gilt*

$$\|f\|^2 \le 6\left\{|I|^2\,\|f'\|^2 + \int_{a'}^{b'} |f|^2\right\}$$

mit $a' := a + (b-a)/3$ und $b' := b - (b-a)/3$.

Den Beweis führen wir wieder in zwei Schritten:

1. Es seien $a = 0$, $b = 1$, $x \in (0,1)$ und $y \in (\frac{1}{3}, \frac{2}{3})$. Dann ist

$$f(x) = f(y) + \int_y^x f',$$

und es folgt

$$|f(x)|^2 \le 2|f(y)|^2 + 2|x-y|\,\|f'\|^2 \le 2|f(y)|^2 + \frac{4}{3}\,\|f'\|^2.$$

Integrieren wir bezüglich y, dann erhalten wir daraus

$$\frac{1}{3}|f(x)|^2 \le 2\int_{1/3}^{2/3} |f|^2 + \frac{4}{9}\,\|f'\|^2$$

oder

$$\|f\|^2 \le 6\int_{1/3}^{2/3} |f|^2 + \frac{4}{3}\,\|f'\|^2.$$

2. Die Substitution $x = \xi(b-a) + a$ ergibt dann wieder die Behauptung. □

Aus jeder dieser beiden Abschätzungen ziehen wir nun eine Folgerung.

Folgerung 6.3.1: *Es seien $f \in \mathcal{C}_2(I) \cap \mathcal{L}^2(I)$ mit $f'' \in \mathcal{L}^2(I)$ und $\varepsilon \in (0,1]$. Dann ist auch $f' \in \mathcal{L}^2(I)$, und es gibt ein $\gamma = \gamma(a,b) > 0$ mit*

$$\|f'\|^2 \le \gamma\left\{\varepsilon\,\|f''\|^2 + \frac{1}{\varepsilon}\,\|f\|^2\right\}$$

Beweis: Wir wählen eine Partition $a = a_0 < a_1 < \cdots < a_n = b$ mit

$$\frac{\sqrt{\varepsilon}(b-a)}{2} < a_i - a_{i-1} < \sqrt{\varepsilon}(b-a) \qquad \text{für alle } i = 1,\dots,n.$$

Dann gilt

$$\int_{a_{i-1}}^{a_i} |f'|^2 \le 41\left\{\frac{4}{\varepsilon|I|^2}\int_{a_{i-1}}^{a_i} |f|^2 + \varepsilon|I|^2\int_{a_{i-1}}^{a_i} |f''|^2\right\}$$

oder mit $\gamma := 41\max(4/|I|^2, |I|^2)$

$$\|f'\|^2 \le \gamma\left\{\varepsilon\,\|f''\|^2 + \frac{1}{\varepsilon}\,\|f\|^2\right\}.$$ □

Folgerung 6.3.2: *Es seien $f \in \mathcal{C}_1(I)$ mit $f' \in \mathcal{L}^2(I)$ und $\varepsilon \in (0,1]$. Dann ist $f \in \mathcal{L}^2(I)$, und es gibt Konstanten a', b', die nur von a, b und ε abhängen, und eine Konstante $\gamma = \gamma(a,b) > 0$ mit*

$$a < a' < b' < b$$

und

$$\|f\|^2 \le \gamma\left\{\varepsilon\,\|f'\|^2 + \int_{a'}^{b'} |f|^2\right\}.$$

Insbesondere ist also $f \in \mathcal{L}^2(I)$.

Zum Beweis wählen wir wieder die Partition $a = a_0 < a_1 < \cdots < a_n = b$ und erhalten

$$\int_{a_{i-1}}^{a_i} |f|^2 \le 6\left\{\varepsilon(b-a)^2\int_{a_{i-1}}^{a_i} |f'|^2 + \int_{a'_{i-1}}^{a'_i} |f|^2\right\}.$$

Daraus folgt mit $\gamma := 6\max(1,|I|^2)$

$$\|f\|^2 \le \gamma\Big\{\varepsilon\,\|f'\|^2 + \int_{a'}^{b'} |f|^2\Big\}. \qquad \square$$

Zu Folgerungen 6.3.1,2 analoge Aussagen gelten auch bezüglich $\Omega \subset \mathbb{R}^n$. Allerdings müssen dann Voraussetzungen an den Rand von Ω gestellt werden. Im Falle der Folgerung 6.3.1 steht auf der linken Seite der Abschätzung eine beliebige erste Ableitung. Man muß deshalb in jeder Richtung in Ω hinein integrieren können und setzt dazu für Ω die *Kegeleigenschaft* voraus. Im Falle von Folgerung 6.3.2 steht links nur f selbst, und zum Abschätzen benötigt man nur eine Differentiationsrichtung. In diesem Falle genügt deshalb die *Segmenteigenschaft*. Näheres hierzu folgt in §6.6.

Es folgen nun zwei wichtige Abschätzungen, die nach Henri Poincaré, 1854–1912, benannt sind.

Erste Poincarésche Abschätzung: *Es seien*

$$G \subset \big\{x \in \mathbb{R}^n \;\big|\; 0 < x_n < d\big\}$$

ein Gebiet in $\mathbb{R}^n$ *und* $u \in \mathring{\mathcal{C}}_1(G)$. *Dann gilt*

$$\|u\|_{\mathcal{L}^2(G)} \le d\,\|\nabla u\|_{(\mathcal{L}^2(G))^n}.$$

Für solche Funktionen ist die $|\cdot|_1$–Halbnorm also eine Norm.

Zum Beweis gehen wir von

$$u(x_1,\dots,x_n) = \int_0^{x_n} \frac{\partial}{\partial x_n}\,u(x_1,\dots,x_{n-1},t)\,dt$$

aus und erhalten

$$|u(x)|^2 \le d\int_0^d |\nabla u|^2(x_1,\dots,x_{n-1},t)\,dt$$

oder

$$\|u\|^2 \le d^2\,\|\nabla u\|^2. \qquad \square$$

Zweite Poincarésche Abschätzung: *Es seien* G *ein Gebiet in* $\mathbb{R}^n$, $n \ge 3$, *und* $u \in \mathring{\mathcal{C}}_1(G)$. *Dann gilt*

$$(n-2)\,\Big\|\frac{u(\cdot)}{|\cdot|}\Big\|_{\mathcal{L}^2(G)} \le 2\,\|\nabla u\|_{(\mathcal{L}^2(G))^n}.$$

Etwas modifizierte Aussagen gelten auch in $\mathbb{R}^1$ und $\mathbb{R}^2$.

Zum Beweis verwenden wir $r := |x|$ und gehen von

$$2\mathrm{Re}\int_G \frac{1}{r}\,\bar u\,\frac{\partial u}{\partial r}\,dx = \int_G \frac{1}{r}\,\frac{\partial}{\partial r}\,|u|^2\,dx = \int_G r^{n-2}\,\frac{\partial |u|^2}{\partial r}\,dr\,d\omega = -(n-2)\int_G \frac{|u|^2}{r^2}\,dx$$

aus. Mit $d\omega$ bezeichnen wir das Flächenelement der Sphäre. Daraus folgt

$$\int_G \Big|\frac{\partial u}{\partial r} + \frac{n-2}{r}\,u\Big|^2\,dx = \int_G \Big|\frac{\partial u}{\partial r}\Big|^2\,dx.$$

Mithin ist

$$(n-2)\Big\|\frac{u}{r}\Big\| \le \Big\|(n-2)\frac{u}{r} + \frac{\partial u}{\partial r}\Big\| + \Big\|\frac{\partial u}{\partial r}\Big\| = 2\Big\|\frac{\partial u}{\partial r}\Big\|. \qquad \square$$

6.4 Spezielle Techniken

6.4.1 Faltungen

Es seien $f, g : \mathbb{R}^n \to \mathbb{K}$ meßbar und

$$F(x) := \big(f * g\big)(x) := \int_{\mathbb{R}^n} f(y)\,g(x-y)\,dy$$

die *Faltung* von f und g. Natürlich wird die Faltung bei beliebigen f, g nicht existieren, man denke nur an $f = g = 1$. Andererseits erkennt man aber auch sofort eine triviale hinreichende Bedingung für die Existenz von $F \in \mathcal{L}^1(\mathbb{R}^n)$, nämlich $f, g \in \mathcal{L}^1(\mathbb{R}^n)$. Es sei in diesem Abschnitt wieder $\varphi \in \mathring{\mathcal{C}}_\infty(\mathbb{R}^n)$ eine Testfunktion. Als erstes zeigen wir dann

Satz 6.4.1: *Es sei $f \in \mathcal{L}^1_{loc}(\mathbb{R}^n)$. Dann ist $f * \varphi \in \mathcal{C}_\infty(\mathbb{R}^n)$, und es gilt*

$$\forall\, \alpha \in \mathbb{N}_0^n \qquad \partial^\alpha (f * \varphi) = f * \partial^\alpha \varphi.$$

Beweis: Aus

$$(f*\varphi)(x+h) - (f*\varphi)(x) - h(f*\varphi')(x) = \int_{\mathbb{R}^n} f(y)\Big\{\varphi(x-y+h) - \varphi(x-y) - h\,\varphi'(x-y)\Big\}dy$$

und

$$\varphi(x-y+h) - \varphi(x-y) = h\,\varphi'(x-y+\theta h)$$

folgt nach dem Satz von Lebesgue die Behauptung. □

Satz 6.4.2: *Es sei $f \in \mathcal{L}^p(\mathbb{R}^n)$, $1 \le p \le \infty$. Dann ist $f * \varphi \in \mathcal{L}^p(\mathbb{R}^n)$, und es gilt*

$$\|f * \varphi\|_{\mathcal{L}^p(\mathbb{R}^n)} \le \|f\|_{\mathcal{L}^p(\mathbb{R}^n)}\, \|\varphi\|_{\mathcal{L}^1(\mathbb{R}^n)}$$

Beweis: Die Fälle $p = 1$ und $p = \infty$ sind klar, es sei also $1 < p < \infty$ und $q := p/(p-1)$. Dann ist

$$\begin{aligned}\big|(f*\varphi)(x)\big| &\le \int |f(y)|\,|\varphi(x-y)|^{\frac{1}{p}}\,|\varphi(x-y)|^{\frac{1}{q}}\,dy \le \\ &\le \Big(\int |f(y)|^p\,|\varphi(x-y)|\,dy\Big)^{\frac{1}{p}}\Big(\int |\varphi(x-y)|\,dy\Big)^{\frac{1}{q}}\end{aligned}$$

oder

$$\int |f*\varphi|^p \le \Big(\int |\varphi|\Big)^{\frac{p}{q}} \iint |f(y)|^p\,|\varphi(x-y)|\,dy\,dx = \big(\|\varphi\|_{\mathcal{L}^1(\mathbb{R}^n)}\big)^p\,\big(\|f\|_{\mathcal{L}^p(\mathbb{R}^n)}\big)^p\,.$$ □

Es sei bemerkt, daß die Voraussetzung $\varphi \in \mathcal{L}^1(\mathbb{R}^n)$ zum Beweis genügt hätte.

Wir geben nun noch eine andere hinreichende Bedingung für die Existenz von $f * g$ an, die von großem praktischen Interesse ist. Um sie zu verstehen, argumentieren wir zunächst heuristisch: Es möge $h := f * g$ existieren. Dann ist für eine Testfunktion φ

$$(h,\varphi) = \iint f(y)g(x-y)\varphi(x)\,dy\,dx = \iint f(y)g(z)\varphi(z+y)\,dy\,dz,$$

also

$$(f*g,\varphi) = \Big(f(x)g(y),\ \psi(x,y)\Big)_{\mathbb{R}^{2n}} \qquad (*)$$

mit $\psi(x,y) := \varphi(x+y)$.

Nun ist aber ψ keine Testfunktion des $\mathbb{R}^{2n}$ mehr, vielmehr hat ψ einen Träger in der Umgebung der Geraden $y = -x$. Deshalb existiert auch $f * g$ im allgemeinen nicht. Wir erhalten aber eine hinreichende Bedingung für die Existenz von $f * g$, indem wir dafür sorgen, daß die rechte Seite von $(*)$ existiert. Es sei

$$C := \operatorname{supp} f \times \operatorname{supp} g \subset \mathbb{R}^n \times \mathbb{R}^n.$$

Dann existiert $f * g$, wenn für alle $\varphi \in \mathring{\mathcal{C}}_\infty(\mathbb{R}^n)$ und $\psi(x,y) := \varphi(x+y)$

$$C \bigcap \operatorname{supp}\psi$$

beschränkt ist. Wir wollen diese Bedingung besser verstehen: Es sei $\operatorname{supp}\varphi \subset B(0,r)$. Dann ist

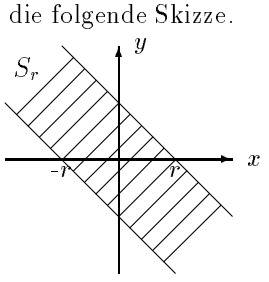
$$\operatorname{supp}\psi \subset S_r := \big\{\langle x,y\rangle \in \mathbb{R}^n \times \mathbb{R}^n \;\big|\; |x+y| < r\big\}\,.$$

Dabei ist S_r ein Streifen der Breite r. Die Faltung $f * g$ existiert also, wenn die *Streifenbedingung*

$$\forall\, r > 0 \qquad C \bigcap S_r \text{ ist beschränkt} \qquad (**)$$

erfüllt ist.

Diese Streifenbedingung ist also zum Beispiel erfüllt, wenn eine der beiden zu faltenden Funktionen beschränkten Träger hat. Man vergleiche die folgende Skizze.



6.4.2 Der Friedrichssche Glättungsoperator

Die nun folgende Glättungstechnik geht auf KURT OTTO FRIEDRICHS, 1901–1983, zurück. Es seien für $x \in \mathbb{R}^n$

$$j(x) := \begin{cases} c\,\exp\Big(-\dfrac{1}{1-|x|^2}\Big) & \text{für } |x|<1 \\ 0 & \text{sonst} \end{cases}$$

und für $\varepsilon > 0$

$$j_\varepsilon(x) := \frac{1}{\varepsilon^n} j\Big(\frac{x}{\varepsilon}\Big)\,.$$

Dabei sei $c > 0$ so gewählt, daß

$$\int_{\mathbb{R}^n} j(x)\,dx = 1$$

ist. Offenbar ist $j_\varepsilon \in \mathring{\mathcal{C}}_\infty(\mathbb{R}^n)$ mit

$$\operatorname{supp} j_\varepsilon = \overline{B(0,\varepsilon)}.$$

Es sei nun für $u \in \mathcal{L}^1_{loc}(\Omega)$, $\Omega \subset \mathbb{R}^n$ offen,

$$\big(J_\varepsilon u\big)(x) := \big(u * j_\varepsilon\big)(x) = \int_\Omega u(y) j_\varepsilon(x-y)\,dy.$$

J_ε ist der Friedrichssche Glättungsoperator, engl. *Friedrichs' Mollifier*. Er ist für alle x mit $\operatorname{dist}(x,\partial\Omega) > \varepsilon$ definiert. Ist u über beschränkte offene Teilmengen von Ω integrierbar, dann existiert $(J_\varepsilon u)(x)$ für alle $x \in \mathbb{R}^n$. Es gilt

Satz 6.4.3: *Es sei $u \in \mathcal{L}^1_{loc}(\Omega)$ über beschränkte offene Teilmengen von Ω integrierbar. Dann ist $J_\varepsilon u \in \mathcal{C}_\infty(\mathbb{R}^n)$. Gilt außerdem $\operatorname{supp} u \subset K \Subset \Omega$ und $\varepsilon < \operatorname{dist}(K,\partial\Omega)$, dann ist $J_\varepsilon u \in \mathring{\mathcal{C}}_\infty(\Omega)$.*

Beweis: Die erste Aussage folgt aus der entsprechenden Eigenschaft von j_ε, weil man wegen des Lebesgueschen Integralsatzes die Reihenfolge von Differentiation und Integration vertauschen kann. Die zweite Aussage ist klar.

Lemma 6.4.4: *Es seien $u \in \mathcal{C}_0(\Omega)$ und $K \Subset \Omega$. Dann konvergiert $J_\varepsilon u$ auf K gleichmäßig gegen u.*

Beweis: Es ist

$$\big(J_\varepsilon u\big)(x) = \frac{1}{\varepsilon^n}\int_{|x-y|<\varepsilon} j\Big(\frac{x-y}{\varepsilon}\Big) u(y)\,dy = \int_{|z|<1} j(z)\,u(x-\varepsilon z)\,dz,$$

also für $2\varepsilon < \operatorname{dist}(K,\,\partial\Omega)$

$$\sup_{x\in K}\big|u - J_\varepsilon u\big|(x) \le \sup_{x\in K}\int_{|z|<1} j(z)\,\big|u(x)-u(x-\varepsilon z)\big|\,dz \le \sup_{x\in K}\sup_{|z|\le 1}\big|u(x)-u(x-\varepsilon z)\big|.$$

Weil u in

$$\big\{x\in\Omega \;\big|\; \operatorname{dist}(x,K)\le\varepsilon\big\}$$

gleichmäßig stetig ist, folgt die Behauptung. □

Satz 6.4.5: *Es sei $u \in \mathcal{L}^p(\Omega)$, $1 \le p < \infty$. Dann gilt*

$$\|J_\varepsilon u - u\|_{\mathcal{L}^p(\Omega)} \to 0.$$

Beweis: Im Falle $\Omega \subsetneq \mathbb{R}^n$ setzen wir u durch $u = 0$ auf den $\mathbb{R}^n$ fort; wir können deshalb o.B.d.A. $\Omega = \mathbb{R}^n$ annehmen. Aus Satz 6.4.2 folgt nun zunächst

$$\|J_\varepsilon u\|_{\mathcal{L}^p(\Omega)} \le \|j_\varepsilon\|_{\mathcal{L}^1(\mathbb{R}^n)}\cdot\|u\|_{\mathcal{L}^p(\Omega)} = \|u\|_{\mathcal{L}^p(\Omega)}.$$

Zu festem $\eta > 0$ wählen wir dann ein $u_\eta \in \mathring{\mathcal{C}}_\infty(\mathbb{R}^n)$ mit

$$\|u - u_\eta\|_{\mathcal{L}^p} < \eta.$$

Dann ist

$$\|J_\varepsilon u - J_\varepsilon u_\eta\|_{\mathcal{L}^p} \le \|u - u_\eta\|_{\mathcal{L}^p} < \eta.$$

Der Träger von u_η sei in $B(0,R)$ enthalten. Dann konvergiert

$$(J_\varepsilon u_\eta)(x) \underset{\varepsilon\to 0}{\longrightarrow} u_\eta(x)$$

gleichmäßig in $B(0,2R)$ nach Lemma 6.4.4, und es ist für $\varepsilon < R$

$$\forall\, x,\ |x| > 2R, \qquad (J_\varepsilon u_\eta)(x) = u_\eta(x) = 0.$$

Mithin gilt für kleine ε

$$\|J_\varepsilon u_\eta - u_\eta\|_{\mathcal{L}^p} < \eta.$$

Daraus folgt

$$\|J_\varepsilon u - u\|_{\mathcal{L}^p} < 3\eta. \qquad \square$$

Der Friedrichssche Mollifier ist in Ω mit der Differentiation vertauschbar, es gilt nämlich

Satz 6.4.6: *Es sei $u \in \mathcal{W}_1(\Omega)$. Dann gilt für alle $x \in \Omega$ mit $\operatorname{dist}(x,\partial\Omega) > \varepsilon > 0$*

$$(\partial J_\varepsilon u)(x) = (J_\varepsilon \partial u)(x).$$

Der Beweis folgt aus

$$\begin{aligned}(\partial J_\varepsilon u)(x) = \partial \int_\Omega j_\varepsilon(x-y)\,u(y)\,dy = \int_\Omega \partial_x j_\varepsilon(x-y)\,u(y)\,dy = \\ = -\int_\Omega \big(\partial_y j_\varepsilon(x-y)\big)u(y)\,dy = \int_\Omega j_\varepsilon(x-y)\,(\partial u)(y)\,dy = (J_\varepsilon \partial u)(x).\end{aligned}$$

$\square$

Folgerung 6.4.7: *Es gilt:*

(1) Es seien $u \in \mathcal{L}^1_{loc}(\Omega)$ und

$$\forall\, \varphi \in \mathring{\mathcal{C}}_\infty(\Omega) \qquad \int_\Omega u\,\varphi = 0.$$

Dann ist $u = o$.

(2) Es seien Ω zusammenhängend, $u \in \mathcal{L}^1_{loc}(\Omega)$ und

$$\forall\, \varphi \in \mathring{\mathcal{C}}_\infty(\Omega) \qquad \int_\Omega u\,\nabla\varphi = 0.$$

Dann ist u fast überall konstant.

Beweis: Aussage (1) folgt aus der Lebesgueschen Integrationstheorie. Direkt sieht man sie so: Für jedes $K \Subset \Omega$ ist $u \in \mathcal{L}^1(K)$, und dort gilt $\|J_\varepsilon u - u\|_{\mathcal{L}^1} \to 0$ für $\varepsilon \to 0$, sowie $J_\varepsilon u = o$, also $u = o$. Zum Nachweis der zweiten Behauptung gehen wir von

$$\int_K (\nabla J_\varepsilon u)\,\varphi = -\int_K (J_\varepsilon u)\,\nabla\varphi = -\int_K u\,J_\varepsilon(\nabla\varphi) = -\int_K u\,\nabla(J_\varepsilon\varphi) = 0$$

für $\varphi \in \mathring{\mathcal{C}}_\infty(K)$ und kleine ε aus, erhalten $\nabla J_\varepsilon u = o$, also $J_\varepsilon u = const$ und damit $u = const$. $\square$

6.4.3 Dirac-Folgen

Die Resultate des letzten Abschnitts geben Anlaß, etwas allgemeiner eine Klasse von Folgen zu betrachten, die man nach Paul Dirac, 1902–1984, benennt.

Definition 6.4.8: *Es seien $\varepsilon > 0$, $\varphi \in \mathcal{L}^1(\mathbb{R}^n)$ mit $\varphi \geq 0$ und*

$$\int_{\mathbb{R}^n} \varphi = 1.$$

Dann heißt

$$\varphi_\varepsilon(x) := \frac{1}{\varepsilon^n}\varphi\Big(\frac{x}{\varepsilon}\Big)$$

Dirac-Folge *zu φ bezüglich $\varepsilon \to 0$.*

Es gilt also wieder

$$\int_{\mathbb{R}^n} \varphi_\varepsilon = 1$$

und für festes $r > 0$

$$\int_{B(0,r)} \varphi_\varepsilon = \int_{B(0,\frac{r}{\varepsilon})} \varphi \underset{\varepsilon \to 0}{\longrightarrow} 1.$$

Es folgt

Satz 6.4.9: *Es seien φ_ε eine Dirac-Folge und $f \in \mathcal{L}^p(\mathbb{R}^n)$ mit $1 \le p < \infty$. Dann gilt für $\varepsilon \to 0$*

$$\|f * \varphi_\varepsilon - f\|_{\mathcal{L}^p} \to 0.$$

Beweis: Es ist

$$\big(f * \varphi_\varepsilon\big)(x) - f(x) = \int_{\mathbb{R}^n} \big(f(y) - f(x)\big)\varphi_\varepsilon(x-y)\, dy,$$

und es folgt wie im Beweis zu Satz 6.4.2

$$\int \Big|\big(f * \varphi_\varepsilon\big)(x) - f(x)\Big|^p dx \le \Big(\int \varphi_\varepsilon\Big)^{\frac{p}{q}} \cdot \iint |f(y) - f(x)|^p \, |\varphi_\varepsilon(x-y)|\, dy\, dx \le$$

$$\le \iint |f(x-h) - f(x)|^p \, |\varphi_\varepsilon(h)|\, dh\, dx.$$

Es sei nun mit $\delta > 0$

$$\varphi_{\varepsilon\delta}(h) := \chi_{B(o,\delta)}(h)\, \varphi_\varepsilon(h) \quad \text{und} \quad \psi_{\varepsilon\delta}(h) := \chi_{\mathbb{R}^n \setminus B(o,\delta)}(h)\, \varphi_\varepsilon(h).$$

Dabei ist χ_G die charakteristische Funktion von G. Sie hat auf G den Wert Eins und verschwindet sonst. Damit folgt

$$\|f * \varphi_\varepsilon - f\|_{\mathcal{L}^p} \le \|\varphi_{\varepsilon\delta}\|_{\mathcal{L}^1}^{1/p} \cdot \sup_{|h|<\delta} \|f(\cdot + h) - f\|_{\mathcal{L}^p} + \|\psi_{\varepsilon\delta}\|_{\mathcal{L}^1}^{1/p} \cdot \sup_{h \in \mathbb{R}^n} \|f(\cdot + h) - f\|_{\mathcal{L}^p}.$$

Nun ist

$$\|\varphi_{\varepsilon\delta}\|_{\mathcal{L}^1} \le 1 \quad \text{und} \quad \sup_{h \in \mathbb{R}^n} \|f(\cdot + h) - f\|_{\mathcal{L}^p} \le 2\|f\|_{\mathcal{L}^p},$$

sowie nach Satz 6.2.1

$$\lim_{\delta \to 0} \sup_{|h|<\delta} \|f(\cdot + h) - f\|_{\mathcal{L}^p} = 0$$

und für festes δ

$$\lim_{\varepsilon \to 0} \|\psi_{\varepsilon\delta}\|_{\mathcal{L}^1} = 0.$$

Daraus folgt die Behauptung. □

Ist $f \in \mathring{\mathcal{C}}(\mathbb{R}^n)$, dann ist der Satz auch für $p = \infty$ richtig. Im allgemeinen ist er jedoch für $p = \infty$ falsch. Das folgende Beispiel soll darauf hinweisen. Es seien in $\mathbb{R}^1$

$$f(x) := \begin{cases} 1 & \text{für } x \ge 0 \\ -1 & \text{für } x < 0 \end{cases}$$

und $\varphi_\varepsilon(-x) = \varphi_\varepsilon(x)$. Dann gilt

$$\big(f * \varphi_\varepsilon\big)(x) \to f(x) \text{ für } x \ne 0$$

und

$$\big(f * \varphi_\varepsilon\big)(0) = 0,$$

also

$$\lim_{\varepsilon \to 0} \sup_{x \in \mathbb{R}^1} \Big|\big(f * \varphi_\varepsilon\big)(x) - f(x)\Big| = 1$$

6.4.4 Die Fouriertransformation

Sicherlich haben Sie im Rahmen einer Vorlesung über Infinitesimalrechnung Fourierreihen kennengelernt. Sie lassen sich als Entwicklung einer Funktion nach den Eigenfunktionen des Differentialoperators

$$u \longmapsto -\mathrm{i}u'$$

in $(-\pi,\pi)$ deuten. Die Fouriertransformation in $\mathbb{R}^1$ bzw. $\mathbb{R}^n$ spielt nun eine ähnliche Rolle wie die Fourierreihe in $(-\pi,\pi)$ bzw. $(-\pi,\pi)^n$. Auch sie läßt sich direkt einführen oder als Entwicklung nach den verallgemeinerten Eigenfunktionen des entsprechenden Differentialoperators deuten. In dieser Vorlesung sollen beide Wege eingeschlagen werden, wobei der zweite, also die Herleitung aus dem Spektralsatz, der interessantere zu sein scheint.

In diesem Abschnitt soll die Fouriertransformation direkt eingeführt werden. Weil das Thema noch einmal aufgegriffen wird, können wir uns kurz fassen.

Es seien $f \in \mathcal{L}^1(\mathbb{R}^n)$ und $k_n := (2\pi)^{-n/2}$. Dann definieren wir

$$\big(Ff\big)(\xi) := \hat{f}(\xi) := k_n \int_{\mathbb{R}^n} \mathrm{e}^{-\mathrm{i}\xi x} f(x)\,dx.$$

Man nennt F die *Fouriertransformation* und $\hat{f}$ die *Fouriertransformierte* von f. Offenbar existiert $\hat{f}$, und es gilt

$$|\hat{f}(\xi)| \le k_n \|f\|_{\mathcal{L}^1}.$$

Aus dem Lebesgueschen Satz folgt auch $\hat{f} \in \mathcal{C}(\mathbb{R}^n)$, es gilt also

$$F \;:\; \mathcal{L}^1(\mathbb{R}^n) \longrightarrow \mathcal{L}^\infty(\mathbb{R}^n) \cap \mathcal{C}(\mathbb{R}^n).$$

Wir geben zwei Beispiele:

1. Es sei für $x \in \mathbb{R}^1$

$$f(x) = \begin{cases} 1 & \text{für } |x| \le c \\ 0 & \text{für } |x| > c. \end{cases}$$

Dann ist

$$\hat{f}(\xi) = \sqrt{2/\pi}\,\frac{\sin c\xi}{\xi}.$$

2. Es ist

$$F\big(\mathrm{e}^{-(\,\cdot\,)^2}\big)(\xi) = \frac{\mathrm{e}^{-\xi^2/4}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{-\left(x+\mathrm{i}\frac{\xi}{2}\right)^2} dx = \frac{\mathrm{e}^{-\xi^2/4}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{-x^2}\,dx = \frac{\mathrm{e}^{-\xi^2/4}}{\sqrt{2}}.$$

Die folgenden Eigenschaften der Fouriertransformation zeigt man unmittelbar:

$$F\big(f(-\,\cdot\,)\big)(\xi) = k_n \int \mathrm{e}^{-\mathrm{i}x\xi} f(-x)\,dx = \hat{f}(-\xi),$$

$$F\big(f(\,\cdot\,+c)\big)(\xi) = k_n \int \mathrm{e}^{-\mathrm{i}x\xi} f(x+c)\,dx = \mathrm{e}^{\mathrm{i}\xi c}\hat{f}(\xi)$$

und für $\lambda > 0$

$$F\big(f(\lambda\,\cdot\,)\big)(\xi) = k_n \int \mathrm{e}^{-\mathrm{i}x\xi} f(\lambda x)\,dx = \frac{1}{\lambda^n}\hat{f}\Big(\frac{\xi}{\lambda}\Big).$$

Mit $x,\xi \in \mathbb{R}^n$, $y,\eta \in \mathbb{R}^m$, $z := \langle x,y\rangle$, $\zeta := \langle \xi,\eta\rangle \in \mathbb{R}^{n+m}$ und $h(z) := f(x)\,g(y)$ ist

$$\hat{h}(\zeta) = \hat{f}(\xi)\,\hat{g}(\eta).$$

Daraus folgt für $x \in \mathbb{R}^n$

$$F\big(\mathrm{e}^{-x^2}\big)(\xi) = \sqrt{\frac{1}{2^n}}\,\mathrm{e}^{-\xi^2/4}$$

und für $a > 0$

$$F\big(\mathrm{e}^{-ax^2}\big)(\xi) = \sqrt{\frac{1}{(2a)^n}}\,\mathrm{e}^{-\xi^2/4a}.$$

Zur weiteren Diskussion der Fouriertransformation ist der Raum $\mathcal{S}$ der *schnell fallenden Funktionen*, auch *Schwartzscher Raum* genannt (nach Laurant Schwartz, *1915), besonders geeignet.

Definition 6.4.10: *Es sei*

$$\mathcal{S}(\mathbb{R}^n) := \Big\{ f \in \mathcal{C}_\infty(\mathbb{R}^n) \;\Big|\; \forall\, \alpha, \beta \in \mathbb{N}_0^n \quad p_{\alpha\beta}(f) < \infty \Big\}$$

mit

$$p_{\alpha\beta}(f) := \sup_{x \in \mathbb{R}^n} \big| x^\beta \partial^\alpha f(x) \big|.$$

Die $p_{\alpha\beta}$ sind Halbnormen auf $\mathcal{S}(\mathbb{R}^n)$. Die Topologie von $\mathcal{S}(\mathbb{R}^n)$ werde von diesen Halbnormen erzeugt, damit ist $\mathcal{S}(\mathbb{R}^n)$ ein Fréchetraum. Die Folgenkonvergenz in $\mathcal{S}$ erhält man durch

Definition 6.4.11: *Die Folge (φ_n) aus $\mathcal{S}$ ist eine $\mathcal{S}$-Nullfolge, wenn*

$$\forall\, \alpha, \beta \in \mathbb{N}_0^n \qquad \lim_{n \to \infty} p_{\alpha,\beta}(\varphi_n) = 0$$

gilt.

Das trifft genau dann zu, wenn die beiden folgenden Aussagen bestehen:

(1) Für alle $\alpha, \beta \in \mathbb{N}_0^n$ gilt

$$\exists\, c_{\alpha\beta} > 0 \quad \forall\, n \in \mathbb{N} \qquad p_{\alpha\beta}(\varphi_n) \le c_{\alpha\beta}.$$

(2) Für alle $\alpha \in \mathbb{N}_0^n$ und alle $K \Subset \mathbb{R}^n$ gilt

$$\sup_{x \in K} \big| \partial^\alpha \varphi_n(x) \big| \underset{n \to \infty}{\longrightarrow} 0.$$

Offenbar sind die Multiplikation mit x^k, also $\varphi(x) \mapsto x^k \varphi(x)$, und die k-te Ableitung ∂^k, also $\varphi \mapsto \partial^k \varphi$, stetige lineare Abbildungen in $\mathcal{S}$.

Unser Hauptresultat ist der folgende

Satz 6.4.12: *Die Fouriertransformation*

$$\begin{aligned} F \;:\; \mathcal{S}(\mathbb{R}^n) &\longrightarrow \mathcal{S}(\mathbb{R}^n), \\ f &\longmapsto \hat{f} \end{aligned}$$

bildet $\mathcal{S}(\mathbb{R}^n)$ linear, stetig und bijektiv auf sich ab. Die Umkehrabbildung hat dieselben Eigenschaften und wird durch

$$\big(F^{-1} f\big)(\xi) := \check{f}(\xi) := k_n \int_{\mathbb{R}^n} \mathrm{e}^{\mathrm{i}\xi x} f(x)\, dx$$

gegeben.

Beweis: Es ist

$$\partial^\alpha \hat{f}(\xi) = (-\mathrm{i})^{|\alpha|}\, k_n \int \mathrm{e}^{-\mathrm{i}\xi x} x^\alpha f(x)\, dx$$

und

$$\mathrm{i}^{|\beta|}\, \xi^\beta \hat{f}(\xi) = k_n \int \mathrm{e}^{-\mathrm{i}\xi x} \partial^\beta f(x)\, dx,$$

also

$$\mathrm{i}^{|\beta|+|\alpha|}\, \xi^\beta \partial^\alpha \hat{f}(\xi) = k_n \int \mathrm{e}^{-\mathrm{i}\xi x} \partial^\beta \big(x^\alpha f(x) \big)\, dx.$$

Mithin ist $\hat{f} \in \mathcal{S}(\mathbb{R}^n)$, und F ist stetig. Die Abbildung F^{-1} ist vom selben Typ wie F. Unser Satz ist deshalb bewiesen, wenn wir

$$F^{-1} F = id \quad \text{und} \quad F\, F^{-1} = id$$

zeigen können. Es ist mit $f, g \in \mathcal{S}$

$$\int g(\xi)\, \hat{f}(\xi)\, \mathrm{e}^{\mathrm{i}x\xi}\, d\xi = k_n \int g(\xi) \int \mathrm{e}^{-\mathrm{i}\xi(y-x)} f(y)\, dy\, d\xi = \int \hat{g}(y-x)\, f(y)\, dy = \int \hat{g}(z)\, f(x+z)\, dz.$$

Ersetzen wir $g(\xi)$ durch $g(\varepsilon\xi)$ mit $\varepsilon > 0$, dann wird daraus

$$\int g(\varepsilon\xi)\, \hat{f}(\xi)\, \mathrm{e}^{\mathrm{i}x\xi}\, d\xi = \frac{1}{\varepsilon^n} \int \hat{g}\left(\tfrac{z}{\varepsilon}\right) f(x+z)\, dz = \int \hat{g}(y)\, f(x+\varepsilon y)\, dy.$$

Wählen wir nun $g(x) := \mathrm{e}^{-|x|^2/2}$, dann erhalten wir für $\varepsilon \to 0$

$$\int \hat{f}(\xi)\, \mathrm{e}^{\mathrm{i}x\xi}\, d\xi = f(x) \int \hat{g}(y)\, dy.$$

Nun ist aber

$$\hat{g}(y) = \mathrm{e}^{-y^2/2} \quad \text{und} \quad \int \hat{g} = \frac{1}{k_n}\big(F\hat{g}\big)(0) = \frac{1}{k_n},$$

also

$$f(x) = k_n \int \mathrm{e}^{\mathrm{i}x\xi} \hat{f}(\xi)\, d\xi$$

oder $F^{-1}F = id$. Der Nachweis von $FF^{-1} = id$ erfolgt analog. □

Korollar 6.4.13: *Für $f, g \in \mathcal{S}(\mathbb{R}^n)$ gilt $\big(F^2 f\big)(x) = f(-x)$, $\overline{F(g)} = F^{-1}(\bar{g})$ und*

$$\int \hat{f}(\xi)\, g(\xi)\, d\xi = \int f(x)\, \hat{g}(x)\, dx, \tag{$*$}$$

$$\int f(\xi)\, \overline{g(\xi)}\, d\xi = \int \hat{f}(x)\, \overline{\hat{g}(x)}\, dx,$$

$$k_n F(f * g) = \hat{f} \cdot \hat{g},$$

$$F(f \cdot g) = k_n\, \hat{f} * \hat{g}.$$

Diese Aussagen wollen wir nicht im einzelnen nachrechnen.

Wichtig für das Folgende ist es nun, die Fouriertransformation auf größere Räume fortzusetzen. Wir verwenden dazu stetige Funktionale und begnügen uns jetzt mit Funktionen $f \in \mathcal{L}^2(\mathbb{R}^n)$. Ein stetiges lineares Funktional T auf $\mathcal{S}(\mathbb{R}^n)$ nennt man *temperierte Distribution*, $T \in \mathcal{S}'(\mathbb{R}^n) := \mathcal{CL}\big(\mathcal{S}(\mathbb{R}^n), \mathbb{K}\big)$. Distributionen werden im nächsten Abschnitt besprochen, wir wollen jetzt schon ein wenig vorgreifen. Die Fouriertransformierte $\hat{T}$ einer temperierten Distribution wird in Übereinstimmung mit Gl. $(*)$ durch

$$\forall\, \varphi \in \mathcal{S}(\mathbb{R}^n) \qquad \hat{T}\varphi := T\hat{\varphi}$$

erklärt. Damit ist auch $\hat{T} \in \mathcal{S}'(\mathbb{R}^n)$. Es seien nun $f \in \mathcal{L}^2(\mathbb{R}^n)$ und T_f die *von f erzeugte temperierte Distribution*, nämlich für $\varphi \in \mathcal{S}$

$$T_f \varphi := \int \varphi(x) f(x)\, dx.$$

Dann folgt für $\hat{T}_f$ die Abschätzung

$$\big|\hat{T}_f \varphi\big| = \big|T_f \hat{\varphi}\big| \le \|\hat{\varphi}\| \cdot \|f\| = \|\varphi\| \cdot \|f\|.$$

Mithin ist $\hat{T}_f$ ein beschränktes lineares Funktional auf $\mathcal{L}^2(\mathbb{R}^n)$, und aufgrund des Rieszschen Darstellungssatzes existiert deshalb ein $\hat{f} \in \mathcal{L}^2(\mathbb{R}^n)$ mit $\|f\| = \|\hat{f}\|$ und

$$\forall\, \varphi \in \mathcal{L}^2(\mathbb{R}^n) \qquad \hat{T}_f \varphi = \int \varphi(x) \hat{f}(x)\, dx.$$

Erst recht gilt dann

$$\forall\, \varphi \in \mathcal{S}(\mathbb{R}^n) \qquad \int_{\mathbb{R}^n} \hat{\varphi}(x) f(x)\, dx = \int_{\mathbb{R}^n} \varphi(x) \hat{f}(x)\, dx.$$

Die Funktion $\hat{f} \in \mathcal{L}^2(\mathbb{R}^n)$ nennt man die Fouriertransformierte von $f \in \mathcal{L}^2(\mathbb{R}^n)$. Es gilt

$$\hat{f}(\xi) = \underset{r\to\infty}{\mathcal{L}^2(\mathbb{R}^n)\text{-lim}}\ k_n \int_{|x|<r} \mathrm{e}^{-\mathrm{i}\xi x} f(x)\, dx.$$

Betrachten wir als Anwendungsbeispiel schließlich noch ein $u \in \mathcal{W}_1(\mathbb{R}^n)$. Es ist

$$\big(F(\partial_j u)\big)(\xi) = \mathrm{i}\xi_j \hat{u}(\xi)$$

und

$$\|u\|^2_{\mathcal{W}_1} = \int_{\mathbb{R}^n} \big(1 + |\xi|^2\big)\, |\hat{u}(\xi)|^2\, d\xi.$$

Speziell in $\mathbb{R}^1$ folgt dann aus

$$u(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathrm{e}^{i\xi x} \hat{u}(\xi)\, d\xi = \frac{1}{\sqrt{2\pi}} \Big\{ \int_{|\xi|<1} \cdots + \int_{|\xi|>1} \cdots \Big\}$$

und

$$\Big| \int_{|\xi|>1} \cdots \Big|^2 \le \int_{|\xi|>1} \frac{d\xi}{|\xi|^2} \cdot \int_{|\xi|>1} |\xi|^2 \, |\hat{u}(\xi)|^2 \, d\xi \le 2 \int_{\mathbb{R}} |\xi|^2 \, |\hat{u}(\xi)|^2 \, d\xi$$

die Abschätzung

$$|u(x)|^2 \le \frac{2}{\pi} \int_{\mathbb{R}} \big(1 + |\xi|^2\big) \, |\hat{u}(\xi)|^2 d\xi,$$

also (vgl. Definition 6.6.3)

$$|u(x)| \le \sqrt{2/\pi} \, \|u\|_{\mathcal{W}_1(\mathbb{R})}.$$

Dies ist ein erste *Regularitätsaussage*. Es gilt $\mathcal{W}_1(\mathbb{R}) \subset \mathcal{L}^\infty(\mathbb{R})$, und aus

$$\big|u(x+h) - u(x)\big|^2 \le \frac{2}{\pi} \int_{\mathbb{R}} \big(1 + |\xi|^2\big) \, \big|\big(\mathrm{e}^{\mathrm{i}\xi h} - 1\big)\hat{u}(\xi)\big|^2 d\xi$$

und dem Lebesgueschen Satz folgt sogar die Stetigkeit von u, also

$$\mathcal{W}_1(\mathbb{R}) \hookrightarrow \mathcal{BC}(\mathbb{R}).$$

6.5 Distributionen

In diesem Abschnitt wollen wir eine weitere Klasse verallgemeinerter Funktionen behandeln, nämlich die *Distributionen*. Die Distributionen spielen in den Anwendungen eine große Rolle. So verwendet man sie, um Lösungen von Differentialgleichungen zu beschreiben.

Es sei zunächst K ein Kompaktum mit $K \Subset \Omega \subset \mathbb{R}^n$, Ω offen, und

$$\mathcal{D}_K(\Omega) := \Big\{ f \in \mathring{\mathcal{C}}_\infty(\Omega) \;\Big|\; \operatorname{supp} f \subset K \Big\}.$$

Die Topologie von $\mathcal{D}_K(\Omega)$ wird von der Familie der Halbnormen

$$p_{m,K}(f) := \sup_{|\alpha| \le m} \sup_{x \in K} |\partial^\alpha f(x)|$$

erzeugt. Auf diese Weise wird $\mathcal{D}_K(\Omega)$ zu einem topologischen Raum.

Wir sind aber nicht an $\mathcal{D}_K(\Omega)$ interessiert, sondern an dem Raum der Testfunktionen $\mathring{\mathcal{C}}_\infty(\Omega)$ und wollen diesen topologisieren. Dazu kann man folgendermaßen vorgehen: Es seien $K_1 \subset K_2$ und die Topologie von $\mathcal{D}_{K_1}(\Omega)$ gleich der relativen Topologie von $\mathcal{D}_{K_1}(\Omega)$ als Teilraum von $\mathcal{D}_{K_2}(\Omega)$. Man wählt

$$\mathcal{D}(\Omega) := \bigcup_{K \Subset \Omega} \mathcal{D}_K(\Omega) = \mathring{\mathcal{C}}_\infty(\Omega)$$

und definiert die Topologie von $\mathcal{D}(\Omega)$ als *induktiven Limes* der $\mathcal{D}_K(\Omega)$. Dieser Raum $\mathcal{D}(\Omega)$ ist nicht mehr quasinormiert. Im wesentlichen besagt das

$$O \subset \mathcal{D}(\Omega) \quad \text{ist offen} \quad :\Longleftrightarrow \quad \forall\, K \Subset \Omega \quad O \cap \mathcal{D}_K(\Omega) \;\text{ ist offen in } \mathcal{D}_K(\Omega).$$

Auf weitere Einzelheiten soll an dieser Stelle nicht eingegangen werden. Es sei zum Beispiel auf *Yosida [1965], S. 27f* verwiesen.

Vielmehr wollen wir hier einfacher vorgehen und nur den Begriff der Folgenkonvergenz benutzen.

Definition 6.5.1: *Für $j \in \mathbb{N}$ sei $\varphi_j \in \mathring{\mathcal{C}}_\infty(\Omega)$. Dann heißt (φ_j) eine $\mathcal{D}$-*Nullfolge*, wenn es ein Kompaktum $K \Subset \Omega$ gibt mit:*

(1) Für alle $j \in \mathbb{N}$ ist $\operatorname{supp} \varphi_j \subset K$.

(2) Es gilt

$$\forall\, \alpha \in \mathbb{N}_0^n \qquad \sup_{x \in \Omega} |\partial^\alpha \varphi_j(x)| \underset{j \to \infty}{\longrightarrow} 0.$$

Es sei betont, daß die letzte Aussage nicht gleichmäßig bezüglich α gelten soll.

Es sei von nun an

$$\mathcal{D}(\Omega) = \mathring{\mathcal{C}}_\infty(\Omega),$$

ausgestattet mit dieser $\mathcal{D}$-Folgenkonvergenz.

Definition 6.5.2: *Ein lineares Funktional auf $\mathcal{D}(\Omega)$, das $\mathcal{D}$-Nullfolgen in Nullfolgen abbildet, heißt* $\mathcal{D}$-stetig.

Nun können wir Distributionen definieren.

Definition 6.5.3: *Eine* Distribution *ist ein $\mathcal{D}$-stetiges lineares Funktional auf $\mathcal{D}(\Omega)$. Es sei $\mathcal{D}'(\Omega)$ der Raum der Distributionen auf $\mathcal{D}(\Omega)$.*

Im Falle $\Omega = \mathbb{R}^n$ schreiben wir kurz $\mathcal{D}$ für $\mathcal{D}(\mathbb{R}^n)$ und $\mathcal{D}'$ für $\mathcal{D}'(\mathbb{R}^n)$. Es gilt

Satz 6.5.4: *Es ist* $T \in \mathcal{D}'(\Omega) \iff$

$$\forall\, K \Subset \Omega \quad \exists\, a_K > 0 \quad \exists\, m \in \mathbb{N}_0 \quad \forall\, \varphi \in \mathcal{D}_K(\Omega) \qquad |T\varphi| \le a_K\, p_{m,K}(\varphi).$$

Beweis:

$\Leftarrow$: Diese Richtung ist trivial. Es sei (φ_j) eine $\mathcal{D}$-Nullfolge. Dann gibt es ein K mit $\varphi_j \in \mathcal{D}_K(\Omega)$ und $p_{m,K}(\varphi_j) \to 0$.

$\Rightarrow$: Wir beweisen indirekt. Es gelte also

$$\exists\, K \Subset \Omega \quad \forall\, j \in \mathbb{N} \quad \exists\, \varphi_j \in \mathcal{D}_K(\Omega) \qquad |T\varphi_j| > j\, p_{j,K}(\varphi_j).$$

Dann ist

$$\psi_j := \frac{\varphi_j}{j\, p_{j,K}(\varphi_j)} \in \mathcal{D}_K(\Omega)$$

eine $\mathcal{D}$-Nullfolge und $|T\psi_j| > 1$. □

Bevor wir weitermachen, geben wir erst einmal eine wichtige Klasse von Distributionen an.

Satz 6.5.5: *Es seien $u \in \mathcal{L}^1_{loc}(\Omega)$ und*

$$\begin{aligned}[u] \;:\; \mathcal{D}(\Omega) &\longrightarrow \mathbb{K},\\ \varphi &\longmapsto \int_\Omega u(x)\varphi(x)dx.\end{aligned}$$

Dann ist $[u] \in \mathcal{D}'(\Omega)$.

Man nennt $[u]$ die *von u erzeugte Distribution*. Im folgenden werden wir meist in der Notation nicht mehr zwischen u und $[u]$ unterscheiden, also die eckigen Klammern fortlassen. Aus dem Zusammenhang wird ersichtlich sein, ob es sich um eine Distribution oder Funktion handeln soll.

Beweis des Satzes: Die Linearität von $[u]$ ist klar, wir zeigen die $\mathcal{D}$-Stetigkeit. Es sei also (φ_j) eine $\mathcal{D}$-Nullfolge. Dann gilt für ein $K \Subset \Omega$

$$\left|[u]\varphi_j\right| = \left|\int_K u(x)\,\varphi_j(x)\,dx\right| \le \|u\|_{\mathcal{L}^1(K)}\, p_{0,K}(\varphi_j) \to 0.$$ □

Wir fassen die Bezeichnungen noch einmal zusammen.

Definition 6.5.6:

(1) Es sei $u \in \mathcal{L}^1_{loc}(\Omega)$. Dann heißt $[u]$ die von u erzeugte Distribution.

(2) Zu $T \in \mathcal{D}'(\Omega)$ gebe es ein $u \in \mathcal{L}^1_{loc}(\Omega)$ mit $T = [u]$. Dann heißt T reguläre Distribution.

Es folgen nun einige Bemerkungen:

1. Der Raum $\mathcal{L}^1_{loc}(\Omega)$ ist zum Unterraum der regulären Distributionen isomorph. Aus $[u] = [v]$ folgt nämlich

$$\forall\, \varphi \in \mathring{\mathcal{C}}_\infty(\Omega) \qquad \int_\Omega (u - v)\,\varphi = 0$$

und daraus $u = v$.

2. Es gibt nichtreguläre Distributionen. Zum Beispiel sei mit $x \in \Omega$

$$\begin{aligned} \delta_x \quad : \quad & \mathcal{D}(\Omega) \longrightarrow \mathbb{K}, \\ & \varphi \longmapsto \varphi(x). \end{aligned}$$

Das ist die *Dirac-Distribution* zu $x \in \Omega$. Offensichtlich ist δ_x eine Distribution. Es sei $x = o$, und es gebe ein $d \in \mathcal{L}^1_{loc}(\Omega)$ mit

$$\delta_o \varphi = \int_{\mathbb{R}^n} d(x)\, \varphi(x)\, dx.$$

Aus $d \in \mathcal{L}^1_{loc}(\Omega)$ folgt dann für ein $\varepsilon > 0$

$$\int_{|x|<\varepsilon} |d(x)| < 1.$$

Wir wählen nun ein $\varphi \in \mathring{\mathcal{C}}_\infty(\Omega)$ mit $\operatorname{supp}\varphi \subset B(0,\varepsilon)$ und $0 \le \varphi \le \varphi(o) \ne 0$. Dann folgt

$$\varphi(o) = \delta_o \varphi = \int_{\mathbb{R}^n} d\,\varphi \le \varphi(o) \int_{|x|<\varepsilon} |d| < \varphi(o).$$

Mithin ist δ_0 keine reguläre Distribution.

3. Symbolisch schreiben wir auch

$$\begin{aligned} \delta_o \varphi =: \delta\varphi =: \int_{\mathbb{R}^n} \delta(x)\varphi(x)\, dx = \varphi(o) \\ \delta_x \varphi =: \int_{\mathbb{R}^n} \delta(y-x)\varphi(y)\, dy = (\varphi * \delta)(x) = \varphi(x). \end{aligned}$$

Dabei sei $\delta(x)$ die *Dirac-Funktion*. Sie ist keine Funktion im üblichen Sinne mehr, läßt sich aber als Grenzwert einer Dirac-Folge deuten.

4. Eine zweite Distribution, die häufig auftritt, ist die *Heaviside-Distribution*, benannt nach Oliver Heaviside, 1850–1925, nämlich

$$H\varphi := H_o\varphi := \int_{\mathbb{R}} \chi_{[0,\infty)}(x)\, \varphi(x)\, dx$$

und analog

$$H_x\varphi := \int_{\mathbb{R}} \chi_{[x,\infty)}(y)\, \varphi(y)\, dy.$$

H ist natürlich eine reguläre Distribution.

Schließlich definieren wir noch die Ordnung einer Distribution.

Definition 6.5.7: *Eine Distribution $T \in \mathcal{D}'(\Omega)$ heißt von* endlicher Ordnung, *wenn folgendes gilt*

$$\exists\, m \in \mathbb{N}_0 \quad \forall\, K \Subset \Omega \quad \exists\, a_K > 0 \quad \forall\, \varphi \in \mathcal{D}_K(\Omega) \qquad |T\varphi| \le a_K\, p_{m,K}(\varphi).$$

Ist T von endlicher Ordnung, dann heißt die kleinstmögliche Zahl m die Ordnung *von T.*

Zum Beispiel hat δ_x die Ordnung Null. Jede reguläre Distribution hat ebenfalls die Ordnung Null. Das folgt aus

$$\big|[u]\varphi\big| \le \|u\|_{\mathcal{L}^1(K)}\, p_{0,K}(\varphi)$$

mit $K := \operatorname{supp}\varphi$.

Nun kommen wir aber zum wesentlichen Inhalt der Distributionentheorie und zeigen, daß Distributionen beliebig oft differenzierbar sind.

Satz 6.5.8: *Es seien $T \in \mathcal{D}'(\Omega)$ und*

$$\begin{aligned} \partial^\alpha T \quad : \quad & \mathcal{D}(\Omega) \longrightarrow \mathbb{K}, \\ & \varphi \longmapsto (-1)^{|\alpha|}\, T(\partial^\alpha \varphi). \end{aligned}$$

Dann ist $\partial^\alpha T \in \mathcal{D}'(\Omega)$. Die Distribution $\partial^\alpha T$ heißt α-te Ableitung von T.

Beweis: Mit φ ist auch $\psi := \partial^\alpha \varphi \in \mathcal{D}(\Omega)$ und ∂^α ist eine stetige Abbildung in $\mathcal{D}(\Omega)$. Mithin ist $\partial^\alpha T \in \mathcal{D}'(\Omega)$. □

Damit ist jede Distribution beliebig oft differenzierbar. Man spricht von der *Ableitung im Distributionensinn*. Wird T von einer glatten Funktion erzeugt, dann stimmen die Ableitungsbegriffe überein.

Berechnen wir die Ableitung von H: Es ist

$$H'\varphi = -H\varphi' = -\int_0^\infty \varphi'(x)dx = \varphi(0) = \delta\varphi,$$

also ist $H' = \delta$.

Zum besseren Arbeiten mit Distributionen wollen wir noch ihren Träger definieren. Dazu erklären wir zunächst, wann eine Distribution auf einer offenen Menge verschwindet.

Definition 6.5.9: *Es seien $\Omega_0 \subset \Omega$ offen und $T \in \mathcal{D}'(\Omega)$. Dann verschwindet T auf Ω_0, wenn*

$$\forall\, \varphi \in \mathring{\mathcal{C}}_\infty(\Omega_0) \qquad T\varphi = 0$$

gilt.

Lemma 6.5.10: *Es seien Λ eine beliebige Indexmenge und $\Omega_\lambda \subset \Omega$ für $\lambda \in \Lambda$ offen. $T \in \mathcal{D}'(\Omega)$ verschwinde auf allen Ω_λ. Dann verschwindet T auf*

$$\Omega_0 := \bigcup_{\lambda\in\Lambda} \Omega_\lambda.$$

Beweis: Es seien $\varphi \in \mathring{\mathcal{C}}_\infty(\Omega_0)$ und $K := \operatorname{supp}\varphi \subset \Omega_0$. Weil K kompakt ist, gibt es $\Omega_1, \ldots, \Omega_n$ mit

$$K \subset \bigcup_{i=1}^n \Omega_i \quad \text{und } \Omega_i \in \left\{\Omega_\lambda \;\middle|\; \lambda \in \Lambda\right\}.$$

Auf K gibt es nun eine *Partition der Eins* zu $\Omega_1, \ldots, \Omega_n$, nämlich Funktionen $\eta_i \in \mathring{\mathcal{C}}_\infty(\Omega_i)$ mit $\eta_i \geq 0$ und

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

Man vergleiche etwa *Barner & Flohr [1982], S. 394* oder Definition 6.6.14 und die anschließende Konstruktion. Damit ist dann

$$T\varphi = \sum_{i=1}^n T(\varphi\eta_i) = 0.$$ □

Definition 6.5.11: *Es sei Ω_0 die Vereinigung aller offenen Mengen, auf denen $T \in \mathcal{D}'(\Omega)$ verschwindet. Dann heißt*

$$\operatorname{supp} T := \Omega\backslash\Omega_0$$

der Träger *von T.*

Zum Beispiel ist

$$\operatorname{supp}\delta = \{o\} \quad \text{und} \quad \operatorname{supp} H = [0, \infty).$$

Für $u \in \mathcal{C}(\Omega)$ ist $\operatorname{supp}[u] = \operatorname{supp} u$.

Satz 6.5.12: *Es sei $T \in \mathcal{D}'(\Omega)$ eine Distribution mit kompaktem Träger $K \Subset \Omega$. Dann hat T endliche Ordnung.*

Beweis: Es sei mit $\delta := \operatorname{dist}(K, \mathbb{R}^n\backslash\Omega) > 0$

$$K_\delta := \left\{x \in \Omega \;\middle|\; \operatorname{dist}(x, K) \leq \tfrac{\delta}{2}\right\}.$$

Dann gibt es nach §6.4.2 ein $\eta \in \mathring{\mathcal{C}}_\infty(K_\delta)$ mit $\eta \geq 0$ und $\eta|K = 1$, und für $\varphi \in \mathcal{D}(\Omega)$ ist

$$\varphi = \eta\,\varphi + (1-\eta)\varphi$$

mit

$$\operatorname{supp}\eta\,\varphi \subset \operatorname{supp}\eta =: K_1 \subset K_\delta \quad \text{und} \quad \operatorname{supp}(1-\eta)\varphi \subset \Omega\backslash K.$$

Daraus folgt

$$T\varphi = T(\eta\varphi),$$

und es gilt nach Satz 6.5.4

$$\exists\, a_{K_1} > 0 \quad \exists\, m > 0 \quad \forall\, \varphi \in \mathcal{D}(\Omega) \qquad |T\varphi| \le a_{K_1}\, p_{m,K_1}(\eta\varphi)$$

oder für $\varphi \in \mathcal{D}_L(\Omega)$ mit $L \Subset \Omega$

$$\exists\, m > 0 \quad \forall\, L \Subset \Omega \quad \exists\, a_L > 0 \quad \forall\, \varphi \in \mathcal{D}_L(\Omega) \qquad |T\varphi| \le a_L\, p_{m,L}(\varphi). \qquad \square$$

Zum Schluß wollen wir noch Klassen spezieller Distributionen angeben, so auch die schon in §6.4.4 aufgetretenen temperierten Distributionen.

Zur Beantwortung spezieller Fragen ist es oft nützlich, die Menge der Distributionen $\mathcal{D}'(\Omega)$ kleiner zu wählen, um so eventuell schärfere Aussagen, zum Beispiel über das Verhalten der Lösungen von Differentialgleichungen, zu gewinnen. Für die Anwendungen ist also $\mathcal{D}'(\Omega)$ oft zu groß. Dazu muß man dann die Klasse der Testfunktionen vergrößern.

Zunächst kann man daran denken, statt mit Funktionen $\varphi \in \mathring{\mathcal{C}}_\infty(\Omega)$ mit Funktionen aus $\mathring{\mathcal{C}}_m(\Omega)$ zu testen und so Distributionen $\mathcal{D}'_m(\Omega)$ zu definieren. Das sind Räume von Distributionen endlicher Ordnung. Größere Bedeutung haben aber in der Praxis die temperierten Distributionen und die Distributionen mit kompaktem Träger gewonnen. Wir wollen sie kurz einführen, beschränken uns aber im folgenden auf $\mathbb{R}^n$ und wählen zwei Räume $\mathcal{S}$ und $\mathcal{E}$ mit

$$\mathcal{D} \subset \mathcal{S} \subset \mathcal{E}.$$

Dabei ist $\mathcal{S}$ der uns schon bekannte Raum der schnell fallenden Funktionen, auch *Schwartzscher Raum* gennannt, und $\mathcal{E} := \mathcal{C}_\infty(\mathbb{R}^n)$, also

$$\mathcal{S} := \left\{ f \in \mathcal{E} \;\middle|\; \forall\, \alpha, \beta \in \mathbb{N}_0^n \quad p_{\alpha\beta}(f) < \infty \right\}$$

mit

$$p_{\alpha\beta}(f) := \sup_{x \in \mathbb{R}^n} \left| x^\beta \partial^\alpha f(x) \right|.$$

Die Folgenkonvergenz in $\mathcal{E}$ wird erklärt in

Definition 6.5.13: *Die Folge (φ_j) aus $\mathcal{E}$ ist eine $\mathcal{E}$-Nullfolge, wenn*

$$\forall\, \alpha \in \mathbb{N}_0^n \quad \forall\, K \Subset \mathbb{R}^n \qquad \sup_{x \in K} |\partial^\alpha \varphi_j(x)| \underset{j\to\infty}{\longrightarrow} 0$$

gilt.

Analog zu $\mathcal{D}'$ erhält man dann die Räume $\mathcal{S}'$ und $\mathcal{E}'$ der $\mathcal{S}$- bzw. $\mathcal{E}$-stetigen linearen Funktionale, und es gilt

$$\mathcal{E}' \subset \mathcal{S}' \subset \mathcal{D}'.$$

Der Raum $\mathcal{S}'$ ist der Raum der *temperierten Distributionen*, und $\mathcal{E}'$ ist der Raum der *Distributionen mit kompaktem Träger*.

6.6 Sobolevräume

Wir behandeln nun Klassen von Funktionen, die in den Anwendungen besonders häufig auftreten. Sie sind nach Sergei Lvovich Sobolev, 1908–1989, benannt. Solche *Sobolevräume* sind uns bereits in der Einführung und in §5.4 begegnet. Über sie läßt sich vieles sagen, bezüglich weiterer Einzelheiten sei auf *Adams [1975]*, verwiesen.

Wie in §5.4 seien $(\cdot,\cdot)$ das $\mathcal{L}^2(\Omega)$-Skalarprodukt und für $k \in \mathbb{N}_0$

$$(u,v)_k := \sum_{0 \le j \le k} \sum_{|\alpha|=j} (\partial^\alpha u, \partial^\alpha v)$$

sowie

$$|u|_k := \sqrt{\sum_{|\alpha|=k} \|\partial^\alpha u\|^2}\,, \qquad \|u\|_k := \sqrt{(u,u)_k}.$$

Definition 6.6.1: *Es seien $\alpha \in \mathbb{N}_0^n$ und $u \in \mathcal{L}^1_{loc}(\Omega)$. Dann heißt $v \in \mathcal{L}^1_{loc}(\Omega)$* schwache α-te Ableitung *von u, wenn $\partial^\alpha[u]$ von v erzeugt wird.*

Es soll also

$$\forall\, \varphi \in \mathcal{D}(\Omega) \qquad \int_\Omega v\,\varphi = (-1)^\alpha \int_\Omega u\,\partial^\alpha \varphi$$

sein.

Definition 6.6.2: *Es sei $m \in \mathbb{N}_0$. Dann heißt $u \in \mathcal{L}^1_{loc}(\Omega)$ m-mal* schwach differenzierbar, *wenn es $v_\alpha \in \mathcal{L}^1_{loc}(\Omega)$ gibt mit*

$$\forall\, \alpha \in \mathbb{N}_0^n,\ |\alpha| \le m, \qquad v_\alpha = \partial^\alpha u \in \mathcal{D}'(\Omega).$$

Eine m-mal klassisch differenzierbare Funktion ist auch m-mal schwach differenzierbar, und beide Ableitungen stimmen überein.

Definition 6.6.3: *Es sei $1 \le p \le \infty$. Dann ist $\mathcal{W}^p_m(\Omega)$ der Raum der m-mal schwach differenzierbaren Funktionen $u \in \mathcal{L}^p(\Omega)$ mit*

$$\forall\, \alpha \in \mathbb{N}_0^n,\ |\alpha| \le m, \qquad \partial^\alpha u \in \mathcal{L}^p(\Omega).$$

Mit der Norm

$$\|u\|_{m,p,\Omega} := \begin{cases} \sum_{|\alpha|\le m} \|\partial^\alpha u\|_{\mathcal{L}^p(\Omega)} & \text{für } p < \infty \\ \max_{|\alpha|\le m} \|\partial^\alpha u\|_{\infty,\Omega} & \text{für } p = \infty \end{cases}$$

wird $\mathcal{W}^p_m(\Omega)$ zu einem normierten Vektorraum.

Im Falle $p = 2$ schreiben wir kurz $\mathcal{W}_m(\Omega)$ statt $\mathcal{W}^2_m(\Omega)$ und verwenden das Skalarprodukt

$$(u,v)_{m,\Omega} := \sum_{|\alpha|\le m} \int_\Omega \partial^\alpha u\, \overline{\partial^\alpha v}.$$

Beispiel 6.6.4: *In $\mathbb{R}^3$ sei*

$$g(x) := \begin{cases} \frac{1}{4\pi|x|} & \text{für } x \ne o \\ 0 & \text{für } x = o. \end{cases}$$

Dann ist g einmal schwach differenzierbar, jedoch nicht zweimal.

Beweis: Es ist $g \in \mathcal{L}^1_{loc}(\mathbb{R}^3)$, und es gilt für $\varphi \in \mathring{\mathcal{C}}_\infty(\mathbb{R}^3)$

$$\int g\,(\partial_i \varphi) = -\int (\partial_i g)\,\varphi$$

mit

$$(\partial_i g)(x) := -\frac{x_i}{4\pi|x|^3} \quad \text{für } x \ne o.$$

Auch $\partial_i g$ gehört zu $\mathcal{L}^1_{loc}(\mathbb{R}^3)$. Wir berechnen nun die zweiten Ableitungen. Es ist

$$\int_{\mathbb{R}^3} (\partial_i\partial_j\varphi) g = \frac{1}{4\pi}\int_{\mathbb{R}^3} (\partial_j\varphi)(x)\,\frac{x_i}{|x|^3}dx = \frac{1}{4\pi}\Big\{\int_{|x|<\varepsilon} \cdots + \int_{|x|>\varepsilon} \cdots\Big\},$$

und aus

$$\partial_j \frac{x_i}{|x|^3} = \frac{\delta_{ji}}{|x|^3} - 3\frac{x_i x_j}{|x|^5}$$

folgt für $x \ne o$

$$\sum_{i=1}^3 \partial_i \frac{x_i}{|x|^3} = 0.$$

Damit erhalten wir

$$\sum_{i=1}^3 \int_{|x|>\varepsilon} (\partial_i\varphi)(x)\frac{x_i}{|x|^3}dx = -\int_{|x|=\varepsilon} \varphi(x)\frac{dx}{|x|^2},$$

also für $\varepsilon \to 0$

$$\sum_{i=1}^3 \int_{\mathbb{R}^3} (\partial_i^2\varphi)\, g = -\varphi(o).$$

Mithin ist $-\Delta g = \delta \notin \mathcal{L}^1_{loc}(\mathbb{R}^3)$. □

Als erstes zeigen wir nun

Satz 6.6.5: *$\mathcal{W}^p_m(\Omega)$ ist vollständig.*

Beweis: Es sei (u_n) eine $\mathcal{W}^p_m$-Cauchyfolge. Dann ist $(\partial^\alpha u_n)$ für jedes $\alpha \in \mathbb{N}_0^n$ mit $|\alpha| \le m$ eine $\mathcal{L}^p$-Cauchyfolge. Der Raum $\mathcal{L}^p(\Omega)$ ist vollständig; es gibt also $v_\alpha \in \mathcal{L}^p(\Omega)$ mit

$$\partial^\alpha u_n \to v_\alpha \quad \text{in} \quad \mathcal{L}^p(\Omega).$$

Mithin gilt für alle Testfunktionen $\varphi \in \mathring{\mathcal{C}}_\infty(\Omega)$

$$\int v_0\, \partial^\alpha \varphi \leftarrow \int u_n\, \partial^\alpha \varphi = (-1)^{|\alpha|} \int (\partial^\alpha u_n)\, \varphi \to (-1)^{|\alpha|} \int v_\alpha \varphi.$$

Es ist also $v_\alpha = \partial^\alpha v_0$ die schwache α-te Ableitung von $v_0 \in \mathcal{L}^p(\Omega)$. Deshalb gilt $v_0 \in \mathcal{W}^p_m(\Omega)$ und

$$\|u_n - v_0\|_{\mathcal{W}^p_m(\Omega)} \to 0. \qquad \square$$

Damit haben wir die Sobolevräume $\mathcal{W}^p_m(\Omega)$ der im schwachen Sinne differenzierbaren Funktionen eingeführt. Neben den *schwachen* Ableitungen, die man also durch *Testen* erhält, benötigt man aber auch *starke* Ableitungen, die durch *Approximieren* gewonnen werden sollen. Um sie zu definieren, bilden wir wie in §5.4

$$\mathcal{C}^{p*}_k(\Omega) := \left\{ u \in \mathcal{C}_k(\Omega) \;\middle|\; \forall\, \alpha, |\alpha| \le k,\ \partial^\alpha u \in \mathcal{L}^p(\Omega) \right\},$$

also $\mathcal{C}^{p*}_k(\Omega) = \mathcal{C}_k(\Omega) \cap \mathcal{W}^p_k(\Omega)$.

Definition 6.6.6: *Der Raum der m-mal* stark differenzierbaren *Funktionen ist*

$$\mathcal{H}^p_m(\Omega) := \left\{ \mathcal{C}^{p*}_m(\Omega);\ \|\cdot\|_{\mathcal{W}^p_m(\Omega)} \right\}^\sim.$$

Im Falle $p = 2$ schreiben wir wieder kurz $\mathcal{H}_m(\Omega)$ für $\mathcal{H}^2_m(\Omega)$. Wichtig ist auch der Teilraum

$$\mathring{\mathcal{H}}^p_m(\Omega) := \left\{ \mathring{\mathcal{C}}_\infty(\Omega);\ \|\cdot\|_{\mathcal{W}^p_m(\Omega)} \right\}^\sim.$$

Es gilt also

$$\mathring{\mathcal{H}}^p_m(\Omega) \subset \mathcal{H}^p_m(\Omega) \subset \mathcal{W}^p_m(\Omega).$$

Die Aussage, daß eine Funktion u mitsamt ihren $(m-1)$-ten Ableitungen am Rande von Ω verschwindet, wird durch $u \in \mathring{\mathcal{H}}^p_m(\Omega)$ verallgemeinert.

Es sei nämlich $\partial\Omega$ glatt und $u \in \mathring{\mathcal{H}}_1(\Omega) \cap \mathcal{C}_1(\bar\Omega)$. Dann sieht man leicht, daß u am Rande $\partial\Omega$ verschwindet: Es seien $\varphi_n \in \mathring{\mathcal{C}}_\infty(\Omega)$ mit

$$\|\varphi_n - u\|_1 \to 0$$

und $f \in \mathcal{C}_1(\bar\Omega)$. Dann ist

$$\int_\Omega (\partial_i \varphi_n)\, f = -\int_\Omega \varphi_n\, (\partial_i f),$$

und daraus folgt

$$\int_\Omega (\partial_i u)\, f = -\int_\Omega u\, (\partial_i f).$$

Andererseits erhält man aber aus dem Gaußschen Satz

$$\int_\Omega (\partial_i u)\, f = \int_{\partial\Omega} n_i\, u\, f - \int_\Omega u\, (\partial_i f).$$

Es gilt also

$$\forall\, f \in \mathcal{C}_1(\bar\Omega) \qquad \int_{\partial\Omega} n_i\, u\, f = 0,$$

und daraus folgt $u|\partial\Omega = o$. Ist $u \in \mathring{\mathcal{H}}_m(\Omega)$, dann erhält man entsprechende Aussagen für höhere Ableitungen.

Wichtig für das Folgende ist nun die Beziehung

$$\mathcal{H}^p_m(\Omega) = \mathcal{W}^p_m(\Omega),$$

„stark gleich schwach", die für $1 \leq p < \infty$ richtig ist. Sie wurde erst relativ spät entdeckt und als allgemeine Aussage von *Meyers & Serrin [1964]* bewiesen. Wir werden den Beweis anschließend geben. Im Falle $p = \infty$ ist sie falsch. In $\Omega = (-1, 1)$ ist nämlich

$$H(x) := \begin{cases} 1 & \text{für } x \geq 0 \\ 0 & \text{für } x < 0 \end{cases}$$

aus $\mathcal{L}^\infty(\Omega)$. Die Heaviside-Funktion H läßt sich aber nicht in der Supremumsnorm durch stetige Funktionen approximieren, denn sonst wäre sie in einer Umgebung des Nullpunktes stetig.

Wir zeigen nun zunächst einen wichtigen Spezialfall der Aussage „stark gleich schwach", nämlich

Satz 6.6.7: *Es ist* $\mathcal{H}_1(\Omega) = \mathcal{W}_1(\Omega)$.

Die Aussage dieses Satzes geht bereits auf FRIEDRICHS [1944] und *Kasuga [1957]* zurück.

Zum Beweis beachten wir, daß $\mathcal{H}_1(\Omega)$ ein abgeschlossener Teilraum von $\mathcal{W}_1(\Omega)$ ist, und projizieren

$$\mathcal{W}_1(\Omega) = \mathcal{H}_1(\Omega) \oplus \mathcal{H}_1(\Omega)^\perp.$$

Es sei $v \in \mathcal{H}_1(\Omega)^\perp$. Dann haben wir $v = o$ zu zeigen. Zunächst gilt

$$\forall\, f \in \mathcal{H}_1(\Omega) \qquad (v, f)_1 = 0,$$

also

$$\forall\, \varphi \in \mathring{\mathcal{C}}_\infty(\Omega) \qquad (v,\, -\Delta\varphi + \varphi) = 0.$$

Damit ist v eine schwache Lösung von

$$(-\Delta + 1)u = 0.$$

Nun gibt es eine zweite wichtige Klasse von Aussagen, auf die wir im folgenden näher eingehen werden, nämlich *Regularitätsaussagen*. Wir verwenden jetzt eine solche Aussage, nämlich

Lemma 6.6.8: *Es seien* $u \in \mathcal{L}^1_{loc}(\Omega)$ *und*

$$\forall\, \varphi \in \mathring{\mathcal{C}}_\infty(\Omega) \qquad (u,\, -\Delta\varphi + \varphi) = 0.$$

Dann gilt fast überall $u \in \mathcal{C}_\infty(\Omega)$.

Dieses Resultat stammt von *Weyl [1940]*. Man nennt es *Weylsches Lemma*, auch seine allgemeinere Formulierung für elliptische Differentialgleichungen mit variablen Koeffizienten heißt so. Benannt wird das Lemma also nach HERMANN WEYL, 1885–1955. Bevor wir es beweisen, führen wir den Beweis des letzten Satzes zu Ende: Aufgrund des Weylschen Lemmas ist $v \in \mathcal{C}_\infty(\Omega) \cap \mathcal{W}_1(\Omega)$ und damit aus $\mathcal{H}_1(\Omega)$. Mithin verschwindet v. □

Zum Beweis des Weylschen Lemmas stellen wir u mit Hilfe einer Grundlösung dar. Aus dieser Darstellung kann man dann die Regularitätsaussage ablesen. Wir beschränken uns auf eine Menge $\Omega \subset \mathbb{R}^3$, wählen zu $x_0 \in \Omega$ ein $r > 0$ mit $\bar{B}(x_0, 2r) \subset \Omega$ und ein $\psi \in \mathring{\mathcal{C}}_\infty(\Omega)$ mit

$$\psi(x) = \begin{cases} 1 & \text{für alle } x \text{ mit } |x - x_0| \leq r \\ 0 & \text{für alle } x \text{ mit } |x - x_0| \geq 2r. \end{cases}$$

Es sei ferner $\varphi \in \mathring{\mathcal{C}}_\infty(\mathbb{R}^3)$.

1. Aus der Voraussetzung, $\nabla(\psi\varphi) = \psi\nabla\varphi + \varphi\nabla\psi$ und

$$\operatorname{div}(\psi\nabla\varphi - \varphi\nabla\psi) = \psi\Delta\varphi - \varphi\Delta\psi$$

folgt

$$\int_\Omega \psi\, u\, (-\Delta + 1)\varphi = \int_\Omega \left\{ -u\varphi\Delta\psi + 2u \operatorname{div}(\varphi\nabla\psi) \right\}.$$

2. Es sei

$$g(x) := \begin{cases} \frac{1}{4\pi} \frac{e^{|x|}}{|x|} & \text{für } x \neq o \\ 0 & \text{für } x = o. \end{cases}$$

Dann folgt wie im Beweis zu Beispiel 6.6.4

$$(-\Delta + 1)g = \delta(o)$$

auf $\mathcal{D}(\mathbb{R}^3)$, das heißt, es gilt

$$\forall\, \varphi \in \mathring{\mathcal{C}}_\infty(\mathbb{R}^3) \qquad \int_{\mathbb{R}^3} g\,(-\Delta + 1)\varphi = \big(g * (-\Delta + 1)\varphi\big)(0) = \varphi(o).$$

3. Es ist

$$\begin{aligned}
\int_\Omega \psi\, u\, \varphi &= \int_{x\in\Omega} (\psi u)(x) \int_{y\in\mathbb{R}^3} g(y)\, \big((-\Delta+1)\varphi\big)(x+y)\, dy\, dx \\
&= \int_{y\in\mathbb{R}^3}\int_{x\in\Omega} \Big\{ \big(u\,\Delta\psi\big)(x)\, \varphi(x+y) + 2\big(u\nabla\psi\big)(x)\, \nabla\varphi(x+y) \Big\}\, g(y)\, dx\, dy \\
&= \int_{z\in\mathbb{R}^3}\int_{x\in\Omega} \Big\{ \big(u\,\Delta\psi\big)(x)\, \varphi(z) + 2\big(u\nabla\psi\big)(x)\, \nabla\varphi(z) \Big\}\, g(z-x)\, dx\, dz \\
&= \int_{z\in\mathbb{R}^3}\int_{x\in\Omega} \Big\{ \big(u\,\Delta\psi\big)(x)\, g(z-x) - 2\big(u\nabla\psi\big)(x)\, \nabla_z g(z-x) \Big\}\, \varphi(z)\, dx\, dz.
\end{aligned}$$

Mithin gilt fast überall

$$(\psi u)(z) = \int_\Omega \Big\{ \big(u\,\Delta\psi\big)(x)\, g(z-x) - 2\big(u\nabla\psi\big)(x)\, \nabla_z g(z-x) \Big\}\, dx.$$

Daraus folgt für $x \in B(x_0, r)$ fast überall

$$u(x) = \int_{\Omega\setminus B(x_0,r)} \Big\{ \big(u\,\Delta\psi\big)(y)\, g(x-y) - 2\big(u\nabla\psi\big)(y)\, \nabla_z g(x-y) \Big\}\, dy,$$

denn in $B(x_0, r)$ verschwinden die Ableitungen von ψ und damit der Integrand. Damit ist fast überall $u \in \mathcal{C}_\infty(B(x_0, r))$. □

Bemerkung 6.6.9:

1. Die Aussage des Weylschen Lemmas läßt sich auch auf schwache Lösungen $u \in \mathcal{L}^1_{loc}(\Omega)$ von

$$(-a_{ik}\partial_i\partial_k + a_i\partial_i + a)u = f$$

verallgemeinern. Dabei seien $a_{ik} = a_{ki} \in \mathcal{C}_{2+\alpha}(\Omega)$, $a_i \in \mathcal{C}_{1+\alpha}(\Omega)$, $a \in \mathcal{C}_1(\Omega)$,

$$\exists\, p > 0 \quad \forall\, \xi \in \mathbb{R}^n \quad \forall\, x \in \Omega \qquad \xi_i\, a_{ik}(x)\, \xi_k \geq p\, |\xi|^2$$

sowie $f \in \mathcal{C}_\alpha(\Omega)$. Dann ist fast überall $u \in \mathcal{C}_{2+\alpha}(\Omega)$. Einen Beweis findet man in *Hellwig [1960], S. 189.*

2. Eine andere Methode zum Nachweis der Regularität besteht darin, direkt den Differenzenquotienten zu bilden und den Grenzübergang auszuführen. Man vergleiche *Nirenberg [1955].* Wir skizziere den Beweis:

Durch Multiplikation mit einer geeigneten Testfunktion läßt sich das Problem lokalisieren. Es sei also $u \in \mathcal{W}_1(\Omega)$ mit $\operatorname{supp} u \Subset U(x_0) \subset \Omega$. Dabei sei $U(x_0)$ eine offene Umgebung von x_0. Wir wollen $u \in \mathcal{W}_2(\Omega)$ zeigen. Dieser Schluß läßt sich dann wiederholen, und es folgt $u \in \mathcal{W}_k(\Omega)$ mit beliebigem k. Aus einem *Sobolevschen Einbettungssatz*, den wir noch besprechen werden, folgt dann die Regularität.

Es seien also $u, v \in \mathcal{W}_1(\Omega)$,

$$\forall\, f \in \mathcal{W}_1(\Omega) \qquad (u, f)_1 = 0$$

sowie $\psi \in \mathring{\mathcal{C}}_\infty(\Omega)$ mit $\psi = 1$ in einer Umgebung von x_0, und $\hat{u} := u\psi$, $\hat{v} := v\psi$. Dann ist

$$(\hat{u}, \hat{v})_1 = \underbrace{(u, \psi^2 v)_1}_{=0} - \big(\nabla\hat{u}, (\nabla\psi)v\big) + \big(u, (\nabla\psi)\nabla\hat{v}\big) + \big(u, (\nabla\psi)^2 v\big).$$

Wichtig ist, daß auf der rechten Seite dieser Gleichung $\nabla\hat{u}$ und $\nabla\hat{v}$ nur linear auftreten. Wir wählen nun $v = u$ und setzen dafür einen Differenzenquotienten $\delta_h u$ ein. Dann folgt mit $c_i = c_i(\|u\|_1, \psi)$

$$\|\delta_h \nabla\hat{u}\|^2 \leq c_1 + c_2\|\delta_h \nabla\hat{u}\|$$

oder

$$\|\delta_h \nabla\hat{u}\| \leq c_3.$$

Nach Satz 4.1.4 gibt es deshalb ein $g \in \mathcal{L}^2(\Omega)$ und eine Folge $h_n \to 0$ mit

$$\underset{h_n \to 0}{\text{w-lim}}\, \delta_{h_n} \nabla\hat{u} = g.$$

Für alle $\varphi \in \mathring{\mathcal{C}}_\infty(\Omega)$ gilt also

$$(g, \varphi) \leftarrow (\delta_{h_n} \nabla\hat{u}, \varphi) = -(\nabla\hat{u}, \delta_{h_n}\varphi) \to -(\nabla\hat{u}, \partial\varphi),$$

das heißt, es ist

$$g = \partial\nabla\hat{u}. \qquad \square$$

In den drei folgenden Sätzen fassen wir einige Resultate über Faltungen mit Testfunktionen zusammen.

Satz 6.6.10: *Es seien $u \in \mathcal{L}^1_{loc}(\mathbb{R}^n)$ m-mal schwach differenzierbar und $\varphi \in \mathring{\mathcal{C}}_\infty(\mathbb{R}^n)$. Dann gilt:*

(1) $u * \varphi \in \mathcal{C}_\infty(\mathbb{R}^n)$.

(2) $\partial^\alpha(\varphi * u) = (\partial^\alpha\varphi) * u$ *für alle* $\alpha \in \mathbb{N}_0^n$.

(3) $\partial^\alpha(\varphi * u) = \varphi * (\partial^\alpha u)$ *für alle* $\alpha \in \mathbb{N}_0^n$ *mit* $|\alpha| \leq m$.

Satz 6.6.11: *Es seien $u \in \mathcal{W}^p_m(\mathbb{R}^n)$ für $1 \leq p \leq \infty$ und $\varphi \in \mathring{\mathcal{C}}_\infty(\mathbb{R}^n)$. Dann gilt:*

(1) $\varphi * u \in \mathcal{C}_\infty(\mathbb{R}^n) \cap \mathcal{W}^p_m(\mathbb{R}^n)$.

(2) $\|\varphi * u\|_{m,p} \leq \|\varphi\|_1 \cdot \|u\|_{m,p}$.

Satz 6.6.12: *Es seien $u \in \mathcal{W}^p_m(\mathbb{R}^n)$ für $1 \leq p < \infty$ und (J_ε) eine Dirac-Folge. Dann gilt*

$$\lim_{\varepsilon \to 0} J_\varepsilon * u = u \quad \text{in } \mathcal{W}^p_m(\mathbb{R}^n),$$

und es folgt

$$\mathcal{H}^p_m(\mathbb{R}^n) = \mathcal{W}^p_m(\mathbb{R}^n) \quad \text{für } 1 \leq p < \infty.$$

6.6.1 Stark gleich Schwach

Der Beweis der letzten Aussage für beliebiges $\Omega \subset \mathbb{R}^n$ ist technisch sehr viel aufwendiger. Wir wollen ihn nun führen und müssen dazu etwas ausholen.

Es sei $\Omega \in \mathbb{R}^n$ wieder eine offene Menge und

$$\Omega_i := \left\{ x \in \Omega \;\middle|\; \operatorname{dist}(x, \partial\Omega) > \frac{1}{i} \quad \text{und} \quad |x| < i \right\}.$$

Dann ist $\Omega_i \subset \Omega$ offen und beschränkt, und es gilt

$$\bar{\Omega}_i \subset \Omega_{i+1} \quad \text{und} \quad \operatorname{dist}(\Omega_i, \partial\Omega_{i+1}) > 0.$$

Die $\{\Omega_i\}$ bilden eine offene Überdeckung von Ω. Wir sind nun im folgenden an einer *lokal finiten Überdeckung* von Ω interessiert.

Definition 6.6.13: *Eine Überdeckung $\{U_j\}_{j \in \mathbb{N}}$ von Ω heißt* lokal finit, *wenn es zu jedem $x \in \Omega$ eine Kugel $B(x, \varepsilon)$ so gibt, daß die Menge*

$$\left\{ i \in \mathbb{N} \;\middle|\; U_i \cap \bar{B}(x, \varepsilon) \neq \emptyset \right\}$$

endlich ist.

Ferner benötigen wir eine *Partition der Eins* auf Ω.

Definition 6.6.14: *Die Menge* $\{\eta_j\}_{j\in\mathbb{N}}$ *mit* $\eta_j \in \mathring{\mathcal{C}}_\infty(\Omega)$ *heißt* Partition der Eins *auf* Ω *zu* $\{U_j\}$, *wenn folgendes gilt:*

(1) $\operatorname{supp}\eta_j \subset U_j$ *und* $\eta_j \geq 0$ *für alle* j.

(2) Für alle $x \in \Omega$ *ist*

$$\sum_{j=1}^{\infty} \eta_j(x) = 1.$$

Wenn die Überdeckung $\{U_j\}$ lokal finit ist, dann enthält diese Summe bei festem x nur endlich viele von Null verschiedene Summanden.

Als erstes geben wir nun eine lokal finite Überdeckung von Ω an. Es seien

$$U_1 := \Omega_3, \;\; U_2 := \Omega_4 \quad \text{und} \quad U_i := \Omega_{i+2}\backslash\bar{\Omega}_{i-2} \quad \text{für } i \geq 3,$$

also $U_3 = \Omega_5\backslash\bar{\Omega}_1$. Offenbar gilt:

1. Die U_i sind beschränkt und offen.

2. $\bar{U}_i \subset \Omega$.

3. $\{U_i\}_{i\in\mathbb{N}}$ ist eine lokal finite Überdeckung von Ω.

Es sei nämlich $x \in \Omega$. Dann gibt es ein i_0 mit

$$\forall\, i \geq i_0 \qquad x \in \Omega_i.$$

Es sei i_0 möglichst klein gewählt. Für $i_0 > 2$ ist x dann nur in

$$U_{i_0+1} = \Omega_{i_0+3}\backslash\bar{\Omega}_{i_0-1}, \;\; \ldots, \;\; U_{i_0-2} = \Omega_{i_0}\backslash\bar{\Omega}_{i_0-4}$$

enthalten.

Als nächstes konstruieren wir eine Partition der Eins auf Ω zu diesen $\{U_j\}$. Es seien

$$K_1 := \bar{\Omega}_2 \quad \text{und} \quad K_i := \bar{\Omega}_{i+1}\backslash\Omega_{i-1} \quad \text{für } i \geq 2.$$

Dann sind die $K_i \Subset \mathbb{R}^n$, und es ist

$$K_i \subset U_i.$$

Ferner gibt es zu jedem $x \in \Omega$ ein K_i mit $x \in K_i$, also ist

$$\Omega = \bigcup_{i=1}^{\infty} K_i.$$

Nach §6.4.2 existieren nun $\hat{\eta}_i \in \mathring{\mathcal{C}}_\infty(U_i)$ mit

$$\hat{\eta}_i = 1 \;\text{ auf }\; K_i, \quad \operatorname{supp}\hat{\eta}_i \subset U_i \quad \text{und} \quad \hat{\eta}_i \geq 0.$$

Ferner gilt

$$\forall\, i \in \mathbb{N} \quad \forall\, j \geq i+2 \qquad \hat{\eta}_j = 0 \quad \text{auf } \Omega_i.$$

Deshalb können wir $\hat{\eta} \in \mathcal{C}_\infty(\Omega)$ definieren vermöge

$$\forall\, i \in \mathbb{N} \quad \forall\, x \in \Omega_i \qquad \hat{\eta}(x) := \sum_{j=1}^{i+1} \hat{\eta}_j(x).$$

Wegen $\Omega = \bigcup K_i$ ist $\hat{\eta} \geq 1$ auf Ω. Die

$$\eta_i := \frac{\hat{\eta}_i}{\hat{\eta}} \in \mathring{\mathcal{C}}_\infty(U_i)$$

leisten deshalb das Gewünschte, sie bilden also eine Partition der Eins auf Ω zu $\{U_i\}$.

Nach diesen Vorbereitungen zeigen wir

Satz 6.6.15: *Für* $1 \leq p < \infty$ *ist*

$$\mathcal{H}_m^p(\Omega) = \mathcal{W}_m^p(\Omega).$$

Beweis: Es seien $u \in \mathcal{W}^p_m(\Omega)$ und $\varepsilon > 0$ vorgegeben und $\{\eta_i\}$ die soeben konstruierte Partition der Eins. Es seien ferner

$$u_k := \begin{cases} \eta_k u & \text{auf } \Omega \\ 0 & \text{sonst.} \end{cases}$$

Dann gilt für alle $k \in \mathbb{N}$ auf Ω_k

$$u = \sum_{j=1}^{k+1} u_j$$

sowie

$$\forall\, \alpha,\ |\alpha| \le m, \qquad \partial^\alpha u = \sum_{j=1}^{k+1} \partial^\alpha u_j .$$

Wir verwenden nun den Friedrichsschen Mollifier (vgl. §6.4.2) und bilden $j_\delta * u_k \in \mathcal{C}_\infty(\mathbb{R}^n) \cap \mathcal{W}^p_m(\mathbb{R}^n)$ zu $\delta > 0$. Dann gilt

$$\lim_{\delta \to 0} \|j_\delta * u_k - u_k\|_{m,p,\mathbb{R}^n} = 0.$$

Wir wählen nun ε_k so klein, daß für $k \ge 4$ und $\varphi_k := j_{\varepsilon_k} * u_k \in \mathcal{C}_\infty(\mathbb{R}^n)$

$$\operatorname{supp} \varphi_k \subset \Omega_{k+3} \backslash \bar{\Omega}_{k-3}$$

und

$$\|\varphi_k - u_k\|_{m,p,\mathbb{R}^n} < \frac{1}{2^k} \frac{\varepsilon}{2N^{1/p}}$$

gilt, mit

$$N := \#\left\{\alpha \in \mathbb{N}_0^n \;\middle|\; |\alpha| \le m\right\} = \binom{m+n}{m}.$$

Es sei also etwa $\varepsilon_k < 1/(k+2)(k+3)$. Dann ist

$$\forall\, k \in \mathbb{N} \quad \forall\, j \ge k+3 \qquad \varphi_j = 0 \quad \text{auf } \Omega_k,$$

und wir können durch

$$\varphi(x) := \sum_{j=1}^{k+2} \varphi_j(x) \quad \text{für } x \in \Omega_k$$

die Abbildung $\varphi \in \mathcal{C}_\infty(\mathbb{R}^n)$ definieren. Für alle $\alpha \in \mathbb{N}_0^n$ mit $|\alpha| \le m$ ist dann

$$\left\|\partial^\alpha \varphi - \partial^\alpha u\right\|_{p,\Omega_k} = \left\| \sum_{j=1}^{k+2} \left(\partial^\alpha \varphi_j - \partial^\alpha u_j\right) \right\|_{p,\Omega_k} \le \sum_{j=1}^{k+2} \left\|\varphi_j - u_j\right\|_{m,p,\Omega_k} \le \sum_{j=1}^{k+2} \left\|\varphi_j - u_j\right\|_{m,p,\mathbb{R}^n} < \frac{\varepsilon}{2N^{1/p}}.$$

Wegen $\Omega = \bigcup \Omega_k$ folgt aus dem Satz von B. Levi über die monotone Konvergenz $\partial^\alpha \varphi - \partial^\alpha u \in \mathcal{L}^p(\Omega)$ und

$$\|\partial^\alpha \varphi - \partial^\alpha u\|_{p,\Omega} \le \frac{\varepsilon}{2N^{1/p}}.$$

Mithin ist $\varphi \in \mathcal{W}^p_m(\Omega)$, weil u es ist. Andererseits folgt aus der Konstruktion $\varphi \in \mathcal{C}_\infty(\Omega)$. Damit gilt

$$\varphi \in \mathcal{W}^p_m(\Omega) \cap \mathcal{C}_\infty(\Omega) \subset \mathcal{H}^p_m(\Omega),$$

und

$$\|\varphi - u\|_{m,p,\Omega} < \varepsilon. \qquad \square$$

Damit haben wir ein wichtiges Resultat bewiesen. Es folgen noch einige Eigenschaften schwach differenzierbarer Abbildungen. Wir beginnen mit der nach Gottfried Wilhelm Leibniz, 1646–1716, benannten Regel.

Die Leibnizsche Regel: *Es seien $1 \le p \le \infty$, $u \in \mathcal{H}^p_m(\Omega)$ und $\psi \in \mathcal{BC}_m(\Omega)$. Dann gilt:*

(1) $\psi u \in \mathcal{H}^p_m(\Omega)$.

(2) Für alle α mit $|\alpha| \le m$ ist

$$\partial^\alpha(\psi u) = \sum_{\beta \le \alpha} \binom{\alpha}{\beta} \left(\partial^\beta \psi\right) \left(\partial^{\alpha-\beta} u\right).$$

Beweis: Im Falle $p = \infty$ ist die Behauptung klassisch, es sei also $p < \infty$: Dann gilt

$$\forall\, \alpha,\, |\alpha| \le m, \qquad \sum_{\beta \le \alpha} \binom{\alpha}{\beta} \left(\partial^\beta \psi\right) \left(\partial^{\alpha-\beta} u\right) \in \mathcal{L}^p(\Omega).$$

Ferner gibt es $u_i \in \mathcal{C}_m^{p*}(\Omega)$ mit

$$\|u_i - u\|_{m,p,\Omega} \to 0.$$

Es sei $\varphi \in \mathring{\mathcal{C}}_\infty(\Omega)$. Dann folgt klassisch

$$\int_\Omega \psi\, u_i\, \partial^\alpha \varphi = (-1)^{|\alpha|} \int_\Omega \left\{ \sum_{\beta \le \alpha} \binom{\alpha}{\beta} \left(\partial^\beta \psi\right) \left(\partial^{\alpha-\beta} u_i\right) \right\} \varphi,$$

und wir erhalten für $i \to \infty$

$$\int_\Omega \psi\, u\, \partial^\alpha \varphi = (-1)^{|\alpha|} \int_\Omega \left\{ \sum_{\beta \le \alpha} \binom{\alpha}{\beta} \left(\partial^\beta \psi\right) \left(\partial^{\alpha-\beta} u\right) \right\} \varphi.$$

Deshalb ist $\psi u \in \mathcal{W}_m^p(\Omega) = \mathcal{H}_m^p(\Omega)$. □

Satz 6.6.16: *Für $1 \le p < \infty$ ist*

$$\left\{ u \in \mathcal{H}_m^p(\Omega) \;\middle|\; \operatorname{supp} u \text{ ist beschränkt} \right\}$$

in $\mathcal{H}_m^p(\Omega)$ dicht.

Diese Aussage ist natürlich nur für unbeschränkte Mengen Ω interessant.

Zum Beweis wählen wir ein $\psi \in \mathring{\mathcal{C}}_\infty(\mathbb{R}^n)$ mit

$$\psi(x) := \begin{cases} 1 & \text{für } |x| \le 1 \\ 0 & \text{für } |x| \ge 2 \end{cases}$$

und zu $u \in \mathcal{H}_m^p(\Omega)$ fast überall in Ω

$$u_k(x) := \psi\left(\frac{x}{k}\right) u(x).$$

Dann hat u_k einen beschränkten Träger, und aus der Leibnizschen Regel folgt

$$\partial^\alpha u_k - \partial^\alpha u = \sum_{\substack{\beta \le \alpha \\ \beta \ne o}} \binom{\alpha}{\beta} \frac{1}{k^{|\beta|}} \left(\partial^\beta \psi\left(\frac{\cdot}{k}\right)\right) \left(\partial^{\alpha-\beta} u\right) + \left(\psi\left(\frac{\cdot}{k}\right) - 1\right) \partial^\alpha u,$$

folglich

$$\left\|\partial^\alpha u_k - \partial^\alpha u\right\|_{\mathcal{L}^p(\Omega)} \le \frac{1}{k} \sum_{\substack{\beta \le \alpha \\ \beta \ne o}} \binom{\alpha}{\beta} \|\partial^\beta \psi\|_\infty \cdot \|\partial^{\alpha-\beta} u\|_{\mathcal{L}^p(\Omega)} + \Big(\int_{|x|>k} |\partial^\alpha u(x)|^p \, dx \Big)^{1/p}.$$

Es gilt damit

$$\lim_{k \to \infty} \|u_k - u\|_{m,p,\Omega} = 0.$$ □

Folgerung 6.6.17: *Für $m \in \mathbb{N}_0$ und $1 \le p < \infty$ ist $\mathring{\mathcal{C}}_\infty(\mathbb{R}^n)$ in $\mathcal{H}_m^p(\mathbb{R}^n)$ dicht, es gilt also*

$$\mathring{\mathcal{H}}_m^p(\mathbb{R}^n) \;=\; \mathcal{H}_m^p(\mathbb{R}^n).$$

6.6.2 Stetige Einbettungen

Nun wollen wir uns einem zweiten interessanten Fragenkomplex zuwenden, nämlich den *Sobolevschen Einbettungssätzen*. Wir beginnen mit einem einfachen Resultat.

Ein Sobolevscher Einbettungssatz: *Es seien $\Omega \subset \mathbb{R}^n$ beschränkt, $j > n/2$ und $u \in \mathring{\mathcal{H}}_j(\Omega)$. Dann ist $u \in \mathcal{C}_0(\bar{\Omega})$, und es gibt eine von u unabhängige Konstante c mit*

$$\|u\|_\infty \le c\, \|u\|_{\mathcal{H}_j(\Omega)}.$$

Die Einbettung von $\mathring{\mathcal{H}}_j(\Omega)$ in $\mathcal{C}_0(\bar{\Omega})$ ist also stetig, dafür schreiben wir auch

$$\mathring{\mathcal{H}}_j(\Omega) \hookrightarrow \mathcal{C}_0(\bar{\Omega}).$$

Beweis: Wir können $u \in \mathring{\mathcal{H}}_j(\Omega)$ durch Null in einen Würfel $W \supset \Omega$ zu $u \in \mathring{\mathcal{H}}_j(W)$ und dann periodisch in $\mathbb{R}^n$ fortsetzen. O.B.d.A. sei W ein Würfel mit den Kantenlängen 2π. Wir entwickeln u dann in eine Fourierreihe. Mit $l \in \mathbb{Z}^n$ und $k_n := (2\pi)^{-n/2}$ bilden die

$$\varphi_l(x) := k_n \, \mathrm{e}^{\mathrm{i}\, lx}$$

ein vollständiges Orthonormalsystem in $\mathcal{L}^2(W)$. Dabei steht lx für $(l,x)_{\mathbb{R}^n}$. Es ist

$$(\partial^\alpha u, \varphi_l) = \mathrm{i}^{|\alpha|} l^\alpha (u, \varphi_l),$$

und die Besselsche Ungleichung für $\partial^\alpha u$ lautet

$$\sum_{l\in\mathbb{Z}^n} (l^\alpha)^2 |(u,\varphi_l)|^2 = \sum_{l\in\mathbb{Z}^n} |(\partial^\alpha u, \varphi_l)|^2 \le \|\partial^\alpha u\|^2_{\mathcal{L}^2(W)}.$$

Daraus folgt die Existenz einer Konstanten c mit

$$\sqrt{\sum_{l\in\mathbb{Z}^n} (1+l^2)^j |(u,\varphi_l)|^2} \le c \, \|u\|_{\mathcal{H}_j(W)}.$$

Der Ausdruck auf der linken Seite dieser Abschätzung definiert eine Norm, und man erhält damit einen zu $\mathcal{H}_j(W)$ äquivalenten Hilbertraum H_j, in dem sich leichter rechnen läßt.

Wir kommen nun zum eigentlichen Beweis. Es ist mit einer von u unabhängigen Konstanten c

$$\begin{aligned}
\sum_{l\in\mathbb{Z}^n} |(u,\varphi_l)\varphi_l(x)| &= k_n \sum_{l\in\mathbb{Z}^n} \left\{ \left(1+l^2\right)^{j/2} \, |(u,\varphi_l)| \right\} \frac{1}{(1+l^2)^{j/2}} \\
&\le k_n \sqrt{\sum_{l\in\mathbb{Z}^n} (1+l^2)^j \, |(u,\varphi_l)|^2} \cdot \sqrt{\sum_{l\in\mathbb{Z}^n} \frac{1}{(1+l^2)^j}} \\
&\le c \, \|u\|_{\mathcal{H}_j(W)},
\end{aligned}$$

letzteres wegen

$$\sum_{l\in\mathbb{Z}^n} \frac{1}{(1+l^2)^j} \le 1 + \omega_n \int_{r>0} \frac{r^{n-1} dr}{(1+r^2)^j} < \infty.$$

Wegen der absoluten und gleichmäßigen Konvergenz der Reihe existiert also

$$\tilde{u}(x) := \sum_{l\in\mathbb{Z}^n} (u,\varphi_l)\, \varphi_l(x) \in \mathcal{C}_0(\bar{W})$$

und stimmt fast überall mit u überein. Es ist

$$\|\tilde{u}\|_\infty \le \gamma \, \|u\|_{\mathcal{H}_j(\Omega)}$$

mit einer von u unabhängigen Konstanten γ. □

Analog erhält man

Satz 6.6.18: *Es seien $\Omega \subset \mathbb{R}^n$ beschränkt und $1 \le p < \infty$. Dann gilt*

$$\mathring{\mathcal{H}}^p_{j+m}(\Omega) \hookrightarrow \mathring{\mathcal{H}}^q_j(\Omega)$$

für $p \le q \le np/(n-mp)$ und $mp < n$, sowie

$$\mathring{\mathcal{H}}^p_{j+m}(\Omega) \hookrightarrow \mathcal{C}_j(\bar{\Omega})$$

für $mp > n$.

Speziell gilt also im Falle $p = q = 2$

$$\mathring{\mathcal{H}}_{j+m}(\Omega) \hookrightarrow \mathcal{C}_j(\bar{\Omega}) \quad \text{für } m > \frac{n}{2}.$$

Entsprechende Sätze gelten für $\mathcal{H}_j(\Omega)$, auch für unbeschränkte Ω. Man benötigt dann aber Glattheitsvoraussetzungen an den Rand von Ω. Es lassen sich auch *Spuren* definieren. In dieser Vorlesung können aber nur Andeutungen geben werden. Wegen weiterer Einzelheiten sei auf das bereits zitierte Buch von *Adams* verwiesen. Insbesondere entnehmen wir diesem Buch folgende Aussage: (Vgl. *Adams [1975], S.54*)

Satz 6.6.19: *Es sei $1 \leq p < \infty$, und $\Omega \subset \mathbb{R}^n$ besitze die strikte Segmenteigenschaft. Dann sind die Restriktionen der Abbildungen aus $\mathring{\mathcal{C}}_\infty(\mathbb{R}^n)$ auf Ω in $\mathcal{W}_m^p(\Omega)$ dicht.*

Wichtig ist, daß man zum Nachweis eine Glattheitsvoraussetzung an den Rand $\partial\Omega$ benötigt, nämlich

Definition 6.6.20: *Eine offene Menge $\Omega \subset \mathbb{R}^n$ hat die* strikte Segmenteigenschaft, *wenn $\partial\Omega$ eine lokal finite offene Überdeckung $\{O_i\}$ mit Vektoren $\{y_i\}$ so besitzt, daß*

$$\forall\, t \in (0,1) \quad \forall\, x \in \bar{\Omega} \cap O_i \qquad x + t\, y_i \in \Omega$$

ist.

Wir zeigen nun folgendes

Lemma 6.6.21: *Es seien*

$$R := \left\{ x \in \mathbb{R}^n \;\middle|\; a_i < x_i < b_i \right\}, \quad R' := \left\{ x' = (x_1, \ldots, x_{n-1}) \in \mathbb{R}^{n-1} \;\middle|\; a_i < x_i < b_i \right\},$$

$\zeta \in (a_n, b_n)$, $1 \leq p < \infty$ und $u \in \mathcal{C}_\infty(R) \cap \mathcal{W}_1^p(R)$. Dann gilt mit $k = k(p, b_n - a_n)$

$$\left\| u(\cdot, \zeta) \right\|_{\mathcal{L}^p(R')} \leq k \left\| u \right\|_{1,p,R}.$$

Das ist also ein *Spursatz*, die Abbildung

$$\begin{aligned} T_\zeta \;:\; \mathcal{H}_1^p(R) &\longrightarrow \mathcal{L}^p(R_\zeta), \\ u &\longmapsto u(\cdot, \zeta) \end{aligned}$$

mit $R_\zeta := \langle R', \zeta \rangle$ ist stetig, das heißt

$$\mathcal{H}_1^p(R) \hookrightarrow \mathcal{L}^p(R_\zeta).$$

Beweis: Weil R die Segmenteigenschaft besitzt, ist $\mathcal{C}_\infty(\bar{R})$ in $\mathcal{H}_1^p(R)$ dicht. Deshalb wählen wir $u \in \mathcal{C}_\infty(\bar{R})$. Dann ist

$$\int_{R'} |u(x', \cdot)|^p \, dx' \in \mathcal{C}_\infty([a_n, b_n]).$$

Aus dem Mittelwertsatz der Integralrechnung folgt deshalb die Existenz eines $\sigma \in [a_n, b_n]$ mit

$$\|u\|_{0,p,R}^p = (b_n - a_n) \int_{R'} |u(x', \sigma)|^p \, dx'.$$

Mit $p =: q(p-1)$ erhält man dann

$$|u(x', \zeta)|^p = \left| u(x', \sigma) + \int_\sigma^\zeta \partial_n u(x', t) \, dt \right|^p \leq 2^{p-1} \left\{ |u(x', \sigma)|^p + |\zeta - \sigma|^{\frac{p}{q}} \int_\sigma^\zeta |\partial_n u(x', t)|^p \, dt \right\},$$

mithin

$$\begin{aligned} \|u(\cdot, \zeta)\|_{0,p,R'}^p &\leq 2^{p-1} \left\{ \|u(\cdot, \sigma)\|_{0,p,R'}^p + (b_n - a_n)^{p-1} \|\partial_n u\|_{0,p,R}^p \right\} \\ &= 2^{p-1} \left\{ \frac{\|u\|_{0,p,R}^p}{b_n - a_n} + (b_n - a_n)^{p-1} \|\partial_n u\|_{0,p,R}^p \right\}. \end{aligned}$$

□

Folgerung 6.6.22: *Es gilt*

$$\mathcal{H}_n^1(R) \hookrightarrow \mathcal{C}(\bar{R}).$$

Beweis: Es sei wieder $u \in \mathcal{C}_\infty(\bar{R})$. Dann ist für $|\alpha| \leq n-1$

$$\|\partial^\alpha u(\cdot, x_n)\|_{0,1,R'} \leq k_1 \|\partial^\alpha u\|_{1,1,R}$$

und

$$\|u(\cdot, x_n)\|_{n-1,1,R'} \leq k_2 \|u\|_{n,1,R}.$$

So kann man fortfahren und erhält

$$\|u(\cdot, x_2, \ldots, x_n)\|_{1,1,(a_1,b_1)} \leq k_3 \|u\|_{n,1,R}.$$

Mit $\sigma \in (a_1, b_1)$ sei

$$\|u(\cdot, x_2, \ldots, x_n)\|_{0,1,(a_1,b_1)} = (b_1 - a_1)\, |u(\sigma, x_2, \ldots, x_n)|.$$

Dann ist wieder

$$\begin{aligned} |u(x)| &\leq |u(\sigma, x_2, \ldots, x_n)| + \int_\sigma^{x_1} \left|\partial_1 u(t, x_2, \ldots, x_n)\right| \, dt \\ &\leq \frac{\|u(\cdot, x_2, \ldots, x_n)\|_{0,1,(a_1,b_1)}}{b_1 - a_1} + \|\partial_1 u(\cdot, x_2, \ldots, x_n)\|_{0,1,(a_1,b_1)} \\ &\leq k_4 \|u\|_{n,1,R}. \end{aligned}$$

□

Um zu Aussagen für allgemeinere Ω zu gelangen, stellt man nun weitere Glattheitsvoraussetzungen an den Rand. Wichtig sind

Definition 6.6.23: *Eine offene Menge $\Omega \subset \mathbb{R}^n$ hat die* Kegeleigenschaft, *wenn es einen endlichen nicht entarteten Kegel C so gibt, daß*

$$\forall\, x \in \Omega \quad \exists\, C(x) \qquad C(x) \subset \Omega$$

ist. Dabei soll $C(x)$ ein zu C kongruenter Kegel mit der Spitze in x sein.

Definition 6.6.24: *Eine beschränkte offene Menge Ω hat die lokale* Lipschitz-Eigenschaft, *wenn es zu jedem Randpunkt $x \in \partial\Omega$ eine Umgebung $U(x)$ so gibt, daß $\partial\Omega \cap U(x)$ der Graph einer Lipschitz-stetigen Funktion ist.*

Besitzt Ω die Kegeleigenschaft und ist $mp < n$, dann gilt für $p \leq q \leq np/(n-mp)$

$$\mathcal{H}^p_{j+m}(\Omega) \hookrightarrow \mathcal{H}^q_j(\Omega).$$

Ist $mp > n$, dann folgt

$$\mathcal{H}^p_{j+m}(\Omega) \hookrightarrow \mathcal{BC}_j(\Omega).$$

Besitzt Ω die Lipschitz-Eigenschaft, dann gilt für $mp > n > (m-1)p$ und $0 < \lambda < m - n/p$ sogar

$$\mathcal{H}^p_{j+m}(\Omega) \hookrightarrow \mathcal{C}_{j,\lambda}(\bar{\Omega}).$$

6.6.3 Kompakte Einbettungen

Neben den stetigen sind die *kompakten Einbettungen* besonders wichtig. Auch hierzu stellen wir einen speziellen Fall dar, verweisen aber für allgemeinere Aussagen wieder auf das Buch von *Adams [1975]*. Zunächst zur Notation: Es seien $\mathcal{X}, \mathcal{Y}$ normierte Räume. Dann schreiben wir

$$\mathcal{X} \hookrightarrow\hookrightarrow \mathcal{Y},$$

wenn $\mathcal{X}$ eine Teilmenge von $\mathcal{Y}$ und die Inklusion

$$id \;:\; \mathcal{X} \longrightarrow \mathcal{Y}$$

kompakt ist. Damit ist gemeint, daß

$$\mathcal{R}\big(id(B(o,1))\big) \subset \mathcal{Y}$$

relativ kompakt ist.

Die folgenden Resultate gehen auf Franz Rellich, 1906–1955, zurück. Sie werden vielfach *Rellichsche Auswahlsätze* genannt.

Satz 6.6.25: *Es sei $\Omega \subset \mathbb{R}^n$ beschränkt. Dann gilt*

$$\mathring{\mathcal{H}}_1(\Omega) \hookrightarrow\hookrightarrow \mathcal{L}^2(\Omega).$$

Der Beweis erfolgt analog zum Sobolevschen Einbettungssatz. Wir setzen $u \in \mathring{\mathcal{H}}_1(\Omega)$ durch Null in einen Würfel W fort und entwickeln in eine Fourierreihe. Es seien $u \in \mathring{\mathcal{H}}_1(\Omega)$ mit $\|u_n\|_1 \leq 1$. Dann folgt

$$\sum_{l \in \mathbb{Z}^n} \left(1 + l^2\right) |(u_n, \varphi_l)|^2 \leq 1.$$

Mithin ist auch $|(u_n, \varphi_l)| \leq 1$, und wir können für $|l| = 0$, $|l| = 1$, usw. jeweils konvergente Teilfolgen von $\{(u_n, \varphi_l)\}$ auswählen. Es sei (u_n) bereits die Diagonalfolge. Dann konvergiert für alle $l \in \mathbb{Z}^n$ die Folge

$$\{(u_n, \varphi_l)\}_{n \in \mathbb{N}}.$$

Wir zeigen nun, daß (u_n) in $\mathcal{L}^2(W)$ konvergiert. Es ist

$$\|u_n - u_m\|^2 = \sum_{l \in \mathbb{Z}^n} \left|(u_n - u_m, \varphi_l)\right|^2 = \sum_{|l| < N} \cdots + \sum_{|l| \geq N} \cdots$$

und

$$\sum_{|l| \geq N} \left|(u_n, \varphi_l)\right|^2 \leq \frac{1}{1 + N^2} \sum_{|l| \geq N} (1 + l^2) \left|(u_n, \varphi_l)\right|^2 \leq \frac{1}{1 + N^2}.$$

Daraus folgt

$$\|u_n - u_m\|^2 \leq \sum_{|l| < N} \left|(u_n - u_m, \varphi_l)\right|^2 + \frac{2}{1 + N^2} < \varepsilon,$$

wenn wir zunächst N und dann bei festem N auch n, m genügend groß wählen. □

Dieses Resultat läßt sich leicht zu

$$\mathring{\mathcal{H}}_{j+1}(\Omega) \hookrightarrow\hookrightarrow \mathring{\mathcal{H}}_j(\Omega)$$

verallgemeinern. Schwieriger wird es jedoch wieder, wenn man analoge Sätze für $\mathcal{H}_j(\Omega)$ beweisen möchte. Für beschränkte Gebiete mit Kegeleigenschaft gilt für $mp < n$

$$\mathcal{H}^p_{j+m}(\Omega) \hookrightarrow\hookrightarrow \mathcal{H}^q_j(\Omega) \quad \text{mit } p \leq q \leq \frac{np}{n - mp},$$

und für $n = mp$, $1 \leq q < \infty$

$$\mathcal{H}^p_{j+m}(\Omega) \hookrightarrow\hookrightarrow \mathcal{H}^q_j(\Omega).$$

Für $mp > n$ gilt auch

$$\mathcal{H}^p_{j+m}(\Omega) \hookrightarrow\hookrightarrow \mathcal{BC}_j(\Omega)$$

und die entsprechende Aussage rechts mit $\mathcal{C}_j(\bar{\Omega})$, wenn Ω die Lipschitz-Eigenschaft besitzt.

Daß Voraussetzungen an den Rand zum Nachweis solcher Einbettungsresultate notwendig sind, wird in *Courant & Hilbert [1993], S. 521*, am einem Beispiel belegt.

Diesen Abschnitt beschließen wir mit der Beweisskizze zu folgendem

Satz 6.6.26: *Die offene Menge $\Omega \subset \mathbb{R}^n$ sei beschränkt und besitze die Segmenteigenschaft. Dann gilt*

$$\mathcal{H}_1(\Omega) \hookrightarrow\hookrightarrow \mathcal{L}^2(\Omega).$$

Beweis:

1. Analog zur Folgerung 6.3.2 zeigt man in $\mathbb{R}^n$ für beschränkte Ω mit Segmenteigenschaft

$$\exists\, \gamma > 0 \quad \forall\, \varepsilon \in (0, 1] \quad \exists\, \Omega_\varepsilon \Subset \Omega \quad \forall\, u \in \mathcal{L}^2_{loc}(\Omega) \text{ mit } \nabla u \in \mathcal{L}^2(\Omega)$$

$$\|u\|^2_{\mathcal{L}^2(\Omega)} \leq \gamma \left\{ \varepsilon \|\nabla u\|^2_{\mathcal{L}^2(\Omega)} + \|u\|^2_{\mathcal{L}^2(\Omega_\varepsilon)} \right\}.$$

2. Es seien nun $u_n \in \mathcal{H}_1(\Omega)$ mit $\|u_n\|_1 \leq 1$,

$$\Omega_i := \left\{ x \in \Omega \;\middle|\; \operatorname{dist}(x, \partial\Omega) > \frac{1}{i} \right\}$$

und

$$\varphi_i \in \mathring{\mathcal{C}}_1(\Omega) \quad \text{mit } \varphi_i|\Omega_i = 1.$$

Dann ist $(\varphi_i u_n)_{n\in\mathbb{N}}$ eine in $\mathring{\mathcal{H}}_1(\Omega)$ beschränkte Folge und enthält eine konvergente Teilfolge. Es sei (u_n) bereits die Diagonalfolge, dann konvergiert (u_n) also für jedes i in $\mathcal{L}^2(\Omega_i)$, und für jedes $\delta \in (0,1]$ gilt

$$\begin{aligned}\left\|u_n - u_m\right\|^2_{\mathcal{L}^2(\Omega)} &\le \gamma \left\{\delta \left\|\nabla(u_n - u_m)\right\|^2 + \left\|u_n - u_m\right\|^2_{\mathcal{L}^2(\Omega_\delta)}\right\} \\ &\le 4\gamma\delta + \gamma \left\|u_n - u_m\right\|^2_{\mathcal{L}^2(\Omega_\delta)} < \varepsilon,\end{aligned}$$

wenn wir zunächst δ genügend klein und anschließend n, m bei festem δ genügend groß wählen. □

6.6.4 Ein Fortsetzungssatz

Es sei schließlich – ohne Beweis – noch ein Fortsetzungssatz angegeben, der auf Alberto Pedro Calderón, *1920, zurückgeht. Dazu benötigen wir

Definition 6.6.27: *Eine offene Menge $\Omega \subset \mathbb{R}^n$ besitzt die* strikte Kegeleigenschaft, *wenn $\partial\Omega$ eine lokal finite Überdeckung $\{O_i\}$ mit Kegeln $\{C_i\}$ – deren Spitzen im Ursprung liegen – so besitzt, daß*

$$\forall\, x \in \Omega \cap O_i \qquad x + C_i \subset \Omega$$

ist.

Dann gilt folgender

Calderónscher Fortsetzungssatz: *Es seien $\Omega \subset \mathbb{R}^n$ beschränkt mit strikter Kegeleigenschaft und $1 < p < \infty$. Dann gibt es eine stetige lineare Abbildung*

$$T \;:\; \mathcal{H}^p_m(\Omega) \longrightarrow \mathcal{H}^p_m(\mathbb{R}^n)$$

mit

$$\forall\, u \in \mathcal{H}^p_m(\Omega) \qquad (Tu)|\Omega = u.$$

Einen Beweis findet man wieder im Buch von *Adams [1975], S. 91.*

Der Calderonsche Fortsetzungssatz liefert also zum Beispiel eine stetige Fortsetzung von $\mathcal{H}_2(\Omega)$ in $\mathcal{H}_2(\mathbb{R}^n)$. In den Anwendungen genügt manchmal die Existenz einer stetigen und lokal kompakten Fortsetzung von $\mathcal{H}_2(\Omega)$ in $\mathcal{H}_1(\mathbb{R}^n)$. Es ist interessant und wichtig, daß man dafür die strikte Kegeleigenschaft nicht benötigt. Es genügt vielmehr die strikte *p-Spitzeneigenschaft.* Diese Eigenschaft wird analog zur strikten Kegeleigenschaft definiert. Man ersetzt nur die Kegel C_i durch *p-Spitzen* X_i, nämlich durch Rotationen von

$$\left\{x = (x', x_n) \in \mathbb{R}^n \;\middle|\; 0 < x_n < h, \;\; \frac{x'}{(x_n)^p} \in \Sigma\right\}.$$

Dabei ist $1 \le p < 2$ und $\Sigma \subset \mathbb{R}^{n-1}$ eine offene beschränkte Menge. Die 1-Spitzeneigenschaft ist die Kegeleigenschaft. Den Beweis findet man in *Witsch [1993].*

6.7 Anwendungen

In diesem Abschnitt behandeln wir das Hausdorffsche Momentenproblem, einen Ergodensatz und Minimalprobleme der Variationsrechnung.

6.7.1 Das Hausdorffsche Momentenproblem

Mit $I := [0,1]$ und $n \in \mathbb{N}$ seien $f_n \in \mathcal{C}(I)$ und $\mu_n \in \mathbb{C}$ vorgegeben. Es soll ein $\alpha \in \mathcal{NBV}(I) \cong \mathcal{C}'(I)$ mit

$$\mu_n = \int_0^1 f_n(x)\, d\alpha(x) \qquad \text{für alle } n \in \mathbb{N}$$

bestimmt werden.

Ist speziell $f_n(x) = x^{n-1}$, dann handelt es sich um Momente bei einer Massenverteilung oder bei einer Verteilung von elektrischen Ladungen. Das Problem tritt auch in der Wahrscheinlichkeitsrechnung auf.

Allgemeiner seien $\mathcal{X}$ ein normierter Raum, (x_n) eine Folge in $\mathcal{X}$ und (μ_n) eine Folge in $\mathbb{K}$.

Definition 6.7.1: *Ein Element $x' \in \mathcal{X}'$ heißt Lösung des* Momentenproblems *zu (x_n) und (μ_n), wenn*

$$\forall\, n \in \mathbb{N} \qquad x' x_n = \mu_n$$

ist.

Als erstes fragt man nach der Eindeutigkeit des Momentenproblems. Es sei

$$M := \operatorname{span}(x_1, x_2, \ldots).$$

Eindeutigkeit liegt vor, wenn aus $x'M = 0$ das Verschwinden von x' folgt. Das ist sicherlich dann der Fall, wenn M in $\mathcal{X}$ dicht enthalten ist.

Etwas schwieriger ist die Frage nach der Existenz einer Lösung. Wir zeigen

Satz 6.7.2: *Das Momentenproblem ist genau dann lösbar, wenn*

$$\exists\, c \in \mathbb{R}_0^+ \quad \forall\, k \in \mathbb{N} \quad \forall\, (\beta_1, \ldots, \beta_k) \in \mathbb{K}^k \qquad \left| \sum_{i=1}^{k} \beta_i \mu_i \right| \le c \left\| \sum_{i=1}^{k} \beta_i x_i \right\| \tag{$*$}$$

gilt.

Beweis:

$\Rightarrow$: Das Momentenproblem sei lösbar, es sei also $\mu_n = x' x_n$. Dann folgt

$$\left| \sum_{i=1}^{k} \beta_i \mu_i \right| = \left| \sum_{i=1}^{k} \beta_i x' x_i \right| = \left| x' \sum_{i=1}^{k} \beta_i x_i \right| \le c \left\| \sum_{i=1}^{k} \beta_i x_i \right\|.$$

$\Leftarrow$: Die Bedingung $(*)$ sei erfüllt. Es sei ferner $m' \in M'$ vermöge

$$\begin{aligned} m' \;:\; & M \longrightarrow \mathbb{K}, \\ & x_n \longmapsto \mu_n \end{aligned}$$

erklärt. Wegen $(*)$ ist m' wohldefiniert und stetig mit $\|m'\| \le c$. Aufgrund des Satzes von Hahn-Banach gibt es deshalb eine Fortsetzung $x' \in \mathcal{X}'$ mit

$$x'\big|M = m' \quad \text{und} \quad \|x'\| = \|m'\|.$$

Dieses x' löst das Momentenproblem. □

Es sei nun die Folge (x_n) in $\mathcal{X}$ vorgegeben. Dann fragt man, für welche (μ_n) das Momentenproblem lösbar ist. Ist beispielsweise (x_n) beschränkt, dann folgt aus $(*)$ unmittelbar, daß auch (μ_n) beschränkt sein muß. Im speziellen Fall $\mathcal{X} = \mathcal{C}(I, \mathbb{R})$, $I = [0,1]$, $x_n(t) = t^{n-1}$ weiß man aufgrund des Weierstraßschen Satzes, daß die Polynome in $\mathcal{X}$ dicht liegen. Es kann also höchstens eine Lösung des Momentenproblems geben. Weil sich Funktionen aus $\mathcal{X}' \cong \mathcal{NBV}(I)$ immer als Differenz zweier monoton wachsender Funktionen schreiben lassen, ist man in diesem Fall besonders an der Frage interessiert, zu welchen Folgen (μ_n) es eine monoton wachsende Lösung $\alpha \in \mathcal{NBV}(I)$ mit

$$\mu_n = \int_0^1 t^{n-1}\, d\alpha(t)$$

gibt. Das ist das eigentliche *Hausdorffsche Momentenproblem.* Benannt ist es nach Felix Hausdorff, 1868–1942. Wir zeigen

Satz 6.7.3: *Das Hausdorffsche Momentenproblem ist genau dann lösbar, wenn* (μ_n) vom positiven Typ *ist. Das soll bedeuten:*

$$\forall\, (\beta_1, \ldots, \beta_k) \in \mathbb{R}^k \qquad \left(\forall\, t \in I \quad \sum_{i=1}^{k} \beta_i x_i(t) \ge 0 \right) \implies \sum_{i=1}^{k} \beta_i \mu_i \ge 0. \tag{$**$}$$

Beweis:

$\Leftarrow$: Die Bedingung $(**)$ sei erfüllt. Es sei ferner

$$h(t) := \sum_{i=1}^{k} \beta_i x_i(t).$$

Dann ist $|h(t)| \leq \|h\| = \|h\|\, x_1$, also

$$\|h\|\, x_1 \pm h(t) \geq 0.$$

Aus $(**)$ folgt deshalb

$$\|h\|\, \mu_1 \pm \sum_{i=1}^{k} \beta_i \mu_i \geq 0$$

oder

$$\Big|\sum_{i=1}^{k} \beta_i \mu_1\Big| \leq \mu_1 \|h\|.$$

Mithin gilt $(*)$ mit $c = \mu_1$, und es gibt ein $\alpha \in \mathcal{NBV}(I)$ mit $V(\alpha) \leq \mu_1$ und

$$\mu_n = \int_0^1 x_n(t)\, d\alpha(t).$$

Nun ist

$$\mu_1 = \alpha(1) - \alpha(0) \leq V(\alpha) \leq \mu_1.$$

Daraus folgt $V(\alpha) = \alpha(1) - \alpha(0)$. Mithin wächst α monoton, und das Hausdorffsche Momentenproblem ist lösbar.

$\Rightarrow$: Das Hausdorffsche Momentenproblem sei lösbar. Es gebe also eine monoton wachsende Funktion α mit

$$\mu_n = \int_0^1 t^{n-1}\, d\alpha(t).$$

Es sei ferner

$$\forall\, t \in I \qquad p(t) := \sum_{i=1}^{k} \beta_i x_i(t) \geq 0.$$

Dann folgt

$$\sum_{i=1}^{k} \beta_i \mu_i = \int_0^1 \sum_{i=1}^{k} \beta_i\, t^{i-1}\, d\alpha(t) = \int_0^1 p(t)\, d\alpha(t) \geq 0.$$

Die Bedingung $(**)$ ist also zur Lösbarkeit notwendig. □

Aus der Bedingung $(**)$ folgen andere Bedingungen an (μ_n), zum Beispiel ist $(**)$ äquivalent zu

$$\forall\, n \in \mathbb{N} \quad \forall\, k \in \mathbb{N}_0 \qquad \sum_{i=0}^{k} (-1)^i \binom{k}{i} \mu_{n+i} \geq 0. \tag{***}$$

Das Momentenproblem hatte große Bedeutung beim Lösen von Gleichungen und vor allem beim Beweis des *Spektralsatzes*. In §9 werden wir diesen Satz ausführlich behandeln. Deshalb erfolgt hier nur ein Hinweis auf einen einfachen Fall. Es seien $\mathcal{H}$ ein Hilbertraum, $I := [m, M]$ und $A : \mathcal{H} \to \mathcal{H}$ ein beschränkter symmetrischer Operator mit

$$\forall\, x \in \mathcal{H} \qquad m\,(x,x) \leq (Ax, x) \leq M\,(x,x).$$

Dann möchte man in Analogie zur Hauptachsentransformation der Matrizen eine *Spektraldarstellung* für A herleiten. Es seien $x \in \mathcal{H}$, $\mu_n := (A^{n-1}x, x)$ und mit $\beta_i \in \mathbb{R}$, $t \in [m, M]$

$$p(t) = \sum_{i=1}^{k} \beta_i t^{i-1} \geq 0$$

sowie

$$p(A) := \sum_{i=1}^{k} \beta_i A^{i-1}.$$

Dann ist auch $\sum_{i=1}^{k} \beta_i \mu_i = \big(p(A)x, x\big) \geq 0$. Einen Beweis dafür findet man in *Riesz & Sz.-Nagy [1956], S. 255*. Die Zuordnung ist also vom positiven Typ. Zu festem $x \in \mathcal{H}$ gibt es deshalb eine monoton wachsende Funktion α_x mit

$$\big(A^i x, x\big) = \int_m^M t^i \, d\alpha_x(t)$$

für alle $i \in \mathbb{N}_0$. Man führt dann eine *Spektralschar* mit Projektoren $P(t)$ ein, schreibt

$$\alpha_x(t) =: \big(P(t)x, x\big)$$

und erhält durch Polarisieren schließlich die Spektraldarstellung für A

$$\forall\, x, y \in \mathcal{H} \qquad (A^i x, y) = \int_m^M t^i \, d(P(t)x, y).$$

6.7.2 Ein Ergodensatz

Das Wort *Ergoden* ist eine Neuschöpfung aus den griechischen $\varepsilon\rho\gamma o\nu$ (Arbeit) und $o\delta o\varsigma$ (Weg). Es wurde von Ludwig Boltzmann, 1844–1906, eingeführt. Eine Darstellung des physikalischen Hintergrundes findet man in *Reed & Simon, l.c. S. 54*. Denken wir zum Beispiel an Lösungen $u(t)$ einer Anfangsrandwertaufgabe mit $u(t) \in \mathcal{H}$, ein Hilbertraum, $t \in \mathbb{R}_0^+$, und

$$\partial_t u + \mathrm{i}Au = o \quad \text{mit } u(0) = u^0 \in \mathcal{H}.$$

Die Voraussetzungen an den Operator A wollen wir jetzt nicht präzisieren. Die Lösung u kann man durch *Halbgruppen* in der Form

$$u(t) = U(t)\, u^0$$

gewinnen. Darauf werden wir in §10 etwas näher eingehen. Für $t \in \mathbb{R}_0^+$ ist zunächst formal

$$U(t) = \mathrm{e}^{-\mathrm{i}At}$$

mit

$$U(0) = id, \;\; U(t+s) = U(t)U(s) \quad \text{und} \quad \|U(t)\| \leq 1.$$

Man fragt dann nach der Existenz des zeitlichen Mittels

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t U(\tau) f \, d\tau$$

für $f \in \mathcal{H}$ und zeigt seine Gleichheit mit einem räumlichen Mittel. So erhält man einen *statistischen Ergodensatz*. Wir beweisen hier das diskrete Analogon, denken Sie an $T := U(1), \ldots, T^n := U(n)$.

Ergodensatz: *Es seien $\mathcal{X}$ ein reflexiver Banachraum und $T \in \mathcal{CL}(\mathcal{X})$ mit*

$$\exists\, c > 0 \quad \forall\, n \in \mathbb{N} \qquad \|T^n\| \leq c.$$

Dann ist

$$\mathcal{X} = \mathcal{N}(T - id) \oplus \overline{\mathcal{R}(T - id)}.$$

Es sei P der stetige Projektor von $\mathcal{X}$ auf $\mathcal{N}(T - id)$. Dann existiert für alle $x \in \mathcal{X}$

$$\lim_{n\to\infty} \frac{1}{n} \Big\{ Tx + T^2 x + \cdots + T^n x \Big\} = Px.$$

Es gibt also einen Fixpunkt $\bar{x}$ der Abbildung T mit

$$\Big\| \frac{1}{n} \Big\{ Tx + T^2 x + \cdots + T^n x \Big\} - \bar{x} \Big\| \to 0.$$

Beweis: Es sei

$$S_n := \frac{1}{n}\sum_{i=1}^{n} T^i.$$

Dann ist $\|S_n\| \leq c$, und es gilt

$$\forall\, x \in \mathcal{X} \quad \forall\, n \in \mathbb{N} \qquad (T-id)S_n x = S_n(T-id)x = \frac{1}{n}(T^{n+1}-T)x,$$

also

$$\|S_n(T-id)x\| \leq \frac{2c}{n}\|x\|.$$

Daraus folgt

$$\forall\, z \in \mathcal{R}(T-id) \qquad \|S_n z\| \to 0$$

und wegen $\|S_n\| \leq c$ sogar

$$\forall\, z \in \overline{\mathcal{R}(T-id)} \qquad \|S_n z\| \to 0.$$

Es sei nun $x \in \mathcal{X}$ fest gewählt und $x_n := S_n x$. Dann gilt

$$\|x_n\| \leq c\,\|x\| \quad \text{und} \quad \|(T-id)x_n\| \to 0.$$

Nach Voraussetzung ist $\mathcal{X}$ reflexiv, also nach Satz 4.1.4 die Kugel $\bar{B}(0,1) \subset \mathcal{X}$ schwach folgenkompakt. Es sei (x_n) bereits eine schwach konvergente Teilfolge. Dann gibt es ein $\bar{x} \in \mathcal{X}$ mit

$$\operatorname*{w\text{-}lim}_{n\to\infty} x_n = \bar{x}.$$

Wegen $T - id \in \mathcal{CL}(\mathcal{X})$ ist auch $(T-id)' \in \mathcal{CL}(\mathcal{X}')$, und es folgt für alle $x' \in \mathcal{X}'$

$$0 \leftarrow x' \cdot (T-id)x_n = (T-id)'x' \cdot x_n \to (T-id)'x' \cdot \bar{x} = x' \cdot (T-id)\bar{x}.$$

Mithin ist

$$(T-id)\bar{x} = o.$$

Daraus folgt für alle $n \in \mathbb{N}$

$$S_n\bar{x} = \bar{x} \quad \text{und} \quad x_n - \bar{x} = S_n(x-\bar{x}).$$

Als nächstes zeigen wir $x - \bar{x} \in \overline{\mathcal{R}(T-id)}$. Daraus folgt dann

$$\|x_n - \bar{x}\| = \|S_n(x-\bar{x})\| \to 0,$$

also die starke Konvergenz. Es gilt nämlich

$$S_n - id = (T-id)\sum_{i=0}^{n-1} \frac{n-i}{n}\, T^i. \qquad (*)$$

Deshalb ist

$$(S_n - id)(x-\bar{x}) = x_n - x \in \mathcal{R}(T-id)$$

und damit

$$\bar{x} - x \in \overline{\mathcal{R}(T-id)},$$

weil $\overline{\mathcal{R}(T-id)} =: \mathcal{X}_0$ schwach folgenabgeschlossen ist. Es sei nämlich $\bar{x}-x \notin \mathcal{X}_0$. Dann gibt es ein $x' \in \mathcal{X}'$ mit $x'|\mathcal{X}_0 = 0$ und $x'(\bar{x}-x) \neq 0$, das besagt der Trennungssatz. Nun ist für alle $n \in \mathbb{N}_0$

$$T^n x - T^{n+1}x = (id-T)T^n x \in \mathcal{X}_0,$$

also

$$x'T^{n+1}x = x'T^n x = \cdots = x'Tx = x'x$$

oder

$$x'(x_n - x) = 0.$$

Daraus folgt der Widerspruch $x'(\bar{x}-x) = 0$.

Definieren wir nun

$$\begin{aligned} P \;:\; & \mathcal{X} \longrightarrow \mathcal{X}, \\ & x \longmapsto \bar{x} = \lim_{n\to\infty} S_n x, \end{aligned}$$

dann ist P eine stetige Projektion auf $\mathcal{R}(P) = \mathcal{N}(T-id)$, und es gilt $\mathcal{N}(P) = \overline{\mathcal{R}(T-id)}$, also

$$\mathcal{X} = \mathcal{N}(T-id) \oplus \overline{\mathcal{R}(T-id)}.$$

Es ist nämlich

$$P^2 x = P\bar{x} = \lim_{n\to\infty} S_n \bar{x} = \bar{x} = Px.$$

Damit ist P eine Projektion mit $\mathcal{R}(P) = \mathcal{N}(T-id)$. Es sei $x \in \mathcal{N}(P)$, also

$$o = Px = \lim_{n\to\infty} S_n x.$$

Nun besagt Gl. $(*)$ $(S_n - id)x \in \mathcal{R}(T-id)$. Deshalb folgt

$$-x = Px - x = \lim_{n\to\infty} (S_n - id)x \in \overline{\mathcal{R}(T-id)}$$

oder

$$\mathcal{N}(P) \subset \overline{\mathcal{R}(T-id)}.$$

Es sei schließlich $z \in \mathcal{R}(T-id)$, also $z = (T-id)x$. Dann ist

$$Pz = PTx - Px = \lim_{n\to\infty} S_n Tx - Px = (T-id)Px = 0,$$

also

$$\overline{\mathcal{R}(T-id)} \;=\; \mathcal{N}(P). \qquad \square$$

6.7.3 Minimalprobleme der Variationsrechnung

In der Einführung wurde bereits angedeutet, wie man klassische Lösungen der Dirichletschen Randwertaufgabe der Potentialtheorie durch Minimieren des Dirichletschen Integrals

$$D(u) := \int_G (\nabla u)^2$$

gewinnen kann. Solche Minimalprobleme spielen eine große Rolle in der Variationsrechnung, der Theorie der partiellen Differentialgleichungen und in vielen anderen Zweigen der Mathematik.

Wir wollen hier nun die in der Einführung skizzierte Aufgabe in etwas allgemeinerem Rahmen behandeln. Es seien G ein beschränktes Gebiet in $\mathbb{R}^n$, $f \in \mathcal{H}_1(G)$ und $g \in \mathcal{L}^2(G)$ vorgegebene Funktionen,

$$J(u) := D(u) + 2(g,u), \quad D(u,v) := \int_G (\nabla u)(\nabla v), \quad E(u,v) := 2D(u,v) + 2(g,v)$$

sowie

$$\mathcal{D} := \left\{ u \in \mathcal{H}_1(G) \;\middle|\; u - f \in \mathring{\mathcal{H}}_1(G) \right\}.$$

Offenbar ist $\mathcal{D} \subset \mathcal{H}_1(G)$ eine konvexe Teilmenge. Eine einfache *Variationsaufgabe* lautet dann: Man bestimme ein $u \in \mathcal{D}$ mit

$$J(u) = \min_{v\in\mathcal{D}} J(v). \qquad (V)$$

Es sei (V) lösbar. Dann folgt aus

$$\forall\, v \in \mathcal{D} \qquad J(u) \le J\big(u + \varepsilon(v-u)\big)$$

und

$$\forall\, \varphi \in \mathring{\mathcal{C}}_\infty(G) \quad \forall\, \varepsilon \in \mathbb{R} \qquad J(u+\varepsilon\varphi) = J(u) + \varepsilon E(u,\varphi) + \varepsilon^2 D(\varphi)$$

die Aussage

$$\forall\, \varphi \in \mathring{\mathcal{C}}_\infty(G) \qquad E(u,\varphi) = 0. \qquad (E)$$

Die Funktion u ist mithin eine schwache Lösung der zum Variationsproblem (V) gehörenden *Eulerschen Differentialgleichung*, hier also

$$\Delta u = g.$$

Natürlich gelten die folgenden Überlegungen auch für allgemeinere Gleichungen mit variablen Koeffizienten.

Ziel dieses Abschnitts soll es nun sein, den folgenden Satz zu zeigen:

Satz 6.7.4: *Es sei $M \subset \mathcal{D}$ eine nichtleere, konvexe und abgeschlossene Teilmenge von $\mathcal{H}_1(G)$. Dann gilt*

$$\overset{1}{\exists}\, u \in M \quad \forall\, v \in M \qquad J(u) \leq J(v),$$

und jede Minimalfolge (u_n) aus M konvergiert in $\mathcal{H}_1(G)$ gegen diese Lösung u.

Wir werden sehen, daß sich die in der Einführung gestellte Aufgabe und auch allgemeinere dieser Aussage unterordnen lassen.

Den Beweis dieses Satzes führen wir in mehreren Schritten.

1. Es seien $u \in M$ eine Lösung, $v \in M$ und $\varepsilon \in [0,1]$. Dann ist

$$J(u) \leq J(u + \varepsilon(v-u)) = J(u) + \varepsilon E(u, v-u) + \varepsilon^2 D(v-u)$$

oder

$$\forall\, v \in M \qquad E(u, v-u) \geq 0. \tag{U}$$

Diese Abschätzung nennt man *Variationsungleichung* des Variationsproblems.

2. Wir zeigen die Eindeutigkeit: Es seien $u_1, u_2 \in M$ zwei Lösungen. Dann ist

$$0 = J(u_1) - J(u_2) = E(u_2, u_1 - u_2) + D(u_1 - u_2),$$

und aus der Variationsungleichung (U) folgt

$$D(u_1 - u_2) = 0.$$

Wegen $u_1 - u_2 \in \mathring{\mathcal{H}}_1(G)$ erhalten wir dann $u_1 = u_2$ aus der Ersten Poincaréschen Ungleichung in §6.3. □

Dem Nachweis der Existenz einer Lösung stellen wir nun zwei Lemmata voran. Wir unterbrechen deshalb den Beweis von Satz 6.7.4 kurz und zeigen

Lemma 6.7.5: *Es sei $u \in \mathcal{H}_1(G)$. Dann ist*

$$\begin{aligned} E(u,\cdot) \;:\; \mathcal{H}_1(G) &\longrightarrow \mathbb{R}, \\ v &\longmapsto 2D(u,v) + 2(g,v) \end{aligned}$$

ein stetiges lineares Funktional.

Der Beweis folgt aus

$$|E(u,v)| \leq c\big(\|\nabla u\|, \|g\|\big)\, \|v\|_1. \qquad \square$$

Lemma 6.7.6: *J ist schwach halbstetig nach unten.*

Beweis: Es seien $u_n \in \mathcal{H}_1(G)$ mit

$$\operatorname*{w\text{-}lim}_{n\to\infty} u_n = u \in \mathcal{H}_1(G).$$

Dann müssen wir

$$J(u) \leq \liminf_{n\to\infty} J(u_n)$$

zeigen. Nun ist

$$J(u_n) = J\big(u + (u_n - u)\big) = J(u) + E(u, u_n - u) + D(u_n - u)$$

mit $D(u_n - u) \geq 0$ und

$$\lim_{n\to\infty} E(u, u_n - u) = 0$$

nach Lemma 6.7.5. Mithin folgt

$$\liminf_{n\to\infty} J(u_n) \geq J(u). \qquad \square$$

Wir setzen den Beweis zu Satz 6.7.4 nun fort:

3. Wir zeigen die Existenz einer Lösung. Es sei $u \in M$. Dann ist $u - f \in \mathring{\mathcal{H}}_1(G)$ und aus der Ersten Poincaréschen Ungleichung folgt

$$\|u - f\| \leq c\,\|\nabla(u - f)\|$$

oder mit $c_1 = c_1(f)$

$$\|u\| \leq \|u - f\| + \|f\| \leq c_1(1 + \|\nabla u\|).$$

Daraus folgt mit $c_2 = c_2(f, g)$

$$\forall\, u \in M \qquad J(u) \geq \tfrac{1}{2}\,\|\nabla u\|^2 - c_2.$$

J ist also von unten beschränkt. Mithin existiert eine Minimalfolge (u_n) in M mit

$$J(u_n) \to d := \inf_{u\in M} J(u).$$

Wegen

$$\|u_n\|_1^2 \leq 2c_1^2 + \big(1 + 2c_1^2\big)\|\nabla u_n\|^2 \leq c_3\big(1 + J(u_n)\big) \quad \text{mit } c_3 = c_3(f, g)$$

ist (u_n) in $\mathcal{H}_1(G)$ beschränkt. $\mathcal{H}_1(G)$ ist reflexiv, mithin existiert nach Satz 4.1.4 eine schwach konvergente Teilfolge. Es sei (u_n) bereits diese Teilfolge. Dann gibt es ein $u \in \mathcal{H}_1(G)$ mit

$$\operatorname*{w\text{-}lim}_{n\to\infty} u_n = u.$$

Nach Satz 4.3.2 ist M schwach folgenabgeschlossen. Also ist $u \in M$, und aus Lemma 6.7.6 folgt

$$d \leq J(u) \leq \liminf_{n\to\infty} J(u_n) = d.$$

Das heißt, es ist

$$J(u) = d.$$

Die Funktion u löst also das Variationsproblem und infolgedessen auch die Variationsungleichung. Es sei (u_n) wieder die ursprüngliche Minimalfolge. Dann folgt aus

$$J(u) + E(u, u_n - u) + D(u_n - u) = J(u_n) \to J(u)$$

und der Variationsungleichung (U)

$$D(u_n - u) \to 0.$$

Wegen $\|u_n - u_m\|_1 \leq (1 + c)\,\|\nabla(u_n - u_m)\|$ ist (u_n) mithin sogar eine starke Cauchyfolge. $\square$

Anwendungen: Mit $M = \mathcal{D}$ erhalten wir wieder die Lösung der in der Einführung beschriebenen Dirichletschen Randwertaufgabe der Potentialtheorie. Weniger einfach ist die folgende Spezialisierung: Es sei fast überall $f \geq 0$ und

$$M := \mathcal{D} \cap \big\{ u \in \mathcal{H}_1(G) \;\big|\; u(x) \geq 0 \quad \text{f.ü.} \big\}$$

M erfüllt die Voraussetzungen des letzten Satzes, und wir erhalten deshalb auch eine Lösung für dieses Variationsproblem. Es handelt sich dabei um eine *Hindernisaufgabe*, die Lösung u muß stets nichtnegativ bleiben. Man beachte, daß statt (E) im allgemeinen nur die Variationsungleichung (U) gilt. Die Eulersche Differentialgleichung wird also nur in der Umgebung $U(x)$ solcher Punkte erfüllt, in der die Lösung u positiv ist. Denn nur dort findet man genügend viele Testfunktionen zum Nachweis der Differentialgleichung.

7 Über das Spektrum linearer, insbesondere kompakter linearer Abbildungen

In diesem Kapitel sollen Abbildungseigenschaften linearer Operatoren weiter untersucht werden. Insbesondere werden die Begriffe der Resolventenmenge und des Spektrums eingeführt. Diskutiert werden abgeschlossene, beschränkte und vor allem kompakte Abbildungen. In den nächsten Kapiteln werden diese Untersuchungen für unbeschränkte lineare Abbildungen weitergeführt. Ziel sind Aussagen zur Fortsetzbarkeit symmetrischer Abbildungen zu selbstadjungierten und der Beweis des Spektralsatzes für unbeschränkte selbstadjungierte Abbildungen.

7.1 Resolvente und Spektrum

Es seien $\mathcal{X}$ ein normierter Raum über $\mathbb{C}$,

$$A \;:\; \mathcal{D}(A) \subset \mathcal{X} \longrightarrow \mathcal{X}$$

eine lineare Abbildung sowie mit $\lambda \in \mathbb{C}$

$$A_\lambda := A - \lambda := A - \lambda\, id.$$

Wir beginnen mit folgender

Definition 7.1.1:

1. *Es sei*
$$\rho(A) := \left\{ \lambda \in \mathbb{C} \;\middle|\; \overline{\mathcal{R}(A_\lambda)} = \mathcal{X},\ \mathcal{N}(A_\lambda) = \{o\} \text{ und } A_\lambda^{-1} \text{ ist stetig} \right\}$$
die Resolventenmenge *von* A.
2. *Es sei* $\sigma(A) := \mathbb{C} \setminus \rho(A)$ *das* Spektrum *von* A. *Es wird disjunkt zerlegt in:*
   2.1 *das* Punktspektrum $P\sigma(A)$,
$$P\sigma(A) := \left\{ \lambda \in \sigma(A) \;\middle|\; \mathcal{N}(A_\lambda) \neq \{o\} \right\},$$
   2.2 *das* kontinuierliche Spektrum $C\sigma(A)$,
$$C\sigma(A) := \left\{ \lambda \in \sigma(A) \;\middle|\; \overline{\mathcal{R}(A_\lambda)} = \mathcal{X} \text{ und } \mathcal{N}(A_\lambda) = \{o\} \right\},$$
   2.3 *das* Residuenspektrum $R\sigma(A)$,
$$R\sigma(A) := \left\{ \lambda \in \sigma(A) \;\middle|\; \overline{\mathcal{R}(A_\lambda)} \neq \mathcal{X} \text{ und } \mathcal{N}(A_\lambda) = \{o\} \right\}.$$

Diese Unterteilung des Spektrums wird später noch verfeinert. Eine Zahl $\lambda \in P\sigma(A)$ nennt man auch *Eigenwert* und ein dazugehörendes $u \in \mathcal{N}(A_\lambda)$, $u \neq o$, *Eigenfunktion (Eigenvektor)* zu λ. Bevor wir Resultate herleiten, beginnen wir mit einigen Beispielen.

Beispiel 7.1.2: *Es seien* $\mathcal{X} = \ell^2$, $x = (\xi_1, \xi_2, \ldots) \in \ell^2$, (α_n) *eine beschränkte Folge aus* $\mathbb{C}$ *und*

$$\begin{aligned} A \;:\; & \mathcal{X} \longrightarrow \mathcal{X}, \\ & x \longmapsto (\alpha_1 \xi_1, \alpha_2 \xi_2, \ldots). \end{aligned}$$

Dann gilt

$$A_\lambda x = y \iff \forall\, n \in \mathbb{N} \quad (\alpha_n - \lambda)\xi_n = \eta_n,$$

also $P\sigma(A) = \{\alpha_1, \alpha_2, \ldots\}$. Es seien

$$y_n := (\eta_1, \eta_2, \ldots, \eta_n, 0, \ldots)$$

und für $\lambda \notin P\sigma(A)$

$$x_n := \left(\frac{\eta_1}{\alpha_1 - \lambda}, \ldots, \frac{\eta_n}{\alpha_n - \lambda}, 0, \ldots \right).$$

Dann ist $A_\lambda x_n = y_n$, und aus der Dichte der $\{y_n\}$ folgt $\overline{\mathcal{R}(A_\lambda)} = \mathcal{X}$. Mithin ist $R\sigma(A) = \emptyset$. Es sei

$$\Lambda_1 := \left\{\lambda \in \mathbb{C} \;\middle|\; \inf_{n\in\mathbb{N}} |\alpha_n - \lambda| > 0\right\}.$$

Dann ist A_λ^{-1} für $\lambda \in \Lambda_1$ beschränkt, also $\Lambda_1 \subset \rho(A)$. Umgekehrt ist

$$\Lambda_2 := \overline{\{\alpha_1, \alpha_2, \ldots\}} \setminus P\sigma(A) = C\sigma(A).$$

Für eine Teilfolge sei nämlich $\alpha_i \to \lambda \in \Lambda_2$.Dann löst

$$z_n := (0, \ldots, 0, \frac{1}{\alpha_n - \lambda}, 0, \ldots)$$

die Gleichung $A_\lambda z_n = e_n$, also $z_n = A_\lambda^{-1} e_n$, und es gilt

$$\sup_{\|y\|=1} \|A_\lambda^{-1} y\| \geq \sup_{n\in\mathbb{N}} \|z_n\| = \sup_{n\in\mathbb{N}} \frac{1}{|\alpha_n - \lambda|} = \infty.$$

A_λ^{-1} ist also unbeschränkt. □

Beispiel 7.1.3: *Es seien* $\mathcal{X} = \ell^2$, $x = (\xi_1, \xi_2, \ldots)$ *und*

$$\begin{aligned} A \;:\; \mathcal{X} &\longrightarrow \mathcal{X}, \\ x &\longmapsto (0, \xi_1, \xi_2, \ldots). \end{aligned}$$

Dann ist

$$\mathcal{R}(A) = \left\{y \in \mathcal{X} \;\middle|\; \eta_1 = 0\right\},$$

und sowohl A als auch A^{-1} sind stetig. Der Nullpunkt gehört also zum Residuenspektrum. Für $\lambda \neq 0$ folgt aus $A_\lambda x = y$

$$\begin{aligned} -\lambda\,\xi_1 &= \eta_1, \\ \xi_1 - \lambda\,\xi_2 &= \eta_2, \\ &\vdots \\ \xi_{n-1} - \lambda\,\xi_n &= \eta_n, \end{aligned}$$

also

$$\xi_n = -\left(\frac{\eta_1}{\lambda^n} + \cdots + \frac{\eta_n}{\lambda}\right).$$

Damit ist $P\sigma(A) = \emptyset$. Zur weiteren Diskussion des Spektrums betrachtet man auch die konjugierte Abbildung A'. Wir bringen ein ähnliches Beispiel am Ende dieses Abschnitts. □

Beispiel 7.1.4: *Wir betrachten den* Volterraschen Integraloperator *mit* $k = 1$, *also* $\mathcal{X} = \mathcal{C}([a,b], \mathbb{R})$ *und*

$$\begin{aligned} A \;:\; \mathcal{X} &\longrightarrow \mathcal{X}, \\ x &\longmapsto Ax(\,\cdot\,) := \int_a^{\cdot} x(s)ds. \end{aligned}$$

Wir wissen bereits, daß

$$(A - id)x = y$$

eindeutig lösbar ist und daß $\lambda = 1$ zur Resolventenmenge gehört. Dasselbe gilt für alle $\lambda \neq 0$. Man nehme nur $k := 1/\lambda$. Wir wissen ferner

$$\mathcal{R}(A) \subset \left\{x \in \mathcal{X} \;\middle|\; x(a) = 0\right\}.$$

Damit gehört der Nullpunkt zum Residuenspektrum, also $\sigma(A) = R\sigma(A) = \{0\}$. □

Beispiel 7.1.5: *Es seien* $\mathcal{X} = \mathcal{L}^2(\mathbb{R}, \mathbb{C})$, $\mathcal{D}(A) = \left\{x \in \mathcal{X} \;\middle|\; t \mapsto tx(t) \in \mathcal{X}\right\}$ *und fast überall*

$$(Ax)(t) = tx(t).$$

Offenbar ist $P\sigma(A) = \emptyset$, und wir erhalten formal aus $A_\lambda x = y$

$$x(t) = \frac{y(t)}{t-\lambda}.$$

Damit ist aber

$$\{\lambda \in \mathbb{C} \mid \operatorname{Im}\lambda \neq 0\} \subset \rho(A).$$

Wir wollen $\sigma(A) = C\sigma(A) = \mathbb{R}$ zeigen. Es seien $y \in \mathcal{X}$, $\lambda \in \mathbb{C}$, $n \in \mathbb{N}$,

$$y_{\lambda n}(t) := \begin{cases} y(t) & \text{für } |t-\lambda| > \frac{1}{n} \\ 0 & \text{sonst} \end{cases}$$

und

$$x_{\lambda n}(t) := \frac{y_{\lambda n}(t)}{t-\lambda} \in \mathcal{D}(A).$$

Dann ist $A_\lambda x_{\lambda n} = y_{\lambda n}$, und aus der Dichte der $\{y_{\lambda n}\}$ folgt $\overline{\mathcal{R}(A_\lambda)} = \mathcal{X}$. Es seien ferner

$$v_{\lambda n}(t) := \begin{cases} 1 & \text{für } t \in (\lambda + \frac{1}{n}, \lambda + \frac{1}{n} + 1) \\ 0 & \text{sonst} \end{cases}$$

und

$$u_{\lambda n}(t) := \frac{v_{\lambda n}(t)}{t-\lambda},$$

also $\|v_{\lambda n}\| = 1$ und $A_\lambda u_{\lambda n} = v_{\lambda n}$. Dann ist

$$\left\|A_\lambda^{-1} v_{\lambda n}\right\|^2 = \|u_{\lambda n}\|^2 = \int_{\lambda+\frac{1}{n}}^{\lambda+\frac{1}{n}+1} \frac{dt}{(t-\lambda)^2} = n - \frac{n}{n+1},$$

also

$$\left\|A_\lambda^{-1}\right\| = \sup_{\|y\|=1} \left\|A_\lambda^{-1} y\right\| \geq \sqrt{\frac{n^2}{n+1}} \to \infty. \qquad \square$$

Beispiel 7.1.6: *Es seien* $\mathcal{X} = \mathcal{C}([0,1], \mathbb{R})$, $\mathcal{D}(A) = \{x \in \mathcal{C}_1([0,1], \mathbb{R}) \mid x(0) = 0\} \subset \mathcal{X}$ *und*

$$\begin{aligned} A \;:\; \mathcal{D}(A) \subset \mathcal{X} &\longrightarrow \mathcal{X}, \\ x &\longmapsto x'. \end{aligned}$$

Dann ist $\mathcal{N}(A_\lambda) = \{o\}$, $\mathcal{R}(A_\lambda) = \mathcal{X}$ und

$$\left(A_\lambda^{-1} y\right)(t) = \mathrm{e}^{\lambda t} \int_0^t \mathrm{e}^{-\lambda s} y(s)\, ds.$$

Mithin ist $\|A_\lambda^{-1}\| \leq (\mathrm{e}^\lambda - 1)/\lambda$, also $\sigma(A) = \emptyset$ und $\rho(A) = \mathbb{C}$. $\square$

Wir zeigen nun erste Resultate aus der Spektraltheorie und beginnen mit

Lemma 7.1.7: *Es seien* $\mathcal{X}$ *ein Banachraum,* $A : \mathcal{D}(A) \subset \mathcal{X} \to \mathcal{X}$ *eine abgeschlossene lineare Abbildung und* $\lambda \in \rho(A)$. *Dann gilt*

$$\mathcal{D}(A_\lambda^{-1}) = \mathcal{R}(A_\lambda) = \mathcal{X}.$$

Beweis: Es seien $\lambda \in \rho(A)$, $y_n \in \mathcal{R}(A_\lambda)$ mit $y_n \to y \in \mathcal{X}$ und

$$x_n := A_\lambda^{-1} y_n.$$

Dann folgt aus der Beschränktheit von A_λ^{-1} die Konvergenz von $x_n \to x \in \mathcal{X}$ und damit $A_\lambda x = y$, weil A abgeschlossen ist. $\square$

Als nächstes zeigen wir

Satz 7.1.8: *Es seien* $\mathcal{X}$ *ein normierter Raum und* $A : \mathcal{D}(A) \subset \mathcal{X} \to \mathcal{X}$ *eine lineare Abbildung. Dann ist* $\rho(A)$ *offen.*

Beweis Es seien $\rho(A) \neq \emptyset$, $\lambda \in \rho(A)$ und

$$r(\lambda) := \left\| A_\lambda^{-1} \right\|^{-1}.$$

Wir zeigen $B\big(\lambda, r(\lambda)\big) \subset \rho(A)$:

1. Für $x \in \mathcal{D}(A)$ und $\mu \in B\big(\lambda, r(\lambda)\big)$ gilt

$$\|A_\mu x\| = \|A_\lambda x - (\mu - \lambda)x\| \geq \|A_\lambda x\| - |\lambda - \mu|\,\|x\| \geq \big(r(\lambda) - |\lambda - \mu|\big)\|x\|.$$

Deshalb existiert A_μ^{-1} und ist beschränkt,

$$\left\| A_\mu^{-1} \right\| \leq \frac{1}{r(\lambda) - |\lambda - \mu|} = \frac{\left\| A_\lambda^{-1} \right\|}{1 - |\lambda - \mu| \left\| A_\lambda^{-1} \right\|}.$$

2. Wir zeigen $\overline{\mathcal{R}(A_\mu)} = \mathcal{X}$. Der Beweis erfolgt indirekt: Zu $\mu \in B\big(\lambda, r(\lambda)\big)$ mit $\overline{\mathcal{R}(A_\mu)} \neq \mathcal{X}$ wählen wir ein $q \in \mathbb{R}$ mit

$$\frac{|\lambda - \mu|}{r(\lambda)} < q < 1.$$

Dann folgt aus dem Rieszschen Lemma die Existenz eines $y \in \mathcal{X}$ mit $\|y\| = 1$ und

$$\operatorname{dist}\Big(y, \overline{\mathcal{R}(A_\mu)}\Big) \geq q.$$

Wegen $\overline{\mathcal{R}(A_\lambda)} = \mathcal{X}$ gibt es $y_n \in \mathcal{R}(A_\lambda)$ mit $y_n \to y$. Damit erhalten wir

$$\begin{aligned} q \leq \left\| y - A_\mu(A_\lambda^{-1} y_n) \right\| = \left\| y - \big(A_\lambda + (\lambda - \mu)\big) A_\lambda^{-1} y_n \right\| = \left\| y - y_n - (\lambda - \mu) A_\lambda^{-1} y_n \right\| \;\leq \\ \leq \|y - y_n\| + \frac{|\lambda - \mu|}{r(\lambda)} \|y_n\| \to \frac{|\lambda - \mu|}{r(\lambda)}, \end{aligned}$$

also der Widerspruch

$$q \leq \frac{|\lambda - \mu|}{r(\lambda)} < q. \qquad \square$$

Satz 7.1.9: *Es seien $\mathcal{X}$ ein Banachraum und $A \in \mathcal{CL}(\mathcal{X})$. Dann ist*

$$\big\{ \lambda \in \mathbb{C} \;\big|\; |\lambda| > \|A\| \big\} \subset \rho(A).$$

Der Beweis folgt aus dem Satz über die Neumannsche Reihe (Satz 1.3.14). Es ist nämlich

$$A - \lambda = -\lambda(id - K)$$

mit $K := A/\lambda$ und $\|K\| < 1$. Die Abbildung A_λ ist deshalb ein linearer Homöomorphismus. $\square$

Für $\lambda \in \rho(A)$, $A : \mathcal{D}(A) \subset \mathcal{X} \to \mathcal{X}$, definieren wir nun

$$(A - \lambda)^{-1} =: R(A, \lambda) =: R(\lambda)$$

und nennen diese Abbildung

$$\begin{aligned} R(A, \cdot) \;\; : \;\; \rho(A) &\longrightarrow \mathcal{CL}\big(\mathcal{R}(A - \cdot\, id), \mathcal{X}\big), \\ \lambda &\longmapsto R(A, \lambda) \end{aligned}$$

Resolvente von A. Ist $\mathcal{X}$ vollständig und A abgeschlossen, dann ist nach Lemma 7.1.7 $\mathcal{R}(A_\lambda) = \mathcal{X}$, also $R(\lambda) \in \mathcal{CL}(\mathcal{X})$. Solche Abbildungen werden uns im folgenden besonders interessieren. Als erstes zeigen wir

Die Resolventenformel: *Für alle $\lambda \in \rho(A)$ sei $\mathcal{R}(A_\lambda) = \mathcal{X}$. Dann gilt:*

1. $R(\lambda) - R(\mu) = (\lambda - \mu) R(\lambda) R(\mu)$.
2. $R(\lambda) R(\mu) = R(\mu) R(\lambda)$.

Der Beweis der ersten Aussage folgt aus

$$R(\lambda) = R(\lambda)A_\mu R(\mu) = R(\lambda)\big((\lambda-\mu)id + A_\lambda\big)R(\mu) = (\lambda-\mu)R(\lambda)R(\mu) + R(\mu).$$

Durch Vertauschen von λ und μ erhält man daraus die zweite. □

Es gilt auch

Satz 7.1.10: *Für alle $\lambda \in \rho(A)$ sei $\mathcal{R}(A_\lambda) = \mathcal{X}$. Dann läßt sich*

$$R(A,\,\cdot\,) \;:\; \rho(A) \longrightarrow \mathcal{CL}(\mathcal{X})$$

lokal in eine Potenzreihe entwickeln mit Koeffizienten aus $\mathcal{CL}(\mathcal{X})$. Diese Reihe konvergiert bezüglich der Operatornorm. Mit $r(\lambda) := \|R(\lambda)\|^{-1}$ gilt

$$\forall\,\lambda \in \rho(A) \quad \forall\,\mu \in B\big(\lambda, r(\lambda)\big) \qquad R(\mu) = \sum_{n=0}^{\infty} (\mu-\lambda)^n \big(R(\lambda)\big)^{n+1}.$$

Beweis: Aus der Resolventenformel erhält man durch vollständige Induktion

$$R(\mu) = \sum_{j=0}^{n} (\mu-\lambda)^j \big(R(\lambda)\big)^{j+1} + \big((\mu-\lambda)R(\lambda)\big)^{n+1} R(\mu).$$

Wegen $|\mu-\lambda| < r(\lambda)$ ist

$$\|(\mu-\lambda)R(\lambda)\| < 1,$$

deshalb konvergiert das Restglied in der Operatornorm gegen Null. □

Folgerung 7.1.11: *Es sei $f' \in \mathcal{CL}(\mathcal{X})'$. Dann läßt sich*

$$\begin{aligned} \rho(A) &\longrightarrow \mathbb{C} \\ \lambda &\longmapsto f'R(\lambda) \end{aligned}$$

lokal in eine Potenzreihe mit Koeffizienten aus $\mathbb{C}$ entwickeln,

$$f'R(\mu) = \sum_{n=0}^{\infty} (\mu-\lambda)^n f'\big(R(\lambda)\big)^{n+1}.$$

Die Abbildung $f'R(\mu)$ ist auf $\rho(A)$ holomorph, und es gilt

$$\left(\frac{d}{d\mu}\right)^n f'R(\mu)\bigg|_{\mu=\lambda} = n!\, f'\big(R(\lambda)\big)^{n+1}.$$

Wir zeigen nun

Satz 7.1.12: *Es seien $\mathcal{X}$ ein Banachraum und $A \in \mathcal{CL}(\mathcal{X})$. Dann gilt*

1. Es existiert der Spektralradius $r_\sigma(A) := \lim_{n\to\infty} \|A^n\|^{1/n}$.

2. Es ist $r_\sigma(A) \le \|A\|$.

3. Aus $|\lambda| > r_\sigma(A)$ folgt $\lambda \in \rho(A)$ und in $\mathcal{CL}(\mathcal{X})$

$$(A-\lambda)^{-1} = -\frac{1}{\lambda}\sum_{n=0}^{\infty}\left(\frac{A}{\lambda}\right)^n.$$

4. Es ist

$$r_\sigma(A) = \sup_{\lambda\in\sigma(A)} |\lambda|.$$

Beweis:

1. Es sei

$$r := \inf_{n\ge 1} \big\|A^n\big\|^{1/n}.$$

Dann gilt

$$\forall\, \varepsilon > 0 \quad \exists\, m \in \mathbb{N} \qquad \left\|A^m\right\|^{1/m} \leq r + \varepsilon$$

und

$$\forall\, n \in \mathbb{N} \quad \exists\, p, q \in \mathbb{N}_0,\, q \leq m-1, \qquad n = pm + q.$$

Wegen $A^n = (A^m)^p A^q$ und $\|AB\| \leq \|A\| \cdot \|B\|$ folgt deshalb

$$\left\|A^n\right\|^{1/n} \leq \left\|A^m\right\|^{\frac{p}{n}} \cdot \left\|A\right\|^{\frac{q}{n}} \leq (r+\varepsilon)^{\frac{mp}{n}} \cdot \left\|A\right\|^{\frac{q}{n}} \leq (r+\varepsilon)\Big(\frac{\|A\|}{r+\varepsilon}\Big)^{q/n}.$$

Weil der letzte Faktor für $n \to \infty$ gegen Eins konvergiert, erhalten wir daraus

$$\limsup_{n\to\infty} \|A^n\|^{1/n} \leq r = \inf_{n\geq 1} \left\|A^n\right\|^{1/n} \leq \liminf_{n\to\infty} \|A^n\|^{1/n}.$$

Mithin existiert

$$\lim_{n\to\infty} \|A^n\|^{1/n} = r.$$

2. Aus $\|A^n\| \leq \|A\|^n$ folgt $r_\sigma(A) \leq \|A\|$.

3. Für $\varepsilon > 0$ sei $|\lambda| \geq r_\sigma(A) + \varepsilon$. Dann gilt

$$\exists\, n(\varepsilon) \in \mathbb{N} \quad \forall\, n \geq n(\varepsilon) \qquad \left\|A^n\right\| \leq \Big(r_\sigma + \frac{\varepsilon}{2}\Big)^n$$

und

$$\left\|\Big(\frac{A}{\lambda}\Big)^n\right\| \leq \Big(\underbrace{\frac{r_\sigma + \frac{\varepsilon}{2}}{r_\sigma + \varepsilon}}_{=:q}\Big)^n = q^n \quad \text{mit } q < 1.$$

Deshalb konvergiert

$$\sum_{n=0}^{\infty} \Big(\frac{A}{\lambda}\Big)^n,$$

und wie im Beweis zum Satz über die Neumannsche Reihe erhält man

$$(A-\lambda)^{-1} = -\frac{1}{\lambda}\Big(id - \frac{A}{\lambda}\Big)^{-1} = -\frac{1}{\lambda}\sum_{n=0}^{\infty}\Big(\frac{A}{\lambda}\Big)^n. \qquad (*)$$

4. Für $\lambda \in \sigma(A)$ ist daher $|\lambda| \leq r_\sigma(A)$, also

$$k := \sup_{\lambda\in\sigma(A)} |\lambda| \leq r_\sigma(A).$$

Gleichung $(*)$ liefert die Laurent-Reihenentwicklung von $(A-\lambda)^{-1}$. Sie konvergiert für alle λ mit $|\lambda| > k$. Es sei nun $\varepsilon > 0$ fest und $|\lambda| := k + \varepsilon$. Dann folgt $\|(A/\lambda)^n\| \to 0$ für $n \to \infty$. Erst recht gilt deshalb

$$\exists\, n(\varepsilon) \in \mathbb{N} \quad \forall\, n \geq n(\varepsilon) \qquad \left\|\Big(\frac{A}{k+\varepsilon}\Big)^n\right\| \leq 1$$

oder

$$\left\|A^n\right\|^{1/n} \leq k + \varepsilon,$$

also

$$r_\sigma(A) = \lim_{n\to\infty} \left\|A^n\right\|^{1/n} \leq k.$$

Daher ist

$$k = \sup_{\lambda\in\sigma(A)} |\lambda| = r_\sigma(A). \qquad \square$$

Beispiel 7.1.13: *Wir betrachten wieder den* Volterraschen Integraloperator *aus §1.3.4,*

$$(Kx)(s) = \int_a^s k(s,t)x(t)\,dt.$$

Aus $\|K^n\| \leq M^n(b-a)^n/n!$ folgt

$$\lim_{n\to\infty} \left\|K^n\right\|^{1/n} \leq M(b-a) \lim_{n\to\infty} \left(\frac{1}{n!}\right)^{1/n} = 0,$$

also ist $r_\sigma(K) = 0$. Das erhält man auch aus $\sigma(K) = R\sigma(K) = \{o\}$. □

Beispiel 7.1.6 hat bereits gezeigt, daß es unbeschränkte Abbildungen mit leerem Spektrum gibt. Für beschränkte Abbildungen gilt dies nicht, dazu zeigen wir

Satz 7.1.14: *Es seien $\mathcal{X}$ ein Banachraum über $\mathbb{C}$ und $A \in \mathcal{CL}(\mathcal{X})$. Dann ist $\sigma(A) \neq \emptyset$.*

Beweis: Für $f' \in \mathcal{CL}(\mathcal{X})'$ sei

$$\begin{aligned} F \;:\; \rho(A) &\longrightarrow \mathbb{C}, \\ \lambda &\longmapsto f'R(\lambda). \end{aligned}$$

Dann gilt für alle $\lambda \in \mathbb{C}$ mit $|\lambda| > \|A\|$

$$|F(\lambda)| \leq \|f'\| \cdot \|R(\lambda)\| = \|f'\| \cdot \left\| \frac{1}{\lambda}\Big(id - \frac{A}{\lambda}\Big)^{-1} \right\| \leq \frac{\|f'\|}{|\lambda| - \|A\|}.$$

Es sei nun $\rho(A) = \mathbb{C}$. Dann ist $F(\lambda)$ auf $\mathbb{C}$ definiert, holomorph und beschränkt, nach dem Satz von Liouville also konstant. Wegen $F(\lambda) \to 0$ für $|\lambda| \to \infty$ folgt daher für alle $\lambda \in \mathbb{C}$

$$F(\lambda) = f'R(\lambda) = 0.$$

Diese Aussage gilt für alle $f' \in \mathcal{CL}(\mathcal{X})'$, mithin ist $R(\lambda) = O$ für alle $\lambda \in \mathbb{C}$ und damit

$$id = (A - \lambda)R(\lambda) = O.$$

Das ist natürlich ein Widerspruch. □

Es ist oft zweckmäßig, eine Abbildung zusammen mit ihrer konjugierten zu betrachten. Wir beginnen mit einer Reihe von Beispielen.

Beispiel 7.1.15: *Es seien $\mathcal{X} = \ell^1$, $x = (\xi_1, \xi_2, \ldots)$ und $x' = (\xi_1', \xi_2', \ldots)$. Wir betrachten*

$$\begin{aligned} A \;:\; \ell^1 &\longrightarrow \ell^1, \\ x &\longmapsto (\xi_2, \xi_3, \ldots). \end{aligned}$$

Es ist $\left(\ell^1\right)' \cong \ell^\infty$, und aus $x'x = \sum_{i=1}^\infty \xi_i'\xi_i$ folgt

$$x'Ax = \xi_1'\xi_2 + \xi_2'\xi_3 + \cdots = (A'x')x,$$

also

$$\begin{aligned} A' \;:\; \ell^\infty &\longrightarrow \ell^\infty, \\ x' &\longmapsto (0, \xi_1', \xi_2', \ldots). \end{aligned}$$

Folgendes wollen wir zeigen:

	σ	$P\sigma$	$C\sigma$	$R\sigma$
A	$\bar{B}(0,1)$	$B(0,1)$	$\partial B(0,1)$	$\emptyset$
A'	$\bar{B}(0,1)$	$\emptyset$	$\emptyset$	$\bar{B}(0,1)$

.

Offenbar ist $\|A\| = 1$ und nach Satz 1.5.4 damit auch $\|A'\| = 1$. Mithin ist

$$r_\sigma(A) \leq 1 \quad \text{und} \quad r_\sigma(A') \leq 1$$

sowie

$$\left\{\lambda \;\middle|\; |\lambda| > 1\right\} \subset \rho(A) \quad \text{und} \quad \left\{\lambda \;\middle|\; |\lambda| > 1\right\} \subset \rho(A').$$

Im folgenden diskutieren wir deshalb nur noch Werte von $\lambda \in \bar{B}(0,1)$.

1. Es ist $P\sigma(A) = B(0,1)$: Es sei nämlich $A_\lambda x = o$. Dann folgt

$$\left.\begin{array}{r} \xi_2 - \lambda\xi_1 = 0 \\ \xi_3 - \lambda\xi_2 = 0 \\ \vdots \end{array}\right\} \quad \rightsquigarrow \quad x = \xi_1 \cdot (1, \lambda, \lambda^2, \dots)$$

und $\|x\| = |\xi_1|\left\{1 + |\lambda| + |\lambda|^2 + \cdots\right\}$. Deshalb ist $x \in \ell^1$ nur für $|\lambda| < 1$ möglich.

2. Es ist $\sigma(A) = \bar{B}(0,1)$, weil $\sigma(A)$ abgeschlossen ist.

3. Es ist $P\sigma(A') = \emptyset$: Es sei nämlich $A'_\lambda x = o$. Dann folgt

$$\left.\begin{array}{r} -\lambda\xi_1 = 0 \\ \xi_1 - \lambda\xi_2 = 0 \\ \vdots \end{array}\right\} \quad \rightsquigarrow \quad x = o.$$

4. Es ist $R\sigma(A) = \emptyset$: Es sei nämlich $\overline{\mathcal{R}(A_\lambda)} \neq \ell^1$ für $\lambda \in \partial B(0,1)$. Dann gibt es ein $y' \neq o$ mit

$$\forall\, x \in \ell^1 \qquad 0 = y' A_\lambda x = (A'_\lambda y')x.$$

Daraus folgt $A'_\lambda y' = o$ im Widerspruch zur dritten Aussage. Mithin ist $\partial B(0,1) = C\sigma(A)$.

5. Es ist $R\sigma(A') = \bar{B}(0,1)$. Das zeigen wir in zwei Schritten:

a) Es sei $\lambda \in B(0,1)$. Dann gibt ein $x_\lambda \neq o$ mit $A_\lambda x_\lambda = o$, und es ist

$$\forall\, y' \in \ell^\infty \qquad 0 = y' A_\lambda x_\lambda = (A'_\lambda y')x_\lambda.$$

Nun folgt aus dem Satz von Hahn-Banach die Existenz eines $z' \in \ell^\infty$ mit

$$z' x_\lambda = \|x_\lambda\| \neq 0.$$

Mithin ist $z' \notin \overline{\mathcal{R}(A'_\lambda)}$.

b) Es sei $\lambda \in \partial B(0,1)$. Dann folgt aus $a, b \in \ell^\infty$, $a = A'_\lambda b$,

$$\left.\begin{array}{l} \alpha_1 = -\lambda\beta_1 \\ \alpha_2 = \beta_1 - \lambda\beta_2 \\ \vdots \end{array}\right\} \quad \rightsquigarrow \quad \begin{array}{l} \beta_1 = -\bar{\lambda}\alpha_1 \\ \beta_2 = -\bar{\lambda}\alpha_2 + \bar{\lambda}\beta_1 = -\bar{\lambda}^3\left(\lambda\alpha_1 + \lambda^2\alpha_2\right) \\ \vdots \end{array}$$

allgemein

$$\beta_n = -\bar{\lambda}^{n+1}\Big(\sum_{k=1}^{n} \lambda^k \alpha_k\Big).$$

Es sei nun $g = (\gamma_1, \gamma_2, \cdots) := (\bar{\lambda}, \bar{\lambda}^2, \dots) \in \ell^\infty$. Dann zeigen wir

$$\mathcal{R}(A'_\lambda) \bigcap B\big(g, \tfrac{1}{2}\big) = \emptyset.$$

Es sei nämlich $d \in \ell^\infty$ mit $\|d - g\|_{\ell^\infty} < \frac{1}{2}$. Dann folgt

$$\operatorname{Re}\lambda^n \delta_n = \operatorname{Re}\lambda^n \gamma_n + \operatorname{Re}\lambda^n(\delta_n - \gamma_n) \geq 1 - \tfrac{1}{2} = \tfrac{1}{2},$$

und es gibt kein $e \in \ell^\infty$ mit $A'_\lambda e = d$. Aus $A'_\lambda e = d$ würde nämlich

$$|\varepsilon_n| = \Big|\sum_{k=1}^{n} \lambda^k \delta_k\Big| \geq \operatorname{Re}\sum_{k=1}^{n} \lambda^k \delta_k \geq \frac{n}{2}$$

im Widerspruch zu $\|e\|_{\ell^\infty} < \infty$ folgen.

Damit sind alle Behauptungen bewiesen. □

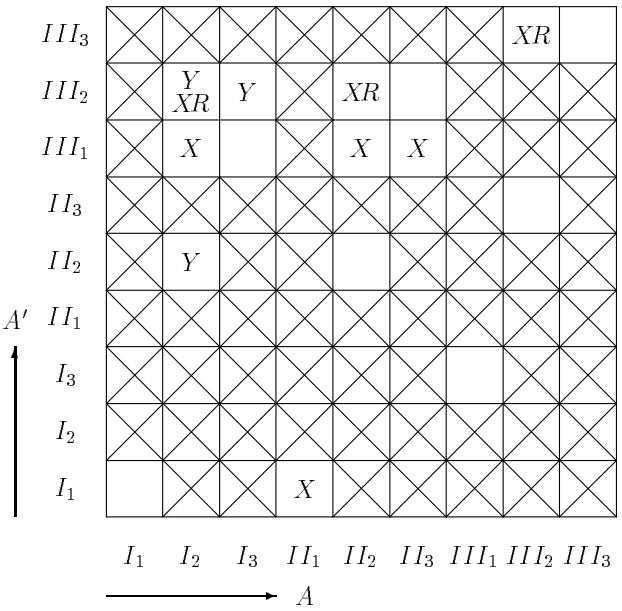
Es seien nun $\mathcal{X}, \mathcal{Y}$ normierte Räume und

$$A \;:\; \mathcal{D}(A) \subset \mathcal{X} \longrightarrow \mathcal{Y}$$

eine lineare Abbildung mit dichtem Definitionsbereich, also $\overline{\mathcal{D}(A)} = \mathcal{X}$. Dann existiert

$$A' \;:\; \mathcal{D}(A') \subset \mathcal{Y}' \longrightarrow \mathcal{X}'$$

mit $y' \mapsto y' \circ A$ und ist nach Satz 5.3.1 abgeschlossen. Ist $\mathcal{R}(A) = \mathcal{Y}$ und existiert $A^{-1} \in \mathcal{CL}(\mathcal{Y}, \mathcal{X})$, dann ist auch $\mathcal{R}(A') = \mathcal{X}'$, und es existiert $A'^{-1} \in \mathcal{CL}(\mathcal{X}', \mathcal{Y}')$. Existiert A^{-1} nicht, dann ist $\overline{\mathcal{R}(A')} \neq \mathcal{X}'$ und umgekehrt. Allgemein ist es nun interessant herauszuarbeiten, wie sich die verschiedenen *Zustände* von A und A' gegenseitig bedingen. Es gilt das folgende Diagramm, das *Taylor & Lay [1980], S. 238*, entnommen ist. Die einzelnen Nachweise wollen wir in dieser Vorlesung nicht bringen, sondern es mit wenigen Bemerkungen genug sein lassen. Es sei vielmehr auf das Buch von *Taylor & Lay [1980]* und die dort gegebenen Zitate verwiesen.



In diesem Diagramm bedeuten

I	$\mathcal{R}(A) = \mathcal{Y}$	1	A^{-1} existiert und ist stetig
II	$\overline{\mathcal{R}(A)} = \mathcal{Y}$ und $\mathcal{R}(A) \neq \mathcal{Y}$	2	A^{-1} existiert und ist nicht stetig
III	$\overline{\mathcal{R}(A)} \neq \mathcal{Y}$	3	A^{-1} existiert nicht

und Entsprechendes für A'. Durchgestrichene Zustände können nicht auftreten, und Markierungen mit X, Y, XR stehen für

X	nicht möglich, wenn $\mathcal{X}$ vollständig und A abgeschlossen ist.
Y	nicht möglich, wenn $\mathcal{Y}$ vollständig ist.
XR	nicht möglich, wenn $\mathcal{X}$ reflexiv und A abgeschlossen ist.

Bevor wir weitermachen, definieren wir den *Annihilator* $M^\perp$ einer Teilmenge $M \subset \mathcal{X}$, und zwar analog zum orthogonalen Komplement in Hilberträumen.

Definition 7.1.16: *Es sei $M \subset \mathcal{X}$ eine Teilmenge des normierten Raumes $\mathcal{X}$. Dann ist*

$$M^\perp := \{x' \in \mathcal{X}' \mid \forall\, x \in M \quad x'x = 0\}$$

der Annihilator von M. Es sei $N \subset \mathcal{X}'$ eine Teilmenge von $\mathcal{X}'$. Dann ist

$$N^\perp := \{x \in \mathcal{X} \mid \forall\, x' \in N \quad x'x = 0\}$$

der Annihilator von N.

Wir zeigen

Satz 7.1.17: *Es seien $\mathcal{X}$ ein normierter Raum und $M \subset \mathcal{X}$ ein Teilraum. Dann gilt*

1. $M^\perp$ ist ein abgeschlossener Teilraum von $\mathcal{X}'$, und M' ist zu $\mathcal{X}'/M^\perp$ kongruent,

$$M' \cong \mathcal{X}'/M^\perp.$$

2. Es sei $M \subset \mathcal{X}$ ein abgeschlossener Teilraum. Dann ist

$$(\mathcal{X}/M)' \cong M^\perp.$$

Beweis:

1. Daß $M^\perp$ ein abgeschlossener Teilraum von $\mathcal{X}'$ ist, ist klar. Es sei $m' \in M'$ gegeben. Dann existiert eine Fortsetzung $x' \in \mathcal{X}'$ von m'. Es sei

$$\begin{aligned} A \;:\; & M' \longrightarrow \mathcal{X}'/M^\perp, \\ & m' \longmapsto [x'] \in \mathcal{X}'/M^\perp. \end{aligned}$$

Diese Abbildung ist wohldefiniert. Ist nämlich $y' \in \mathcal{X}'$ eine andere Fortsetzung von m', dann folgt $x' - y' \in M^\perp$, also $[x'] = [y']$. Weil jedes $y' \in [x'] = Am'$ eine Fortsetzung von m' ist, gilt

$$\|m'\| \leq \inf\left\{\|y'\| \;\middle|\; y' \in [x'] = Am'\right\} = \|Am'\|.$$

Andererseits folgt aus dem Satz von Hahn-Banach die Existenz einer Fortsetzung y' mit $\|y'\| = \|m'\|$. Also ist $\|m'\| = \|Am'\|$. Wegen $x' \,|\, M \in M'$ ist A ein isometrischer Isomorphismus von M' auf $\mathcal{X}'/M^\perp$.

2. Es sei φ die stetige kanonische Abbildung von $\mathcal{X}$ auf $\mathcal{X}/M$ (vgl. Lemma 1.3.5). Es sei $z' \in (\mathcal{X}/M)'$ gegeben und $Az' := z'\varphi$. Dann ist $A \in \mathcal{L}\big((\mathcal{X}/M)', \mathcal{X}'\big)$, und aus $\|\varphi\| = 1$ folgt

$$\|Az'\| = \sup_{x \in S} \big|(Az')x\big| = \sup_{x \in S} \big|z'\varphi(x)\big| = \|z'\|.$$

Mithin ist A eine lineare Isometrie. Es gilt auch

$$\forall\, m \in M \quad \forall\, z' \in (\mathcal{X}/M)' \qquad (Az')m = z'\varphi(m) = 0,$$

also $\mathcal{R}(A) \subset M^\perp$. Es folgt die andere Richtung: Es sei $x' \in M^\perp$. Dann ist $z'[x] := x'x$ ein wohldefiniertes Funktional auf $\mathcal{X}/M$ mit

$$\sup_{[x] \in S} \big|z'[x]\big| = \sup_{x \in S} |x'x| = \|x'\|.$$

Es ist also $z' \in (\mathcal{X}/M)'$ und $Az' = x' \in \mathcal{R}(A)$. Die Abbildung A ist daher eine isometrische lineare Abbildung von $(\mathcal{X}/M)'$ auf $M^\perp$. □

Wir zeigen auch

Lemma 7.1.18: *Folgendes gilt:*

1. Es sei $M \subset \mathcal{X}$ ein Teilraum. Dann ist

$$\bar{M} = (M^\perp)^\perp.$$

2. Es sei $N \subset \mathcal{X}'$ ein Teilraum. Dann ist

$$\bar{N} \subset (N^\perp)^\perp.$$

In reflexiven Räumen ist $\bar{N} = (N^\perp)^\perp$.

Beweis:

1. Aus $m \in M$ erhalten wir

$$\forall\, x' \in M^\perp \qquad x'm = 0,$$

also $m \in (M^\perp)^\perp$, und daraus $\bar{M} \subset (M^\perp)^\perp$. Es sei $x \in (M^\perp)^\perp \backslash \bar{M}$, $x \neq o$, dann gibt es ein $x' \in M^\perp$ mit $\|x'\| = 1$ und $x'x \neq 0$. Aus $x \in (M^\perp)^\perp$ folgt aber $x'x = 0$, mithin ist $\bar{M} = (M^\perp)^\perp$.

2. Aus $n' \in N$ erhält man

$$\forall\, x \in N^\perp \qquad n'x = 0,$$

also $n' \in (N^\perp)^\perp$, und damit $\bar{N} \subset (N^\perp)^\perp$. Hier gilt das Gleichheitszeichen nur in reflexiven Räumen. □

Es seien nun wieder $\mathcal{X}, \mathcal{Y}$ normierte Räume und A eine in $\mathcal{X}$ dicht definierte lineare Abbildung. Die konjugierte Abbildung $A' : \mathcal{D}(A') \subset \mathcal{Y}' \to \mathcal{X}'$ soll also existieren. Dann zeigen wir

Satz 7.1.19: *Es gilt*

1. $\overline{\mathcal{R}(A)}^\perp = \mathcal{R}(A)^\perp = \mathcal{N}(A')$.
2. $\overline{\mathcal{R}(A)} = \mathcal{N}(A')^\perp$.
3. $\overline{\mathcal{R}(A)} = \mathcal{Y} \iff (A')^{-1}$ *existiert.*

Der Beweis folgt unmittelbar aus Lemma 7.1.18 und den Definitionen. Der Nachweis der entsprechenden Aussagen für A' ist etwas komplizierter.

Satz 7.1.20: *Es sei A abgeschlossen. Dann gilt*

1. $\mathcal{R}(A')^\perp \cap \mathcal{D}(A) = \mathcal{N}(A)$.
2. $\overline{\mathcal{R}(A')} \subset \mathcal{N}(A)^\perp$.
3. $\mathcal{R}(A')$ *ist eine totale Teilmenge von* $\mathcal{X}' \iff A^{-1}$ *existiert.*

Beweis:

1. Aus $x \in \mathcal{N}(A)$ folgt

$$\forall\, y' \in \mathcal{D}(A') \qquad 0 = y'Ax = (A'y')x,$$

also $\mathcal{N}(A) \subset \mathcal{R}(A')^\perp \cap \mathcal{D}(A)$.
Es sei $x \in \mathcal{R}(A')^\perp \cap \mathcal{D}(A)$. Dann folgt

$$\forall\, y' \in \mathcal{D}(A') \qquad y'Ax = (A'y')x = 0.$$

Nun ist $\mathcal{D}(A')$ nach Satz 5.3.2 eine totale Teilmenge von $\mathcal{Y}'$, also $x \in \mathcal{N}(A)$.

2. Nach Lemma 7.1.18 und der ersten Aussage ist

$$\overline{\mathcal{R}(A')} \subset (\mathcal{R}(A')^\perp)^\perp = \mathcal{N}(A)^\perp.$$

3. Es sei $\mathcal{R}(A')$ total. Dann ist $\mathcal{R}(A')^\perp = \{o\}$, also $\mathcal{N}(A) = \{o\}$. Mithin existiert A^{-1}. Umgekehrt, A^{-1} möge existieren. Wir wählen $x_0 \neq o$ in $\mathcal{X}$. Dann liegt $\langle x_0, o\rangle$ nicht im Graphen von A, und weil A abgeschlossen ist, gibt es ein $z' \in (\mathcal{X} \times \mathcal{Y})'$ mit

$$z'\langle x_0, o\rangle \neq 0 \quad \text{und} \quad \forall\, x \in \mathcal{D}(A) \quad z'\langle x, Ax\rangle = 0.$$

Wir definieren $x' \in \mathcal{X}'$ und $y' \in \mathcal{Y}'$ vermöge

$$x'x := z'\langle x, o\rangle \quad \text{und} \quad y'y := z'\langle o, y\rangle.$$

Dann ist

$$x'x_0 = z'\langle x_0, o\rangle \neq 0$$

und

$$\forall\, x \in \mathcal{D}(A) \qquad 0 = z'\langle x, Ax\rangle = x'x + y'Ax,$$

also $y' \in \mathcal{D}(A')$ und

$$x' = -A'y' \in \mathcal{R}(A').$$

Damit ist $\mathcal{R}(A')$ total. □

Es sei noch bemerkt, daß in der zweiten Aussage die Gleichheit gilt, wenn $\mathcal{X}$ reflexiv und $\mathcal{D}(A) = \mathcal{X}$ ist.

Als nächstes zeigen wir

Satz 7.1.21:

$$\mathcal{R}(A') = \mathcal{X}' \iff A^{-1} \text{ existiert und ist stetig.}$$

Zum Beweis verwenden wir

Lemma 7.1.22: *A besitze keine stetige Inverse. Dann gibt es eine Folge (x_n) in $\mathcal{D}(A)$ mit $\|x_n\| \to \infty$ und $Ax_n \to o$.*

Beweis: Nach Satz 1.3.7 existiert $A^{-1} \in \mathcal{L}\big(\mathcal{R}(A), \mathcal{D}(A)\big)$ und ist stetig genau dann, wenn

$$\exists\, m > 0 \quad \forall\, x \in \mathcal{D}(A) \qquad m\,\|x\| \le \|Ax\|$$

gilt. Dabei können wir uns auf $x \in S_{\mathcal{X}} \cap \mathcal{D}(A)$ beschränken. Es gilt also

$$\forall\, n \in \mathbb{N} \quad \exists\, s_n \in S_{\mathcal{X}} \cap \mathcal{D}(A) \qquad \tfrac{1}{n} > \|As_n\|.$$

Es sei nun

$$x_n := \begin{cases} s_n/\sqrt{\|As_n\|}, & \text{falls } As_n \neq o \text{ ist} \\ n s_n, & \text{falls } As_n = o \text{ ist.} \end{cases}$$

Dann ist (x_n) die gewünschte Folge. □

Beweis von Satz 7.1.21:

$\Rightarrow$: Es sei $\mathcal{R}(A') = \mathcal{X}'$, und A habe keine stetige Inverse. Wir wählen (x_n) wie in Lemma 7.1.22. Dann gilt

$$\forall\, y' \in \mathcal{Y}' \qquad y'Ax_n \to 0.$$

Wegen $\mathcal{R}(A') = \mathcal{X}'$ folgt daraus

$$\forall\, x' \in \mathcal{X}' \qquad x'x_n \to 0.$$

Dann ist (x_n) aber beschränkt, im Widerspruch zu $\|x_n\| \to \infty$.

$\Leftarrow$: A besitze eine stetige Inverse. Für festes $x' \in \mathcal{X}'$ ist dann

$$y \longmapsto x'(A^{-1}y)$$

ein stetiges lineares Funktional auf $\mathcal{R}(A)$. Wir setzen es auf $\mathcal{Y}$ stetig fort. Es gibt also ein $y' \in \mathcal{Y}'$ mit

$$\forall\, y \in \mathcal{R}(A) \qquad y'y = x'(A^{-1}y).$$

Daraus folgt

$$\forall\, x \in \mathcal{D}(A) \qquad y'Ax = x'x.$$

Deshalb ist $y' \in \mathcal{D}(A')$ und $A'y' = x'$. □

Wir beschließen diesen Abschnitt mit dem

Satz vom abgeschlossenen Wertebereich: *Es seien $\mathcal{X}, \mathcal{Y}$ Banachräume und A eine abgeschlossene, in $\mathcal{X}$ dicht definierte lineare Abbildung in $\mathcal{Y}$. Dann sind folgende Aussagen äquivalent:*

(1) $\mathcal{R}(A)$ *ist abgeschlossen.*

(2) $\mathcal{R}(A')$ *ist abgeschlossen.*

(3) $\mathcal{R}(A) = \mathcal{N}(A')^{\perp}$.

(4) $\mathcal{R}(A') = \mathcal{N}(A)^{\perp}$.

(5) $\mathcal{R}(A')$ *ist schwach-stern abgeschlossen.*

Beweis:

(1) $\Leftrightarrow$ (3) gilt nach Satz 7.1.19.2.

(1) $\Rightarrow$ (4): Es seien $\mathcal{R}(A)$ abgeschlossen und $x' \in \mathcal{N}(A)^{\perp}$. Dann gibt es ein $z' \in (\mathcal{X}/\mathcal{N}(A))'$ mit

$$z'[x] = x'x,$$

man vergleiche den Beweis zu Satz 7.1.17.2. Es sei $\hat{A}$ die bijektive Abbildung, die von A induziert wird. Weil $\mathcal{R}(\hat{A}) = \mathcal{R}(A)$ abgeschlossen ist, besitzt $\hat{A}$ eine beschränkte Inverse

$$\hat{A}^{-1} \;:\; \mathcal{R}(A) \to \mathcal{X}/\mathcal{N}(A).$$

Deshalb ist $z' \circ \hat{A}^{-1} \in \mathcal{R}(A)'$, nämlich

$$\forall\, x \in \mathcal{D}(A) \qquad (z' \circ \hat{A}^{-1})Ax = x'x.$$

Es sei y' eine stetige Fortsetzung von $z' \circ \hat{A}^{-1}$ auf $\mathcal{Y}'$. Dann folgt $y' \in \mathcal{D}(A')$ und $A'y' = x'$. Es ist also $\mathcal{N}(A)^\perp \subset \mathcal{R}(A')$. Aus Satz 7.1.20,2 erhält man dann $\mathcal{N}(A)^\perp = \mathcal{R}(A')$.

(4) $\Rightarrow$ (2) folgt, weil $\mathcal{N}(A)^\perp$ abgeschlossen ist.

(2) $\Rightarrow$ (1): Es seien $\mathcal{R}(A')$ abgeschlossen und $A_0 : \mathcal{D}(A) \to \overline{\mathcal{R}(A)}$ mit $A_0 x = Ax$. Diese Abbildung ist ebenfalls abgeschlossen. Weil jedes Element von $\overline{\mathcal{R}(A)}'$ eine stetige Fortsetzung auf $\mathcal{Y}'$ besitzt, ist $\mathcal{R}(A_0') = \mathcal{R}(A')$, also abgeschlossen. Dann ist auch $\mathcal{R}(A_0)$ abgeschlossen, das folgt aus Lemma 7.1.7 und dem Zustandsdiagramm. Damit ist $\mathcal{R}(A)$ abgeschlossen.

Den Nachweis der Äquivalenz mit (5) findet man in *Taylor & Lay [1980], S. 162.* □

7.2 Kompakte lineare Abbildungen

Wir greifen nun eine wichtige Klasse linearer Abbildungen heraus, nämlich die kompakten.

Definition 7.2.1: *Es seien $\mathcal{X}, \mathcal{Y}$ normierte Räume und $S_\mathcal{X}$ die Einheitssphäre in $\mathcal{X}$. Dann heißt $K \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$* kompakt, *wenn $KS_\mathcal{X}$ in $\mathcal{Y}$ relativ kompakt ist. Es sei*

$$\mathcal{KL}(\mathcal{X}, \mathcal{Y}) := \left\{ K \in \mathcal{L}(\mathcal{X}, \mathcal{Y}) \;\middle|\; K \text{ ist kompakt} \right\}.$$

Als erstes zeigen wir

Satz 7.2.2: *Es seien $\mathcal{W}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}$ normierte Räume. Dann gilt:*

1. *$\mathcal{KL}(\mathcal{X}, \mathcal{Y})$ ist ein linearer Teilraum von $\mathcal{CL}(\mathcal{X}, \mathcal{Y})$.*
2. *Aus $A \in \mathcal{CL}(W, \mathcal{X})$, $K \in \mathcal{KL}(\mathcal{X}, \mathcal{Y})$ und $B \in \mathcal{CL}(\mathcal{Y}, \mathcal{Z})$ folgt $K \circ A \in \mathcal{KL}(W, \mathcal{Y})$ und $B \circ K \in \mathcal{KL}(\mathcal{X}, \mathcal{Z})$.*
3. *Aus $A \in \mathcal{CL}(\mathcal{X}, \mathcal{Y})$ mit $\dim \mathcal{R}(A) < \infty$ folgt $A \in \mathcal{KL}(\mathcal{X}, \mathcal{Y})$.*
4. *$id_\mathcal{X} \in \mathcal{KL}(\mathcal{X}, \mathcal{X}) \iff \dim \mathcal{X} < \infty$.*
5. *Es sei $K \in \mathcal{KL}(\mathcal{X}, \mathcal{X})$. Dann ist $\mathcal{R}(K)$ separabel.*

Beweis:

1. Die Linearität ist klar, $KS_\mathcal{X}$ ist beschränkt, mithin existiert für $K \in \mathcal{KL}(\mathcal{X}, \mathcal{Y})$

$$\|K\| = \sup_{s \in S_\mathcal{X}} \|Ks\|.$$

2. Es sei zum Beispiel $S_\mathcal{W}$ die Sphäre in $\mathcal{W}$. Dann ist $AS_\mathcal{W}$ beschränkt und damit $KAS_\mathcal{W}$ in $\mathcal{Y}$ relativ kompakt.

3. $AS_\mathcal{X}$ ist in $\mathcal{Y}$ beschränkt. Wegen $\dim \mathcal{R}(A) < \infty$ enthält (As_n) aufgrund des Satzes von Bolzano-Weierstraß eine konvergente Teilfolge.

4. Das ist Satz 1.3.20.

5. Es seien für $n \in \mathbb{N}$

$$B_n := \left\{ x \in \mathcal{X} \;\middle|\; \|x\| \le n \right\}.$$

Dann ist

$$\mathcal{R}(K) = \bigcup_{n=1}^{\infty} KB_n,$$

und wir haben nur noch die Separabilität der KB_i zu zeigen. Die folgt aber aus der Kompaktheit. Zu jedem $\varepsilon > 0$ gibt es ein $n(\varepsilon)$ und y_j mit

$$KB_i \subset \bigcup_{j=1}^{n(\varepsilon)} B(y_j, \varepsilon).$$

Mithin existiert eine abzählbare überall dichte Teilmenge. □

Beispiel 7.2.3: *Es seien* $J = [a,b]$, $\mathcal{X} = \mathcal{Y} = \mathcal{C}(J,\mathbb{R})$, $k \in \mathcal{C}(J \times J, \mathbb{R})$ *und*

$$\begin{aligned} K \;:\; \mathcal{X} &\longrightarrow \mathcal{X}, \\ f &\longmapsto Kf := \int_a^b f(t)\, k(\,\cdot\,,t)\, dt. \end{aligned}$$

Dann ist K *kompakt.*

Beweis: Es sei

$$M := \sup_{s,t} |k(s,t)|.$$

Dann folgt aus $M < \infty$

$$\|Kf\| \le (b-a)\, M\, \|f\|.$$

Deshalb ist $K \in \mathcal{CL}(\mathcal{X})$. Es seien $f_n \in S_{\mathcal{X}}$ und $x_n := Kf_n$. Dann ist

$$\|x_n\| \le \|K\|$$

und

$$\big|x_n(s_1) - x_n(s)\big| \le \int_a^b \big|k(s_1,t) - k(s,t)\big|\, dt.$$

Nun ist k auf $J \times J$ gleichmäßig stetig, es gilt daher

$$\forall\, \varepsilon > 0 \quad \exists\, \delta > 0 \quad \forall\, s_1, s, t \in J,\ |s_1 - s| < \delta, \qquad \big|k(s_1,t) - k(s,t)\big| < \varepsilon$$

oder

$$\big|x_n(s_1) - x_n(s)\big| \le (b-a)\varepsilon.$$

Mithin ist die Folge (x_n) beschränkt und gleichartig stetig. Aus dem Satz von Arzelà-Ascoli folgt daher die Kompaktheit von K. □

Es sei noch bemerkt, daß sich die Voraussetzung $k \in \mathcal{C}(J \times J)$ abschwächen läßt. Beispielsweise kann man eine singuläre Abbildung wie

$$k(s,t) := \frac{1}{\sqrt{|s-t|}}$$

zulassen.

Beispiel 7.2.4: *Es seien* $I = (a,b)$, $\mathcal{X} = \mathcal{Y} = \mathcal{L}^2(I,\mathbb{R})$, $k \in \mathcal{L}^2(I \times I, \mathbb{R})$ *und*

$$\begin{aligned} K \;:\; \mathcal{X} &\longrightarrow \mathcal{X}, \\ f &\longmapsto Kf := \int_a^b f(t)\, k(\,\cdot\,,t)\, dt. \end{aligned}$$

Dann ist K *kompakt.*

Beweis:

1. Fast überall ist

$$|(Kf)(s)|^2 \le \int_a^b |f(t)|^2\, dt \cdot \int_a^b |k(s,t)|^2\, dt,$$

also

$$\|Kf\|^2 \le \|f\|^2 \cdot \int_{I\times I} |k(s,t)|^2\, ds\, dt$$

oder

$$\|K\| \le \sqrt{\int_{I\times I} |k(s,t)|^2\, ds\, dt} = \|k\|_{\mathcal{L}^2(I\times I)}.$$

Das heißt, es ist $K \in \mathcal{CL}(\mathcal{X})$.

2. Wir approximieren nun K durch $K_n \in \mathcal{KL}(\mathcal{X})$ in der Operatornorm. Aus dem nächsten Satz folgt dann $K \in \mathcal{KL}(\mathcal{X})$.

Es sei (u_n) ein vollständiges Orthonormalsystem in $\mathcal{X}$. Dann gilt

$$\sum_{n=1}^{\infty} \|Ku_n\|^2 = \|k\|^2. \qquad (*)$$

Es ist nämlich

$$(Ku_n)(s) = \int_a^b k(s,t)\,u_n(t)\,dt = \big(k(s,\cdot\,),\bar{u}_n\big).$$

Mit (u_n) ist auch $(\bar{u}_n)$ ein vollständiges Orthonormalsystem, aus der Parsevalschen Gleichung folgt deshalb fast überall

$$\sum_{n=1}^{\infty} \big|(Ku_n)(\,\cdot\,)\big|^2 = \int_a^b k^2(\,\cdot\,,t)\,dt =: f \in \mathcal{L}^1(I).$$

Es sei

$$f_n(s) := \sum_{i=1}^{n} \big|(Ku_i)(s)\big|^2.$$

Dann ist $f_n \in \mathcal{L}^1(I)$, und fast überall gilt $f_n \geq 0$, $f_n \nearrow f$. Aus dem Satz von B. Levi über die monotone Konvergenz folgt daher

$$\lim_{n\to\infty} \int_a^b f_n = \int_a^b f,$$

also Gl. $(*)$.

Für $f \in \mathcal{X}$ ist $f = \sum_i (f,u_i)u_i$. Weil K stetig ist, folgt daraus

$$Kf = \sum_{i=1}^{\infty} (f,u_i)\,Ku_i.$$

Es sei

$$K_n f := \sum_{i=1}^{n} (f,u_i)\,Ku_i.$$

Offenbar ist $K_n \in \mathcal{CL}(\mathcal{X})$ mit $\dim \mathcal{R}(K_n) \leq n$. Mithin ist K_n kompakt, und es folgt

$$\big\|(K_n - K)f\big\|^2 \leq \Big(\sum_{i=n+1}^{\infty} \big|(f,u_i)\big|\,\big\|Ku_i\big\|\Big)^2 \leq \Big(\sum_{i=n+1}^{\infty} \big|(f,u_i)\big|^2\Big)\cdot\Big(\sum_{i=n+1}^{\infty} \big\|Ku_i\big\|^2\Big) \leq$$

$$\leq \|f\|^2 \cdot \sum_{i=n+1}^{\infty} \|Ku_i\|^2.$$

Dies gilt für alle $f \in \mathcal{X}$, deshalb folgt aus $\sum_{i=1}^{\infty} \|Ku_i\|^2 = \|k\|^2$

$$\big\|K_n - K\big\| \leq \sqrt{\sum_{i=n+1}^{\infty} \big\|Ku_i\big\|^2} \to 0.$$

Das war zu zeigen. □

Satz 7.2.5: *Es seien $\mathcal{X}$ ein normierter Raum und $\mathcal{Y}$ ein Banachraum. Dann ist $\mathcal{KL}(\mathcal{X},\mathcal{Y})$ in $\mathcal{CL}(\mathcal{X},\mathcal{Y})$ abgeschlossen.*

Beweis: Es seien A aus dem Abschluß von $\mathcal{KL}(\mathcal{X},\mathcal{Y})$ in $\mathcal{CL}(\mathcal{X},\mathcal{Y})$ und $B := \big\{x \in \mathcal{X} \;\big|\; \|x\| \leq 1\big\}$. Nach Folgerung 1.2.34 genügt es dann zu zeigen, daß AB in $\mathcal{Y}$ präkompakt ist. Zu $\varepsilon > 0$ wählen wir ein $K \in \mathcal{KL}(\mathcal{X},\mathcal{Y})$ mit

$$\big\|A - K\big\| < \frac{\varepsilon}{2}$$

und ein $n = n(\varepsilon)$ sowie $y_1,\ldots,y_n$ mit

$$KB \subset \bigcup_{i=1}^{n} B\big(y_i, \tfrac{\varepsilon}{2}\big).$$

Dann ist

$$AB \subset \bigcup_{i=1}^{n} B(y_i,\varepsilon).$$

Es seien nämlich $x \in B$ und $y_i \in B(Kx,\varepsilon/2)$. Dann ist

$$\big\|Ax - y_i\big\| \leq \big\|(A-K)x\big\| + \big\|Kx - y_i\big\| < \varepsilon.$$ □

Wir zeigen auch

Satz 7.2.6: *Es seien $\mathcal{X}, \mathcal{Y}$ normierte Räume und $K \in \mathcal{KL}(\mathcal{X}, \mathcal{Y})$. Dann ist*

$$K' \in \mathcal{KL}(\mathcal{Y}', \mathcal{X}').$$

Beweis: Es sei (y'_n) eine Folge aus $B_{\mathcal{Y}'} := \{y' \in \mathcal{Y}' \mid \|y'\| \leq 1\}$. Wir wollen zeigen, daß $(K'y'_n)$ eine konvergente Teilfolge enthält. Dazu sei mit $D := \overline{KB_{\mathcal{X}}}$

$$f_n := y'_n | D.$$

Dann ist D kompakt und $f_n \in \mathcal{CL}(D, \mathbb{C})$. Für alle $n \in \mathbb{N}$ und $y, y_0 \in D$ gilt

$$\big|f_n(y) - f_n(y_0)\big| \leq \|f_n\| \cdot \|y - y_0\| \leq \|y - y_0\|$$

und

$$|f_n(y)| \leq \|y\|.$$

Mithin ist (f_n) beschränkt und gleichartig stetig. Es gibt deshalb eine gleichmäßig konvergente Teilfolge, die wir wieder mit (f_n) bezeichnen. Dann folgt

$$\big\|K'y'_m - K'y'_n\big\| = \sup_{x \in B_{\mathcal{X}}} \big|(K'y'_m - K'y'_n)x\big| = \sup_{x \in B_{\mathcal{X}}} \big|(y'_m - y'_n)Kx\big| \leq \sup_{y \in D} \big|f_m(y) - f_n(y)\big| \to 0.$$

Damit ist $(K'y'_n)$ in $\mathcal{X}'$ eine Cauchyfolge. $\mathcal{X}'$ ist vollständig, es gibt also ein $x' \in \mathcal{X}'$ mit

$$K'y'_n \to x'.$$

Das war zu zeigen. □

Es gilt auch die andere Richtung dieser Aussage, nämlich

Satz 7.2.7: *Es seien $\mathcal{X}, \mathcal{Y}$ normierte Räume, $\mathcal{Y}$ vollständig, $A \in \mathcal{CL}(\mathcal{X}, \mathcal{Y})$ und $A' \in \mathcal{KL}(\mathcal{Y}', \mathcal{X}')$. Dann ist $A \in \mathcal{KL}(\mathcal{X}, \mathcal{Y})$.*

Beweisskizze: Es ist $A'' \in \mathcal{KL}(\mathcal{X}'', \mathcal{Y}'')$ und nach Lemma 1.5.6

$$J_{\mathcal{Y}} \circ A \;=\; A'' \circ J_{\mathcal{X}}.$$

Daraus folgt

$$AS_{\mathcal{X}} = J_{\mathcal{Y}}^{-1} A'' J_{\mathcal{X}} S_{\mathcal{X}} \subset J_{\mathcal{Y}}^{-1} A'' S_{\mathcal{X}''}.$$

Weil $\mathcal{Y}$ vollständig ist, ist $\overline{AS_{\mathcal{X}}}$ kompakt. □

Wir wollen nun schließlich auch den Begriff der schwachen Konvergenz zur Charakterisierung kompakter Abbildungen heranziehen und zeigen

Satz 7.2.8: *Es seien $\mathcal{X}, \mathcal{Y}$ normierte Räume, $K \in \mathcal{KL}(\mathcal{X}, \mathcal{Y})$, $x \in \mathcal{X}$ und (x_n) eine Folge in $\mathcal{X}$. Dann gilt*

$$\operatorname*{w\text{-}lim}_{n \to \infty} x_n = x \;\Longrightarrow\; \Big(Kx_n \to Kx\Big).$$

Beweis: Es sei $\text{w-lim}_{n\to\infty} x_n = o$, und es gelte

$$\neg \Big(Kx_n \to o\Big).$$

Dann gibt es eine Teilfolge $(x_{1,n})$ und ein $c > 0$ mit

$$\forall\, n \in \mathbb{N} \quad \|Kx_{1,n}\| \geq c.$$

Die schwach konvergente Folge $(x_{1,n})$ ist beschränkt. Mithin gibt es eine Teilfolge $(x_{2,n})$ und ein $y \in \mathcal{Y}$ mit

$$Kx_{2,n} \to y.$$

Dabei ist $\|y\| \geq c > 0$, und es folgt für alle $y' \in \mathcal{Y}'$

$$y'y = y'Kx_{2,n} - y'\big(Kx_{2,n} - y\big) = \underbrace{(K'y')}_{\in \mathcal{X}'} x_{2,n} - y'\big(Kx_{2,n} - y\big) \to 0.$$

Das ist ein Widerspruch. □

Es folgt

Satz 7.2.9: *Es seien $\mathcal{X}, \mathcal{Y}$ normierte Räume, $\mathcal{X}$ reflexiv und $K \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$. Dann sind äquivalent:*

(1) $K \in \mathcal{KL}(\mathcal{X}, \mathcal{Y})$.

(2) Für jede Folge (x_n) in $\mathcal{X}$ und $x \in \mathcal{X}$ gilt

$$\text{w-}\lim_{n\to\infty} x_n = x \implies \Big(Kx_n \to Kx \Big).$$

Beweis:

(1) $\Rightarrow$ (2) wurde soeben bewiesen.

(2) $\Rightarrow$ (1): Nach Satz 4.1.4 ist $\bar{B}(o,1) \subset \mathcal{X}$ schwach folgenkompakt. Es sei (x_n) eine Folge aus $S_{\mathcal{X}}$. Dann gibt es eine Teilfolge, die wir wieder (x_n) nennen, und ein $x \in \mathcal{X}$ mit

$$\text{w-}\lim_{n\to\infty} x_n = x.$$

Daraus folgt aber $Kx_n \to Kx$ und damit $K \in \mathcal{KL}(\mathcal{X}, \mathcal{Y})$. □

Wenn $\mathcal{X}$ reflexiv ist, gilt also

$$\text{w-}\lim_{n\to\infty} x_n = x \implies \Big(K \in \mathcal{KL}(\mathcal{X}, \mathcal{Y}) \iff Kx_n \to Kx \Big).$$

Ist $\mathcal{X}$ nicht reflexiv, dann ist die Implikation (2) $\Rightarrow$ (1) im letzten Satz jedoch im allgemeinen falsch. Zum Beispiel seien $\mathcal{X} = \mathcal{Y} = \ell^1$. Dann folgt nach Satz 4.2.4

$$\text{w-}\lim_{n\to\infty} x_n = x \implies x_n \to x.$$

Die Abbildung $K = id$ ist aber nicht kompakt, denn für $x_n = e_n$ enthält (Ke_n) wegen

$$\|e_n - e_m\| = 2\,(1 - \delta_{nm})$$

keine konvergente Teilfolge.

7.3 Die Riesz-Schauder Theorie

In diesem Abschnitt wollen wir uns wieder dem Lösen von Gleichungen

$$Ax \stackrel{!}{=} y \qquad (*)$$

zuwenden. Dabei seien $\mathcal{X}$ ein Banachraum, $x, y \in \mathcal{X}$ und A eine kompakte Störung der Identität, also mit $K \in \mathcal{KL}(\mathcal{X})$

$$A = id + K.$$

Wir wollen zeigen, daß für solche Gleichungen wie in $\mathbb{R}^n$ die Fredholmsche Alternative gilt. Dieses Resultat geht auf Frigyes Riesz und Juliusz Pawel Schauder, 1899–1943, zurück. Es stammt aus den Jahren 1918–1930 und verallgemeinert ein Ergebnis, das Fredholm um 1900 bewiesen hatte.

Wir beginnen mit

Satz 7.3.1: *Es ist $\mathcal{N}(A)$ ein abgeschlossener endlichdimensionaler Teilraum von $\mathcal{X}$.*

Beweis: Die Abgeschlossenheit von $\mathcal{N}(A)$ ist klar. Wir zeigen die Kompaktheit der Sphäre $S_{\mathcal{N}(A)}$ in $\mathcal{N}(A)$. Aus dem Rieszschen Lemma folgt dann die Behauptung. Es sei also (x_n) eine Folge aus $S_{\mathcal{N}(A)}$. Dann gibt es eine Teilfolge, die wir wieder (x_n) nennen, und ein $x \in \mathcal{X}$ mit

$$Kx_n \to -x.$$

Wegen $Kx_n = -x_n$ gilt dann auch $x_n \to x \in S_{\mathcal{N}(A)}$. □

Satz 7.3.2: *Es seien $x \in \mathcal{X}$ und $d(x) := \operatorname{dist}(x, \mathcal{N}(A))$. Dann gilt*

$$\exists\, M > 0 \quad \forall\, x \in \mathcal{X} \qquad d(x) \le M\,\|Ax\|.$$

Den Beweis führen wir indirekt. Es sei (x_n) eine Folge aus $\mathcal{X}$ mit

$$\forall\, n \in \mathbb{N} \qquad d(x_n) > n\, \|Ax_n\|.$$

Wegen $\dim \mathcal{N}(A) < \infty$ existieren $y_n \in \mathcal{N}(A)$ mit

$$d(x_n) = \|x_n - y_n\|.$$

Es seien

$$z_n := \frac{x_n - y_n}{d(x_n)}.$$

Dann ist $\|z_n\| = 1$, $d(z_n) = 1$ und

$$\|Az_n\| = \frac{1}{d(x_n)}\, \|Ax_n\| < \frac{1}{n}.$$

Nun ist K kompakt. Folglich gibt es eine Teilfolge, die wir wieder (z_n) nennen, und ein $z \in \mathcal{X}$ mit

$$Kz_n \to -z$$

und

$$z_n = Az_n - Kz_n \to z.$$

Dann ist $z \in \mathcal{N}(A)$, also $d(z) = 0$. Andererseits folgt aber

$$d(z) \geq \big|d(z_n) - |z - z_n|\big| \to 1. \qquad \square$$

Folgerung 7.3.3: *Es gilt*

$$\forall\, y \in \mathcal{R}(A) \quad \exists\, x \in \mathcal{X} \qquad Ax = y \quad \textit{und} \quad \|x\| \leq M\, \|y\|.$$

Beweis: Es seien $y \in \mathcal{R}(A)$ und $z \in \mathcal{X}$ mit $Az = y$. Wir wählen $n \in \mathcal{N}(A)$ mit $\|z - n\| = d(z)$ und

$$x := z - n.$$

Dann ist $Ax = Az = y$ und

$$\|x\| = \|z - n\| = d(z) \leq M\, \|Az\| = M\, \|y\|. \qquad \square$$

Satz 7.3.4: $\mathcal{R}(A)$ *ist abgeschlossen.*

Beweis: Es seien $y_n \in \mathcal{R}(A)$ mit $y_n \to y \in \mathcal{X}$. Wir wählen $x_n \in \mathcal{X}$ mit $Ax_n = y_n$ und $\|x_n\| \leq M\, \|y_n\|$. Weil K kompakt ist, gibt es eine Teilfolge, die wir wieder (x_n) nennen, mit $Kx_n \to x$. Dann folgt

$$x_n = Ax_n - Kx_n = y_n - Kx_n \to y - x$$

und

$$y_n = Ax_n \to A(y - x),$$

also

$$y = A(y - x) \in \mathcal{R}(A). \qquad \square$$

Mit A betrachten wir auch die Potenzen A^n von A. Es ist für $n \in \mathbb{N}_0$

$$A^n = id + K_n \quad \text{mit } K_0 := O \text{ und } K_n := \sum_{i=1}^{n} \binom{n}{i} K^i \text{ für } n \in \mathbb{N}.$$

Dabei ist K_n ebenfalls kompakt, folglich sind $\mathcal{N}(A^n)$ und $\mathcal{R}(A^n)$ abgeschlossen. Es seien

$$\mathcal{N}_n := \mathcal{N}(A^n) \quad \text{und} \quad \mathcal{R}_n := \mathcal{R}(A^n).$$

Offenbar ist

$$\{o\} = \mathcal{N}_0 \subset \mathcal{N}_1 \subset \mathcal{N}_2 \subset \cdots$$
$$\mathcal{X} = \mathcal{R}_0 \supset \mathcal{R}_1 \supset \mathcal{R}_2 \supset \cdots.$$

Wir zeigen nun vier Hilfssätze:

Lemma 7.3.5: *Es gibt eine Zahl $r \in \mathbb{N}_0$ mit*

$$\mathcal{N}_0 \subsetneq \cdots \subsetneq \mathcal{N}_r = \mathcal{N}_{r+1} = \cdots .$$

Beweis: Es sei

$$\mathcal{M} := \left\{ n \in \mathbb{N}_0 \;\middle|\; \mathcal{N}_n = \mathcal{N}_{n+1} \right\} .$$

$\mathcal{M}$ ist nicht leer. Es sei nämlich $\mathcal{M} = \emptyset$. Dann sind für alle $n \in \mathbb{N}$

$$\mathcal{N}_{n-1} \subsetneq \mathcal{N}_n$$

abgeschlossene Unterräume, und aus dem Rieszschen Lemma folgt die Existenz eines $x_n \in \mathcal{N}_n$ mit $\|x_n\| = 1$ und

$$\operatorname{dist}(x_n, \mathcal{N}_{n-1}) \geq \frac{1}{2}.$$

Für $n > m \geq 1$ gilt also

$$Kx_m - Kx_n = x_n - y \qquad \text{mit} \quad y := x_m + Ax_n - Ax_m \in \mathcal{N}_{n-1}.$$

Daraus folgt

$$\left\| Kx_m - Kx_n \right\| \geq \frac{1}{2}$$

im Widerspruch zur Kompaktheit von K. Die Menge $\mathcal{M}$ ist also nicht leer. Es sei $n \in \mathcal{M}$. Dann ist auch $n+1 \in \mathcal{M}$. Das folgt aus

$$A^{-1}\mathcal{N}_n = \mathcal{N}_{n+1},$$

also

$$\mathcal{N}_{n+1} = A^{-1}\mathcal{N}_n = A^{-1}\mathcal{N}_{n+1} = \mathcal{N}_{n+2}.$$

Mithin leistet

$$r := \min \mathcal{M}$$

das Gewünschte. □

Die in Lemma 7.3.5 gefundene Zahl r nennen wir *Rieszzahl* $r(A)$ von A.

Lemma 7.3.6: *Es gibt eine Zahl $q \in \mathbb{N}_0$ mit*

$$\mathcal{R}_0 \supsetneq \cdots \supsetneq \mathcal{R}_q = \mathcal{R}_{q+1} = \cdots .$$

Der Beweis ähnelt dem zu Lemma 7.3.5. Es sei jetzt

$$\mathcal{M} := \left\{ n \in \mathbb{N}_0 \;\middle|\; \mathcal{R}_n = \mathcal{R}_{n+1} \right\} .$$

Dann ist $\mathcal{M}$ nicht leer. Es sei nämlich $\mathcal{M}$ leer. Dann ist für alle $n \in \mathbb{N}_0$

$$\mathcal{R}_{n+1} \subsetneq \mathcal{R}_n,$$

und aus dem Rieszschen Lemma folgt die Existenz eines $x_n \in \mathcal{R}_n$ mit $\|x_n\| = 1$ und

$$\operatorname{dist}(x_n, \mathcal{R}_{n+1}) \geq \frac{1}{2}.$$

Für $m > n \geq 0$ gilt jetzt

$$Kx_m - Kx_n = x_n - y \qquad \text{mit} \quad y := x_m + Ax_n - Ax_m \in \mathcal{R}_{n+1},$$

also

$$\left\| Kx_m - Kx_n \right\| \geq \frac{1}{2}$$

im Widerspruch zur Kompaktheit von K. Die Menge $\mathcal{M}$ ist also nicht leer. Es sei $n \in \mathcal{M}$. Dann ist auch $n+1 \in \mathcal{M}$. Das folgt aus

$$\mathcal{R}_{n+1} = A\mathcal{R}_n = A\mathcal{R}_{n+1} = \mathcal{R}_{n+2}.$$

Mithin leistet

$$q := \min \mathcal{M}$$

das Gewünschte. □

Lemma 7.3.7: *Es ist* $q = r(A)$.

Beweis:

1. Es sei $q < r$. Dann wählen wir $x \in \mathcal{N}_r$ mit $x \notin \mathcal{N}_{r-1}$. Wegen

$$A^{r-1}x \in \mathcal{R}_{r-1} = \mathcal{R}_r$$

gibt es ein $y \in \mathcal{X}$ mit

$$A^{r-1}x = A^r y,$$

und wir erhalten

$$o = A^r x = A^{r+1} y,$$

also

$$y \in \mathcal{N}_{r+1} = \mathcal{N}_r$$

oder

$$A^{r-1}x = A^r y = o.$$

Das ist ein Widerspruch.

2. Es sei $q > r$. Dann wählen wir $x \in \mathcal{R}_{q-1}$ mit $x \notin \mathcal{R}_q$, und es gibt ein $y \in \mathcal{X}$ mit

$$x = A^{q-1}y.$$

Wegen

$$A^q y \in \mathcal{R}_q = \mathcal{R}_{q+1}$$

gibt es ein $z \in \mathcal{X}$ mit

$$A^q y = A^{q+1} z$$

oder

$$A^q(y - Az) = o.$$

Das bedeutet

$$y - Az \in \mathcal{N}_q = \mathcal{N}_{q-1}$$

oder

$$x - A^q z = A^{q-1}(y - Az) = o.$$

Wir erhalten also den Widerspruch

$$x = A^q z \in \mathcal{R}_q.$$ □

Lemma 7.3.8: *Es gilt*

$$\mathcal{X} = \mathcal{N}_r \oplus \mathcal{R}_r.$$

Beweis: Es sei $x \in \mathcal{N}_r \cap \mathcal{R}_r$. Dann gibt es ein $y \in \mathcal{X}$ mit $x = A^r y$, und es folgt

$$A^{2r} y = o.$$

Damit ist $y \in \mathcal{N}_{2r} = \mathcal{N}_r$, also $x = o$. Das heißt, es ist

$$\mathcal{N}_r \cap \mathcal{R}_r = \{o\}.$$

Es sei nun $x \in \mathcal{X}$. Dann gibt es wegen $A^r x \in \mathcal{R}_r = \mathcal{R}_{2r}$ ein $y \in \mathcal{X}$ mit $A^r x = A^{2r} y$. Es folgt also

$$A^r(x - A^r y) = o$$

oder

$$x - A^r y \in \mathcal{N}_r.$$

Das liefert die gewünschte Zerlegung

$$x = \underbrace{A^r y}_{\in \mathcal{R}_r} + \underbrace{(x - A^r y)}_{\in \mathcal{N}_r}.$$ □

Wir fassen unsere bisherigen Ergebnisse zusammen im

Satz 7.3.9: *Für alle $n \in \mathbb{N}$ sind die Räume $\mathcal{N}(A^n)$ und $\mathcal{R}(A^n)$ in $\mathcal{X}$ abgeschlossen. Es ist* $\dim \mathcal{N}(A^n) < \infty$, *und es gibt eine Zahl $r \in \mathbb{N}_0$, $r = r(A)$, die Rieszzahl von A, mit*

$$\{o\} = \mathcal{N}(A^0) \subsetneq \cdots \subsetneq \mathcal{N}(A^r) = \mathcal{N}(A^{r+1}) = \cdots$$
$$\mathcal{X} = \mathcal{R}(A^0) \supsetneq \cdots \supsetneq \mathcal{R}(A^r) = \mathcal{R}(A^{r+1}) = \cdots .$$

Ferner gilt

$$\mathcal{X} = \mathcal{N}(A^r) \oplus \mathcal{R}(A^r).$$

Folgerung 7.3.10: *Die Gleichung*

$$Ax \stackrel{!}{=} y \qquad (*)$$

ist in $\mathcal{R}_r$ eindeutig lösbar.

Beweis: Wegen $y \in \mathcal{R}_r = \mathcal{R}_{r+1}$ gibt es ein $x \in \mathcal{R}_r$ mit $y = Ax$. Dieses $x \in \mathcal{R}_r$ ist eindeutig bestimmt. Es sei nämlich $x = A^r z$. Dann folgt aus $Ax = o$

$$z \in \mathcal{N}_{r+1} = \mathcal{N}_r,$$

also $x = o$. □

Ein wichtiger Spezialfall ist natürlich der Fall $r = 0$. Jetzt ist Gl. $(*)$ für alle $y \in \mathcal{X}$ eindeutig lösbar, und für die Lösung x gilt die Abschätzung

$$\|x\| \leq M \, \|y\|.$$

Auch im Fall $r = 1$ erhält man aus

$$\mathcal{X} = \mathcal{N}(A) \oplus \mathcal{R}(A)$$

eine Charakterisierung von $\mathcal{R}(A)$. Schwieriger wird es für größere r. Eine befriedigende Lösungstheorie erhält man dann erst, wenn zusammen mit A auch der zu A konjugierte Operator diskutiert wird. Bevor wir eine solche Theorie entwickeln, soll im folgenden Beispiel zur Illustration an den Fall $\mathcal{X} = \mathbb{R}^2$ erinnert werden.

Beispiel 7.3.11: *Der Fall $\mathcal{X} = \mathbb{R}^2$.*

Wir greifen drei typische Abbildungen heraus:

1. Es sei $A = id$. Dann ist $r(A) = 0$ und Gl. $(*)$ eindeutig lösbar.

2. Es sei

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Nun gilt

$$id = A^0 \neq A = A^2 = \cdots .$$

Damit ist $r(A) = 1$,

$$\mathcal{N}(A) = \operatorname{span}(e_2) \quad \text{und} \quad \mathcal{R}(A) = \operatorname{span}(e_1)$$

sowie

$$\mathcal{X} = \mathcal{N}(A) \oplus \mathcal{R}(A).$$

In $\mathcal{R}(A)$ ist Gl. $(*)$ eindeutig lösbar.

3. Es sei

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Nun gilt

$$id = A^0 \neq A \neq A^2 = \cdots = O.$$

Damit ist $r(A) = 2$ und $\mathcal{R}(A^2) = \{o\}$. Die Aussage

$$\mathcal{X} = \mathcal{N}(A^2) \oplus \mathcal{R}(A^2)$$

hilft jetzt also nicht weiter.

In diesem Fall leistet die bisher entwickelte Theorie noch nicht genug. Aus der Matrizenrechnung weiß man, daß nun neben den Eigenvektoren auch Hauptvektoren ins Spiel kommen. In den beiden ersten Fällen war A diagonalähnlich, während man jetzt für A die Jordansche Normalform erhält.

Insbesondere zeigt sich, daß man zur besseren Beschreibung des Sachverhalts auch die transponierte Abbildung A^{tr} mitbetrachten sollte.

Um das Beispiel abzuschließen, bemerken wir nur noch, daß jetzt $e := e_1$ Eigenvektor von A ist, also $\mathcal{N}(A) = \text{span}(e)$, und $\mathcal{R}(A) = \mathcal{N}(A)$. Damit ist

$$\mathcal{X} \neq \mathcal{N}(A) \oplus \mathcal{R}(A).$$

Es gilt aber

$$\mathcal{X} = \mathcal{N}(A^{tr}) \oplus \mathcal{R}(A),$$

und das liefert die notwendige und hinreichende Lösungsbedingung für y. Es ist $\mathcal{N}(A^{tr}) = \text{span}(e_2)$, und $h := e_2$ ist ein *Hauptvektor* von A; es gilt nämlich $Ah = e$ und $A^2h = o$. □

Im folgenden soll also die konjugierte Abbildung A' mit einbezogen werden. Dabei halten wir die Darstellung zunächst etwas allgemeiner als notwendig und gehen von einem Paar $\mathcal{X}, \mathcal{Y}$ normierter Räume aus und von Abbildungen

$$K \in \mathcal{KL}(\mathcal{X}) \quad \text{und} \quad L \in \mathcal{KL}(\mathcal{Y})$$

sowie

$$A := id + K \quad \text{und} \quad B := id + L.$$

Dabei denken wir an drei Fälle:

1. Es seien $\mathcal{X} = \mathcal{Y}$ ein Hilbertraum und $B = A^*$.
2. Es seien $\mathcal{X} = \mathcal{Y} = \mathcal{C}(J, \mathbb{C})$, zusätzlich ausgestattet mit dem $\mathcal{L}^2(J)$-Skalarprodukt, und $B = A^*$ bezüglich dieses Skalarproduktes. Dabei soll J der Abschluß eines beschränkten Gebietes in $\mathbb{R}^n$ sein.
3. Es seien $\mathcal{X}$ ein Banachraum, $\mathcal{Y} = \mathcal{X}'$ und $B = A'$.

Im ersten und dritten Fall genügt es, $K \in \mathcal{KL}(\mathcal{X})$ vorauszusetzen, $L \in \mathcal{KL}(\mathcal{Y})$ folgt dann. Im zweiten Fall läßt sich $L \in \mathcal{KL}(\mathcal{Y})$ aber im allgemein nicht aus $K \in \mathcal{KL}(\mathcal{X})$ ableiten.

Wir nehmen nun an, daß es eine sesquilineare Form (oder eine bilineare Form)

$$\begin{aligned} [\,\cdot\,,\cdot\,] \;:\; \mathcal{X} \times \mathcal{Y} &\longrightarrow \mathbb{C}, \\ \langle x, y\rangle &\longmapsto [x, y] \end{aligned}$$

gibt mit

$$\{x \in \mathcal{X} \mid \forall\, y \in \mathcal{Y} \quad [x, y] = 0\} = \{o\}$$
$$\{y \in \mathcal{Y} \mid \forall\, x \in \mathcal{X} \quad [x, y] = 0\} = \{o\}$$

und

$$\forall\, \langle x, y\rangle \in \mathcal{X} \times \mathcal{Y} \qquad [Ax, y] = [x, By].$$

Die soeben genannten Spezialfälle lassen sich hier unterordnen, man nehme das Skalarprodukt bzw. die Abbildung $x'x$.

Wir beginnen mit folgendem

Lemma 7.3.12: *Es seien $x_1, \ldots, x_n \in \mathcal{X}$ linear unabhängig. Dann gibt es $y_1, \ldots, y_n \in \mathcal{Y}$ mit*

$$\forall\, 1 \leq i, j \leq n \qquad [x_i, y_j] = \delta_{ij}.$$

Entsprechendes gilt, wenn man $y_1, \ldots, y_n \in \mathcal{Y}$ vorgibt.

Beweis:

1. Für $n = 1$ ist die Aussage richtig.
2. Die Aussage sei für n Elemente $x_1, \ldots, x_n$ richtig. Wir zeigen sie für $n + 1$ Elemente. Es seien also $x_1, \ldots, x_{n+1} \in \mathcal{X}$ linear unabhängig. Dann gilt

$$\forall\, l \in \{1, 2, \ldots, n+1\} \quad \forall\, x_1, \ldots, x_{l-1}, x_{l+1}, \ldots, x_{n+1} \quad \exists\, y_1^l, \ldots, y_{l-1}^l, y_{l+1}^l, \ldots, y_{n+1}^l$$

$$\forall\, i, j \neq l \qquad [x_i, y_j^l] = \delta_{ij}.$$

Nun gibt es ein $\eta_l \in \mathcal{Y}$ mit

$$\alpha_l := \left[x_l,\, \eta_l - \sum_{\substack{k=1 \\ k\neq l}}^{n+1} y_k^l\, \overline{[x_k, \eta_l]}\right] = \left[x_l - \sum_{\substack{k=1 \\ k\neq l}}^{n+1} [x_l, y_k^l]\, x_k,\, \eta_l\right] \neq 0,$$

denn sonst wäre

$$x_l = \sum_{\substack{k=1\\k\neq l}}^{n+1} [x_l, y_k^l]\, x_k.$$

Wir wählen nun

$$y_l := \frac{1}{\alpha_l}\left\{\eta_l - \sum_{\substack{k=1\\k\neq l}}^{n+1} y_k^l\, \overline{[x_k, \eta_l]}\right\}.$$

Dann ist

$$[x_l, y_l] = 1$$

und

$$\forall\, i \neq l \qquad [x_i, y_l] = \frac{1}{\alpha_l}\left\{[x_i, \eta_l] - \sum_{\substack{k=1\\k\neq l}}^{n+1} \left[x_i,\, y_k^l\, \overline{[x_k, \eta_l]}\right]\right\} = \frac{1}{\alpha_l}\left[x_i - \sum_{\substack{k=1\\k\neq l}}^{n+1} [x_i, y_k^l]\, x_k,\, \eta_l\right] = 0. \qquad \square$$

Es seien nun

$$\begin{aligned}
\mathcal{R}(A)^a &:= \{y \in \mathcal{Y} \mid \forall x \in \mathcal{X} \quad [Ax, y] = 0\}\\
\mathcal{R}(B)^a &:= \{x \in \mathcal{X} \mid \forall\, y \in \mathcal{Y} \quad [x, By] = 0\}\\
\mathcal{N}(A)^a &:= \{y \in \mathcal{Y} \mid \forall\, x \in \mathcal{N}(A) \quad [x, y] = 0\}\\
\mathcal{N}(B)^a &:= \{x \in \mathcal{X} \mid \forall\, y \in \mathcal{N}(B) \quad [x, y] = 0\}
\end{aligned}$$

Annihilatoren bezüglich der Form $[x, y]$ und

$$\operatorname{codim}\mathcal{R}(A) := \dim X/\mathcal{R}(A).$$

Dann zeigen wir unser Hauptresultat, nämlich

Satz 7.3.13: *Es ist*

$$\mathcal{R}(A)^a = \mathcal{N}(B),\ \ \mathcal{R}(B)^a = \mathcal{N}(A),\ \ \mathcal{N}(A)^a = \mathcal{R}(B),\ \ \mathcal{N}(B)^a = \mathcal{R}(A)$$

und

$$\dim\mathcal{N}(A) = \dim\mathcal{N}(B) = \operatorname{codim}\mathcal{R}(A) = \operatorname{codim}\mathcal{R}(B) \in \mathbb{N}_0.$$

Der Beweis des Satzes ist etwas langwierig. Wir zerlegen ihn in fünf Schritte:

1. Es seien $m := \dim\mathcal{N}(A)$, $n := \dim\mathcal{N}(B)$ und

$$\mathcal{N}(A) = \operatorname{span}(x_1, \ldots, x_m), \qquad \mathcal{N}(B) = \operatorname{span}(y_1, \ldots, y_n).$$

Aus dem gerade bewiesenen Lemma folgt dann die Existenz von $u_1, \ldots, u_n \in \mathcal{X}$ und $v_1, \ldots, v_m \in \mathcal{Y}$ mit

$$\begin{aligned}
\forall\, 1 \leq i, j \leq m \qquad & [x_i, v_j] = \delta_{ij}\\
\forall\, 1 \leq i, j \leq n \qquad & [u_i, y_j] = \delta_{ij}\,.
\end{aligned}$$

Ferner sieht man unmittelbar

$$\mathcal{R}(A)^a = \mathcal{N}(B),\ \ \mathcal{R}(B)^a = \mathcal{N}(A),\ \ \mathcal{N}(A)^a \supset \mathcal{R}(B),\ \ \mathcal{N}(B)^a \supset \mathcal{R}(A).$$

2. Es sei $m \leq n$. Dann definieren wir $\tilde{A} \in \mathcal{CL}(\mathcal{X})$ vermöge

$$\tilde{A}x := Ax + \sum_{i=1}^{m} [x, v_i]\, u_i.$$

Es sei $x \in \mathcal{N}(\tilde{A})$. Dann folgt für alle i, $1 \leq i \leq m$,

$$0 = \left[\tilde{A}x, y_i\right] = \left[Ax, y_i\right] + \left[x, v_i\right] = \left[x, By_i\right] + \left[x, v_i\right] = \left[x, v_i\right],$$

also $\tilde{A}x = Ax = o$ oder $x \in \mathcal{N}(A)$. Damit ist

$$x = \sum_{i=1}^{m} \xi_i\, x_i = \sum_{i=1}^{m} [x, v_i]\, x_i = 0,$$

das heißt, es ist $\mathcal{N}(\tilde{A}) = \{o\}$. Wegen

$$\tilde{A} = id + K + \sum_{i=1}^{m} [\,\cdot\,, v_i]\, u_i =: id + \tilde{K}$$

mit $\tilde{K} \in \mathcal{KL}(\mathcal{X})$ verschwindet die Rieszzahl von $\tilde{A}$. Daraus folgt $\mathcal{R}(\tilde{A}) = \mathcal{X}$.

3. Wir zeigen $m = n$. Aus Symmetriegründen genügt es dafür, den Fall $m < n$ auszuschließen. Es sei also $m < n$. Dann ist $\tilde{A}$ wohldefiniert und wegen $\mathcal{R}(\tilde{A}) = \mathcal{X}$ existiert ein $x \in \mathcal{X}$ mit

$$\tilde{A}x = u_{m+1}.$$

Es folgt der Widerspruch

$$1 = \left[\tilde{A}x, y_{m+1}\right] = \left[Ax, y_{m+1}\right] = \left[x, By_{m+1}\right] = 0.$$

4. Wir zeigen $\mathcal{N}(B)^a \subset \mathcal{R}(A)$ und $\mathcal{N}(A)^a \subset \mathcal{R}(B)$. Aus Symmetriegründen genügt wieder der Nachweis von

$$\mathcal{N}(B)^a \subset \mathcal{R}(A),$$

zusammen mit der ersten Aussage folgt dann $\mathcal{N}(B)^a = \mathcal{R}(A)$. Es sei $f \in \mathcal{N}(B)^a$ vorgegeben. Wegen $m = n$ ist $\tilde{A}$ wohldefiniert und

$$\tilde{A}x = f$$

lösbar. Dann folgt für alle $1 \le j \le m = n$

$$0 = [f, y_j] = \left[\tilde{A}x, y_j\right] = [Ax, y_j] + [x, v_j] = \left[x, By_j\right] + [x, v_j] = [x, v_j]$$

oder

$$Ax = \tilde{A}x = f,$$

also

$$\mathcal{N}(B)^a \subset \mathcal{R}(A).$$

5. Wir zeigen $\operatorname{codim} \mathcal{R}(A) = n$ und $\operatorname{codim} \mathcal{R}(B) = n$. Wegen der Symmetrie genügt wieder der Nachweis von

$$\operatorname{codim} \mathcal{R}(A) = n.$$

Weil $\mathcal{R}(A)$ abgeschlossen ist, existiert nach Lemma 1.3.6 die kanonische Abbildung

$$\varphi \;:\; \mathcal{X} \longrightarrow \mathcal{X}/\mathcal{R}(A).$$

Wir wollen zeigen, daß $\{\varphi(u_1), \ldots, \varphi(u_n)\}$ eine Basis von $\mathcal{X}/\mathcal{R}(A)$ ist. Es sei $x \in \mathcal{X}$ vorgegeben. Dann ist für alle $i = 1, \ldots, n$

$$\Big[x - \sum_{j=1}^{n} [x, y_j]\, u_j,\; y_i\Big] = [x, y_i] - [x, y_i] = 0,$$

also

$$x - \sum_{j=1}^{n} [x, y_j]\, u_j \in \mathcal{N}(B)^a = \mathcal{R}(A)$$

oder

$$\varphi(x) = \sum_{j=1}^{n} [x, y_j]\, \varphi(u_j)$$

und

$$\operatorname{span}\big(\varphi(u_1), \ldots, \varphi(u_n)\big) = \mathcal{X}/\mathcal{R}(A).$$

Wir haben nur noch die lineare Unabhängigkeit der $\varphi(u_1), \ldots, \varphi(u_n)$ zu zeigen. Es sei mit $\alpha_i \in \mathbb{C}$

$$\sum_{j=1}^{n} \alpha_j\, \varphi(u_j) = o.$$

Dann ist

$$\sum_{j=1}^{n} \alpha_j u_j \in \mathcal{R}(A) = \mathcal{N}(B)^a,$$

und es folgt

$$\forall\, i = 1, \ldots, n \qquad 0 = \left[\sum_{j=1}^{n} \alpha_j u_j, y_i \right] = \alpha_i.$$

Damit ist Satz 7.3.13 bewiesen. □

Es seien nun $\mathcal{X} = \mathcal{Y}$ ein Hilbertraum, $B = A^*$ und r die Rieszzahl von A. Dann lautet unser Ergebnis:

a) Entweder ist $r = 0$. Dann sind die Gleichungen

$$Ax \stackrel{!}{=} y \quad \text{und} \quad Bu \stackrel{!}{=} v \tag{$**$}$$

für alle $y, v \in \mathcal{X}$ eindeutig lösbar.

b) Oder es ist $r > 0$. Dann ist $0 < \dim\mathcal{N}(A) = \dim\mathcal{N}(B) < \infty$. Notwendig und hinreichend für die Lösbarkeit der Gleichungen $(**)$ sind die Bedingungen

$$y \in \mathcal{N}(B)^\perp \quad \text{bzw.} \quad v \in \mathcal{N}(A)^\perp.$$

Das ist die *Fredholmsche Alternative*. Die letzte Aussage erhält man aus

$$\mathcal{X} = \mathcal{R}(A) \oplus \mathcal{N}(A^*) \quad \text{orthogonal.}$$

Sind $\mathcal{X}, \mathcal{X}'$ Banachräume und $B = A'$, dann lauten die Lösungsbedingungen entsprechend

$$y \in \mathcal{R}(A) = \mathcal{N}(A')^a \quad \text{bzw.} \quad v \in \mathcal{R}(A') = \mathcal{N}(A)^a.$$

Wir beschließen diesen Abschnitt mit dem

Beispiel 7.3.14: *Es seien* $I = (0, 2\pi)$, $\mathcal{X} = \mathcal{L}^2(I, \mathbb{C}^2)$,

$$\mathcal{D}(D) = \left\{ u \in \mathcal{H}_1(I, \mathbb{C}^2) \;\middle|\; u(0) = u(2\pi) \right\}$$

und

$$D \;:\; u \mapsto \mathrm{i}\left\{u' + Au\right\} \quad \textit{mit } A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Offenbar ist D eine symmetrische Abbildung. Wir zeigen $\mathrm{i} \in \rho(D)$. Völlig analog folgt auch $-\mathrm{i} \in \rho(D)$ und wegen

$$\mathcal{D}(A^*) = \mathcal{D}(A) \oplus \mathcal{R}(A + \mathrm{i}\, id)^\perp \oplus \mathcal{R}(A - \mathrm{i}\, id)^\perp \tag{$*$}$$

ist D dann selbstadjungiert. Gl. $(*)$ werden wir in Satz 8.1.2 für abgeschlossene symmetrische Abbildungen A beweisen. Für unsere Abbildung D läßt sich die Selbstadjungiertheit aber auch leicht direkt verifizieren.

Wir zeigen also $\mathrm{i} \in \rho(D)$. Dazu lösen wir zunächst die homogene Gleichung

$$u' \stackrel{!}{=} (id - A)u \quad \text{mit } u(0) = u_0 \in \mathbb{C}^2$$

und finden

$$u(t) = \mathrm{e}^{(id-A)t} u_0 = \mathrm{e}^t \mathrm{e}^{-At} u_0.$$

Mit

$$\mathrm{e}^{At} := \sum_{n=0}^{\infty} \frac{1}{n!} A^n t^n$$

ist dann

$$\mathrm{e}^{At} = id\, \cos t + A\, \sin t = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix},$$

letzteres wegen

$$A^2 = -id, \;\; A^3 = -A \;\; \text{und } A^4 = id.$$

Die Lösung $u \in \mathcal{D}(D)$ von

$$(D - \mathrm{i}\, id)u \stackrel{!}{=} f \in \mathcal{X}$$

erhält man dann durch

$$u(t) = \mathrm{e}^{t}\mathrm{e}^{-At}u_0 - \mathrm{i}\,\mathrm{e}^{t}e^{-At}\int_0^t \mathrm{e}^{-s}\mathrm{e}^{As}f(s)\,ds.$$

Dabei ist $u_0 \in \mathbb{C}^2$ so zu wählen, daß $u(0) = u(2\pi)$ ist, also

$$u_0 \stackrel{!}{=} \mathrm{e}^{2\pi}u_0 - \mathrm{i}\,\mathrm{e}^{2\pi}\int_0^{2\pi} \mathrm{e}^{-s}e^{As}f(s)\,ds.$$

Es sei also

$$u_0 := \frac{-\mathrm{i}\,\mathrm{e}^{2\pi}}{(1-\mathrm{e}^{2\pi})}\int_0^{2\pi} \mathrm{e}^{-s}\mathrm{e}^{As}f(s)\,ds.$$

Mithin existiert die Abbildung

$$\begin{aligned} M \;:\; \mathcal{X} &\longrightarrow \mathcal{D}(D) \subset \mathcal{X}, \\ f &\longmapsto (D - \mathrm{i}\, id)^{-1}f, \end{aligned}$$

und es gilt

$$\|Mf\| \le c\,\|f\|$$

sowie

$$\|Mf\|_1 = \sqrt{\|Mf\|^2 + \|(Mf)'\|^2} \le c\,\|f\|.$$

Aufgrund des Rellichschen Auswahlsatzes (Satz 6.6.26) ist deshalb $M \in \mathcal{KL}(\mathcal{X})$.

Nun wollen wir die Gleichung

$$Du \stackrel{!}{=} f \tag{1}$$

lösen. Wegen $D = (D - \mathrm{i}\, id) + \mathrm{i}\, id$ und $M = (D - \mathrm{i}\, id)^{-1}$ folgt aus Gl. (1)

$$(id + \mathrm{i}\, M)u \stackrel{!}{=} Mf. \tag{2}$$

Hat man umgekehrt eine Lösung $u \in \mathcal{X}$ von Gl. (2), dann folgt aus Gl. (2)

$$u = M(f - \mathrm{i}\, u) \in \mathcal{D}(D).$$

Das heißt, u löst Gl. (1). Es genügt also, Gl. (2) zu behandeln. Für sie gelten unsere Resultate aus der Riesz-Schauder Theorie, und notwendig und hinreichend für ihre Lösbarkeit ist

$$\forall\, v \in \mathcal{N}\big(id + (\mathrm{i}\, M)^*\big) \qquad (Mf, v) = 0,$$

oder

$$\forall\, v \in \mathcal{N}\big(id + (\mathrm{i}\, M)^*\big) \qquad (f, v) = 0.$$

Nun ist

$$M^* = \big((D - \mathrm{i}\, id)^{-1}\big)^* = \big((D - \mathrm{i}\, id)^*\big)^{-1} = (D + \mathrm{i}\, id)^{-1}.$$

Deshalb folgt aus $v \in \mathcal{N}\left(id + (\mathrm{i}\, M)^*\right)$ das Verschwinden von

$$Dv = (D + \mathrm{i}\, id - \mathrm{i}\, id) = (D + \mathrm{i}\, id)(id - \mathrm{i}\, M^*)v$$

und umgekehrt. Unsere Lösungsbedingung für Gl. (1) lautet also

$$\forall\, v \in \mathcal{N}(D) \qquad (f, v) = 0. \tag{3}$$

Dabei ist mit beliebigem $v^0 \in \mathbb{C}^2$

$$v(t) = \mathrm{e}^{-At}v_0 = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} v_0.$$

Es gilt also

$$\dim \mathcal{N}(D) = \operatorname{codim} \mathcal{R}(D) = 2$$

und $\mathcal{R}(D)^{\perp} = \mathcal{N}(D)$. □

7.4 Die Integralgleichungsmethode in der Potentialtheorie

Als Anwendung der Riesz-Schauder Theorie lösen wir nun Randwertaufgaben der Potentialtheorie mit der klassischen Integralgleichungsmethode.

Es sei $G \subset \mathbb{R}^3$ ein beschränktes *glatt berandetes Gebiet* (eine zusammenhängende offene Menge mit $\mathcal{C}_1$-Rand); auch $G_a := \mathbb{R}^3 \backslash G$, das Äußere von G, sei zusammenhängend. Wir suchen Lösungen der Potentialgleichung

$$\Delta u = o$$

in G oder G_a, die bis zum Rande stetig bzw. stetig differenzierbar sind, mit

$$u|\partial G = f \quad \text{bzw.} \quad \frac{\partial u}{\partial n}\Big|\partial G = g.$$

Es soll also die *Dirichletsche* bzw. die *Neumannsche Randwertaufgabe* in G oder G_a gelöst werden. Dabei sind $f, g \in \mathcal{C}(\partial G)$ vorgegebene Funktionen und n der nach außen weisende Normalenvektor von ∂G. Im Falle der *Außenraumaufgabe* (in G_a) benötigt man zur Eindeutigkeit noch eine Bedingung für das Verhalten der Lösungen im Unendlichen. In $\mathbb{R}^3$ lautet eine solche Bedingung etwa

$$u(x) = \mathcal{O}\left(\frac{1}{|x|}\right) \quad \text{und} \quad (\nabla u)(x) = \mathcal{O}\left(\frac{1}{|x|^2}\right) \quad \text{für} \quad |x| \to \infty,$$

und zwar gleichmäßig bezüglich $x_0 := x/|x|$. Aus der Gaußschen Umformung für G bzw. G_a

$$\int_{\partial G} u \frac{\partial u}{\partial n} = \int_G \nabla(u \nabla u) = \int_G (\nabla u)^2$$

folgt dann die Eindeutigkeit der Dirichletschen Randwertaufgaben und der Neumannschen Außenraumaufgabe, während die homogene Neumannsche *Innenraumaufgabe* (in G) Konstanten als Lösungen besitzt. Es sei hier nur bemerkt, daß zum Nachweis der Eindeutigkeit $u \in \mathcal{C}_1(\bar{G})$ nicht benötigt wird, vielmehr genügt $u \in \mathcal{C}(\bar{G})$ bzw. die Existenz der Ableitung in Normalenrichtung. Die Aussage $u \in \mathcal{C}_1(\bar{G})$ ist dann für Lösungen zu verschwindenden Randwerten beweisbar.

Zur Lösung solcher Randwertaufgaben hat man schon in der Mitte des letzten Jahrhunderts spezielle Ansätze eingeführt, und zwar *Potentiale mit der einfachen Flächenbelegung* $\mu \in \mathcal{X} := \mathcal{C}(\partial G, \mathbb{R})$

$$v(x) := \frac{1}{2\pi} \int_{\partial G} \mu(y) \frac{1}{|x-y|} dy$$

und Potentiale mit der *doppelten Flächenbelegung (Dipolbelegung)* $\nu \in \mathcal{X}$

$$w(x) := \frac{1}{2\pi} \int_{\partial G} \nu(y) \frac{\partial}{\partial n_y} \frac{1}{|x-y|} dy.$$

Beide Funktionen lösen sowohl in G als auch in G_a die Potentialgleichung. Die Funktion v ist überall in $\mathbb{R}^3$ definiert und stetig; und mit etwas Rechnung folgt, daß die Ableitung von v in Normalenrichtung bei Annäherung an den Rand springt, während die Tangentialableitungen stetig bleiben. Die Potentialfunktion w hingegen ist in $\bar{G}$ und in $\bar{G}_a$ stetig, springt aber am Rand, während für $\nu \in \mathcal{C}_{1+\alpha}(\partial G)$ ihre Normalableitungen stetig bleiben. Es gilt

$$\begin{array}{ll}
w(x) \to \quad \nu(z) + (K\nu)(z) & \text{für } x \to z \in \partial G,\ x \in G_a \\
w(x) \to -\nu(z) + (K\nu)(z) & \text{für } x \to z \in \partial G,\ x \in G \\
\frac{\partial v}{\partial n}(x) \to -\mu(z) + (K^{tr}\mu)(z) & \text{für } x \to z \in \partial G,\ x \in G_a \\
\frac{\partial v}{\partial n}(x) \to \quad \mu(z) + (K^{tr}\mu)(z) & \text{für } x \to z \in \partial G,\ x \in G.
\end{array}$$

Dabei ist

$$(K\nu)(z) := \frac{1}{2\pi} \int_{\partial G} \nu(y) \frac{\partial}{\partial n_y} \frac{1}{|z-y|} dy \quad \text{für } z \in \partial G,$$

$$(K^{tr}\mu)(z) := \frac{1}{2\pi} \int_{\partial G} \mu(y) \frac{\partial}{\partial n_z} \frac{1}{|z-y|} dy \quad \text{für } z \in \partial G$$

und K^{tr} der bezüglich des $\mathcal{L}^2(\partial G)$-Skalarproduktes zu K adjungierte Operator. Die *Sprungrelationen* muß man im einzelnen nachrechnen. Dabei geht besonders ein, daß wegen der Glattheit von ∂G für $y, z \in \partial G$ und $|z-y| \to 0$

$$\frac{\partial}{\partial n_y} \frac{1}{|z-y|} = \frac{n(y)(z-y)}{|z-y|^3} = \mathcal{O}\left(\frac{1}{|z-y|}\right)$$

ist. Damit *glättet* der Operator K, es folgt

$$K,\ K^{tr}\ :\ \mathcal{C}(\partial G) \longrightarrow \mathcal{C}_{0,\alpha}(\partial G) \quad \text{mit } 0 < \alpha < \frac{1}{2}$$

und aus dem Satz von Arzelà-Ascoli dann $K, K^{tr} \in \mathcal{KL}(\mathcal{X})$. Einzelheiten sollen in dieser Vorlesung nicht besprochen werden, man vergleiche Beispiel 7.2.3.

Für die Dirichletsche Randwertaufgabe macht man nun den Ansatz $u = w$ und für die Neumannsche den Ansatz $u = v$. So erhält man für μ bzw. ν die Integralgleichungen

	Dir. RWA	Neum. RWA
G	$f = (-id + K)\nu$	$g = (id + K^{tr})\mu$
G_a	$f = (id + K)\nu$	$g = (-id + K^{tr})\mu$

und kann die Riesz-Schauder Theorie in $\mathcal{X}$ anwenden. Benötigt wird dazu

Lemma 7.4.1: *Es gilt*

$$r(id - K) = 0 \quad \textit{und} \quad r(id + K) = 1$$

und mit $e \in \mathcal{X}$, $e : z \mapsto e(z) = 1$,

$$\mathcal{N}(id + K) = \operatorname{span}(e).$$

Beweisskizze:

1. Es sei $\varphi - K\varphi = o$. Weil K glättet, ist $\varphi \in \mathcal{C}_{1+\alpha}(\partial G)$. Es sei ferner

$$w(x) := \frac{1}{2\pi} \int_{\partial G} \varphi(y) \frac{\partial}{\partial n_y} \frac{1}{|x - y|} dy.$$

Dann verschwindet w an der Innenseite von ∂G, wegen der Eindeutigkeit der Dirichletschen Randwertaufgabe mithin überall in G. Wegen $\varphi \in \mathcal{C}_{1+\alpha}(\partial G)$ existiert die Normalableitung von w am Rande und verschwindet dort. Weil auch die Neumannsche Außenraumaufgabe höchstens eine Lösung besitzt, verschwindet w überall. Dann ist aber

$$o = w|_a = \varphi + K\varphi = 2\varphi.$$

2. Es sei nun $\psi + K\psi = o$. Auch jetzt folgt $\psi \in \mathcal{C}_{1+\alpha}(\partial G)$, und es sei

$$w(x) := \frac{1}{2\pi} \int_{\partial G} \psi(y) \frac{\partial}{\partial n_y} \frac{1}{|x - y|} dy.$$

Dann erhält man aus $w|_a = o$ wegen der Eindeutigkeit $w|G_a = o$. Damit existiert und verschwindet die Normalableitung von w am Rande. Man zeigt nun leicht, daß die homogene Neumannsche Randwertaufgabe für G nur die Konstanten zur Lösung hat. Mithin ist

$$const = w|_i = -\psi + K\psi = -2\psi,$$

und die Lösung $\psi = e$ tritt wirklich auf; das rechnet man leicht nach.

3. Es seien nun $\varphi \in \mathcal{N}(id + K^{tr})$, $\varphi \neq o$ und

$$v(x) := \frac{1}{2\pi} \int_{\partial G} \varphi(y) \frac{dy}{|x - y|}.$$

Wir wollen

$$(e, \varphi) \neq 0$$

zeigen und nehmen dazu $(e, \varphi) = 0$ an. Dann ist

$$\left.\frac{\partial v}{\partial n}\right|_i = \varphi + K^{tr}\varphi = o,$$

also $v|G$ konstant, und

$$\left.\frac{\partial v}{\partial n}\right|_a = -\varphi + K^{tr}\varphi = -2\varphi.$$

Damit folgt aus dem Gaußschen Satz

$$\int_{G_a} (\nabla v)^2 = \int_{G_a} \nabla(v\nabla v) = -\int_{\partial G_a} v\,\frac{\partial v}{\partial n} = 2\,(v|\partial G, \varphi) = 0,$$

also $v|G_a = 0$, und wegen der Stetigkeit $v = 0$. Dann ist aber

$$\varphi = -\frac{1}{2}\,\frac{\partial v}{\partial n}\bigg|_a = o$$

im Widerspruch zur Annahme.

4. Als letztes zeigen wir $r(id + K) = 1$. Es sei $x \notin \mathcal{N}(id + K)$ und

$$(id + K)^2 x = o.$$

Dann ist

$$(id + K)x \in \mathcal{N}(id + K) = \mathrm{span}(e).$$

Es sei etwa $(id + K)x = e$. Dann folgt mit

$$\varphi \in \mathcal{N}(id + K^{tr}),\ \varphi \neq o,$$

$$0 = \big((id + K^{tr})\varphi, x\big) = \big(\varphi, (id + K)x\big) = (\varphi, e) \neq 0,$$

und das ist ein Widerspruch. □

Folgerung 7.4.2: *Die Dirichletsche Innenraumaufgabe und die Neumannsche Außenraumaufgabe sind eindeutig lösbar.*

Beweis: Die Existenz einer Lösung folgt aus $r(id - K) = 0$. □

Folgerung 7.4.3: *Notwendig und hinreichend zur Lösbarkeit der Neumannschen Innenraumaufgabe ist die Bedingung*

$$(e, g) = \int_{\partial G} g(y)\,dy = 0.$$

Beweis: Aus

$$\int_{\partial G} \frac{\partial u}{\partial n} = \int_G \Delta u = 0$$

folgt die Notwendigkeit der angegebenen Bedingung. Daß sie auch hinreicht, folgt aus $r(id + K) = 1$ und $\mathcal{N}(id + K) = \mathrm{span}(e)$. □

Schwieriger ist die Behandlung der Dirichletschen Außenraumaufgabe. Auch hier gilt die Eindeutigkeit, wegen $r(id + K) = 1$ ist die entsprechende Integralgleichung jedoch nicht immer lösbar; es gilt jedoch

$$\mathcal{X} = \mathcal{R}(id + K) \oplus \mathcal{N}(id + K) = \mathcal{R}(id + K) \oplus \mathrm{span}(e).$$

Man kann daher die Randwerte f in ihre Komponenten zerlegen, und eine spezielle Potentialfunktion u mit $u|\partial G = e$ läßt sich direkt angeben. Mit diesem Kunstgriff wurde lange Zeit die Dirichletsche Außenraumaufgabe in der Potentialtheorie gelöst. Das führt aber zu Schwierigkeiten, wenn man beispielsweise die Lösung u in Abhängigkeit von den Randdaten f diskutieren will.

Viel systematischer ist folgendes Vorgehen: Offenbar entsprechen Dirichletsche Außenraumaufgabe und Integralgleichung $(id + K)\nu = f$ einander nicht; der Ansatz mit dem Dipolpotential w ist in diesem Fall schlecht gewählt. Nimmt man statt dessen

$$W(x) := \frac{1}{2\pi} \int_{\partial G} \nu(y) \Big\{ \frac{\partial}{\partial n_y} + \mathrm{i} \Big\} \frac{1}{|x - y|} dy,$$

dann erhält man eine Integralgleichung der Form

$$(id + L)\nu = f$$

mit $r(id + L) = 0$.

Es sei noch betont, daß die eingangs zitierte Abschätzung

$$n(y)(z-y) = \mathcal{O}\big(|z-y|^2\big)$$

besonders wichtig ist. Aus ihr folgt ja die Kompaktheit der auftretenden Operatoren. Leider hängt sie stark von der Glattheit des Randes und der speziellen Gestalt der Grundlösung der Potentialtheorie ab. Bei nur stückweise glattem Rand ist eine solche Abschätzung aber nicht mehr richtig; auch bei anderen Differentialgleichungen bzw. Differentialgleichungssystemen aus der mathematischen Physik, wie den Maxwellschen oder den Elastizitätsgleichungen, gilt eine analoge Abschätzung nicht, schon bei glattem Rand. Hier wird man dann schnell auf *singuläre Integralgleichungen* geführt, über die viel gearbeitet worden ist. Aber letztlich bedeutet dies doch eine starke Beschränkung der Methode.

Ähnliches gilt bei unbeschränktem Rand ∂G. Zum Abschluß bringen wir hierzu noch das

Beispiel 7.4.4: *Wir betrachten in $\mathcal{L}^2(\mathbb{R}^+)$ die Fourier-Sinus Transformation, nämlich mit $k(x,y) := \sqrt{2/\pi}\sin(xy)$*

$$Kf \;:=\; \int_0^\infty k(\,\cdot\,,y)f(y)\,dy.$$

Es sei für $\alpha \in \mathbb{R}^+$

$$g_\alpha(x) := \sqrt{\pi/2}\,\mathrm{e}^{-\alpha x}.$$

Dann ist

$$\big(Kg_\alpha\big)(x) = \int_0^\infty \sin(xy)\;\mathrm{e}^{-\alpha y}\,dy = \frac{1}{\alpha}\int_0^\infty x\cos(xy)\;\mathrm{e}^{-\alpha y}\,dy = \frac{x}{\alpha^2} - \frac{x^2}{\alpha^2}\big(Kg_\alpha\big)(x),$$

also

$$\big(Kg_\alpha\big)(x) = \frac{x}{\alpha^2+x^2} =: h_\alpha(x).$$

Wir werden in §8.3 im Anschluß an den Spektralsatz

$$K = K^* = K^{-1}$$

zeigen, also

$$Kh_\alpha = g_\alpha.$$

Damit folgt

$$K(g_\alpha + h_\alpha) = h_\alpha + g_\alpha$$

oder $g_\alpha + h_\alpha \in \mathcal{N}(id-K)$. Es ist also

$$\dim\mathcal{N}(id-K) = \infty.$$

7.5 Das Spektrum kompakter linearer Abbildungen

In §7.3 haben wir für kompakte Störungen der Identität das Gelten der Fredholmschen Alternative bewiesen. Wir wollen nun das Spektrum kompakter Abbildungen diskutieren, geben vorher jedoch noch

Definition 7.5.1: *Es seien $\mathcal{X},\mathcal{Y}$ normierte Räume und $A \in \mathcal{L}\big(\mathcal{D}(A)\subset\mathcal{X},\mathcal{Y}\big)$ eine abgeschlossene Abbildung. Dann heißt A* Fredholmoperator, *wenn folgendes gilt:*

1. $\dim\mathcal{N}(A) < \infty$,
2. $\mathcal{R}(A) \subset \mathcal{Y}$ *ist abgeschlossen,*
3. $\operatorname{codim}\mathcal{R}(A) < \infty$.

Ist A ein Fredholmoperator, dann heißt die Zahl

$$\operatorname{ind} A := \dim\mathcal{N}(A) - \operatorname{codim}\mathcal{R}(A)$$

Index *von A.*

Offenbar ist *id* ein Fredholmoperator mit dem Index Null, und in §7.3 haben wir bewiesen

Satz 7.5.2: *Es seien $\mathcal{X}$ ein normierter Raum und $K \in \mathcal{KL}(\mathcal{X})$. Dann ist auch $id + K$ ein Fredholmoperator mit*

$$\operatorname{ind}(id + K) = \operatorname{ind} id.$$

Der Satz bleibt richtig, wenn man *id* durch einen Fredholmoperator $A \in \mathcal{CL}(\mathcal{X})$ ersetzt. Einen Beweis findet man in *Kato [1966], S. 238*. Dieses Buch enthält auch Resultate über kleine Störungen von Fredholmoperatoren.

Definition 7.5.3: *Es seien $\mathcal{X}$ ein normierter Raum und $A \in \mathcal{L}\big(\mathcal{D}(A) \subset \mathcal{X},\ \mathcal{X}\big)$ eine abgeschlossene Abbildung. Dann heißen*

$$\Phi(A) := \Big\{\lambda \in \mathbb{C} \ \Big|\ A - \lambda\, id \text{ ist ein Fredholmoperator}\Big\}$$

die Fredholmmenge *und*

$$\Sigma(A) := \Big\{\lambda \in \sigma(A) \ \Big|\ \lambda \notin \Phi(A)\Big\}$$

das wesentliche Spektrum *von A.*

Es ist $\Phi(A)$ offen (vgl. *Kato [1966], S. 235*) und $\Sigma(A) = \mathbb{C}\backslash\Phi(A)$ abgeschlossen. Das wesentliche Spektrum enthält das kontinuierliche, Eigenwerte unendlicher Vielfachheit und Häufungspunkte von Eigenwerten.

Wir beginnen nun mit

Satz 7.5.4: *Es seien $\mathcal{X}$ ein Banachraum und $K \in \mathcal{KL}(\mathcal{X})$. Dann gilt*

1. $\sigma(K)\backslash\{0\} = P\sigma(K)\backslash\{0\}$.
2. *Für $\lambda \in \mathbb{C}\backslash\{0\}$ ist $K - \lambda\, id$ ein Fredholmoperator mit dem Index Null.*
3. *Es sei $\dim\mathcal{X} < \infty$. Dann ist K ein Fredholmoperator mit dem Index Null.*
4. *Es sei $\dim\mathcal{X} = \infty$. Dann ist K kein Fredholmoperator.*

Es gilt dann also

$$\Sigma(K) = \begin{cases} \emptyset & \text{falls } \dim\mathcal{X} < \infty \\ \{0\} & \text{falls } \dim\mathcal{X} = \infty. \end{cases}$$

Beweis: Für $\lambda \in \mathbb{C}\backslash\{0\}$ ist

$$K - \lambda\, id = -\lambda\Big(id - \frac{1}{\lambda}K\Big)$$

ein Fredholmoperator mit dem Index Null. Das beweist die beiden ersten Aussagen. Die dritte Aussage ist trivial, wir zeigen die vierte: Es sei K ein Fredholmoperator. Dann existiert die stetige kanonische Abbildung

$$\varphi \ : \ \mathcal{X} \longrightarrow \mathcal{X}/\mathcal{N}(K),$$

und wir können

$$\tilde{K} \ : \ \mathcal{X}/\mathcal{N}(K) \longrightarrow \mathcal{R}(K)$$

vermöge $\tilde{K}[x] := Kx$ definieren. Dann ist $\tilde{K} \in \mathcal{KL}\big(\mathcal{X}/\mathcal{N}(K), \mathcal{R}(K)\big)$. Nun sind $\mathcal{X}/\mathcal{N}(K)$ und $\mathcal{R}(K)$ Banachräume; aus dem Satz von der inversen Abbildung folgt deshalb

$$\tilde{K}^{-1} \in \mathcal{CL}\left(\mathcal{R}(K),\, \mathcal{X}/\mathcal{N}(K)\right).$$

Damit ist

$$id_{\mathcal{X}/\mathcal{N}(K)} = \tilde{K}^{-1} \circ \tilde{K} \in \mathcal{KL}\left(\mathcal{X}/\mathcal{N}(K)\right),$$

und aus Satz 7.2.2,4 folgt $\dim\mathcal{X}/\mathcal{N}(K) < \infty$. Das ist ein Widerspruch. □

Satz 7.5.5: *Es seien $\mathcal{X}$ ein normierter Raum und $K \in \mathcal{KL}(\mathcal{X})$. Dann ist $\sigma(K)$ abzählbar und besitzt höchstens Null als Häufungspunkt.*

Beweis: Wir zeigen, daß für alle $n \in \mathbb{N}$ die Mengen

$$\Lambda_n := \Big\{\lambda \in \sigma(K) \ \Big|\ |\lambda| \geq \frac{1}{n}\Big\}$$

endlich sind. Es gebe nämlich eine Folge $(\lambda_k)_{k\in\mathbb{N}}$ mit

$$\lambda_k \in \sigma(K),\ \ \lambda_i \neq \lambda_j \text{ für } i \neq j \quad \text{und} \quad |\lambda_k| \geq \frac{1}{n}.$$

Dann gibt es Eigenfunktionen $x_k \neq o$ mit

$$Kx_k = \lambda_k x_k.$$

Diese Eigenfunktionen sind linear unabhängig. Das folgt leicht durch vollständige Induktion. Es seien $x_1, \ldots, x_{m-1}$ bereits linear unabhängig und

$$x_m = \sum_{i=1}^{m-1} \alpha_i x_i.$$

Dann folgt

$$\lambda_m x_m = Kx_m = \sum_{i=1}^{m-1} \alpha_i \lambda_i x_i,$$

also

$$o = \sum_{i=1}^{m-1} \Big(1 - \frac{\lambda_i}{\lambda_m}\Big)\alpha_i x_i$$

und $\alpha_i = \cdots = \alpha_{m-1} = 0$. Es sei nun

$$\mathcal{X}_m \ := \ \operatorname{span}(x_1, \ldots, x_m).$$

Dann ist $\mathcal{X}_{m-1} \subsetneq \mathcal{X}_m$, und aus dem Rieszschen Lemma folgt die Existenz von $y_m \in \mathcal{X}_m$ mit $\|y_m\| = 1$ und

$$\operatorname{dist}(y_m, \mathcal{X}_{m-1}) \geq \frac{1}{2}.$$

Für $m > l \geq 1$ gilt

$$Ky_m - Ky_l = \lambda_m(y_m - x)$$

mit

$$x := \frac{1}{\lambda_m}\Big(Ky_l + \lambda_m y_m - Ky_m\Big) \in \mathcal{X}_{m-1},$$

also

$$\big\|Ky_m - Ky_l\big\| = |\lambda_m| \, \big\|y_m - x\big\| \geq \frac{1}{2n} > 0$$

im Widerspruch zu $K \in \mathcal{KL}(\mathcal{X})$. Mithin ist Λ_n endlich. □

Beispiel 7.5.6: *Der Wert $0 \in \Sigma(K)$ kann zum kontinuierlichen, Residuen- oder Punktspektrum gehören.*

1. Es sei $K = O$. Dann ist $\mathcal{N}(K) = \mathcal{X}, \mathcal{R}(K) = \{o\}$ und
$$\sigma(K) = P\sigma(K) = \{0\}.$$

2. Es seien $\mathcal{X} = \ell^2$ und $Kx := \big(\xi_1, \frac{1}{2}\xi_2, \frac{1}{3}\xi_3, \ldots\big)$. Es ist $K^* = K$. Es sei ferner $K_N \in \mathcal{KL}(\mathcal{X})$ mit
$$K_N x := \big(\xi_1, \tfrac{1}{2}\xi_2, \ldots, \tfrac{1}{N}\xi_N, 0, \ldots\big).$$
Dann gilt für $M < N$
$$\|K_N - K_M\| = \sup_{\|x\|=1} \Big\| \sum_{i=M+1}^{N} \frac{\xi_i}{i} e_i \Big\| \leq \sqrt{\sup_{\|x\|=1} \sum_{i=M+1}^{N} \xi_i^2} \sqrt{\sum_{i=M+1}^{N} \frac{1}{i^2}} \leq \sqrt{\sum_{i=M+1}^{\infty} \frac{1}{i^2}} \to 0.$$
Mithin ist $K \in \mathcal{KL}(\mathcal{X})$. Als erstes bestimmen wir die von Null verschiedenen Eigenwerte von K. Aus $K_\lambda x = o$ folgt
$$\xi_1 = \lambda\,\xi_1, \ \ \frac{1}{2}\xi_2 = \lambda\,\xi_2, \ \ \ldots \ .$$

Für jedes $n \in \mathbb{N}$ ist $\lambda_n := 1/n$ mithin Eigenwert zu e_n, also

$$P\sigma(K) = \left\{1, \frac{1}{2}, \frac{1}{3}, \ldots\right\}.$$

Damit liegt $\lambda_0 := 0$ in $\sigma(K)$, gehört jedoch nicht zum Punktspektrum. Es ist $\mathcal{R}(K) \neq \mathcal{X}$, denn für

$$f := \left(1, \frac{1}{2}, \frac{1}{3}, \ldots\right) \in \mathcal{X}$$

ist

$$Kx \stackrel{!}{=} f$$

in $\mathcal{X}$ nicht lösbar. Es wäre nämlich $x = (1, 1, \ldots)$. Es ist jedoch $\overline{\mathcal{R}(K)} = \mathcal{X}$. Dazu muß man ein vorgegebenes $f \in \mathcal{X}$ nur durch eine endliche Summe $\Sigma\varphi_i e_i$ approximieren. Damit gilt $\lambda_0 \in C\sigma(K)$.

3. Es sei wieder $\mathcal{X} = \ell^2$ und $Kx := \left(0, \xi_1, \frac{1}{2}\xi_2, \frac{1}{3}\xi_3, \ldots\right)$. Dann ist K kompakt und

$$K^*x = \left(\xi_2, \tfrac{1}{2}\xi_3, \tfrac{1}{3}\xi_4, \ldots\right),$$

also $\mathcal{N}(K) = \{o\}$ und $\mathcal{N}(K^*) = \operatorname{span}(e_1)$. Es folgen

$$\sigma(K) = \Sigma(K) = R\sigma(K) = \{0\},$$

$$\sigma(K^*) = \Sigma(K^*) = P\sigma(K^*) = \{0\}. \qquad \square$$

Es sei auch an Beispiel 7.1.4 (7.2.4) erinnert. Wir fassen unsere Resultate zusammen im

Spektralsatz für kompakte Abbildungen: *Es seien $\mathcal{X}$ ein Banachraum und $K \in \mathcal{KL}(\mathcal{X})$. Dann gilt*

1. *$\sigma(K)\backslash\{0\}$ besteht höchstens aus abzählbar vielen Eigenwerten mit Null als einzig möglichem Häufungspunkt.*
2. *$id - \lambda K$ ist ein Fredholmoperator mit dem Index Null. Die Rieszzahl $r(\lambda)$ von $id - \lambda K$ existiert, und es gilt*
$$\mathcal{X} = \mathcal{N}\left((id - \lambda K)^{r(\lambda)}\right) \oplus \mathcal{R}\left((id - \lambda K)^{r(\lambda)}\right).$$
3. *Eigenfunktionen zu verschiedenen Eigenwerten von K sind linear unabhängig.*
4. *Im Falle $\dim \mathcal{X} = \infty$ ist $0 \in \Sigma(K)$, andernfalls ist $0 \in \Phi(K)$.*

7.6 Das Spektrum kompakter normaler Abbildungen

In diesem Abschnitt seien $\mathcal{X}$ ein Hilbertraum und $K \in \mathcal{KL}(\mathcal{X})$. Wir wollen uns mit der besonders wichtigen Klasse kompakter normaler Abbildungen beschäftigen.

Mit K ist nach Satz 7.2.6 auch die konjugierte Abbildung K' kompakt. In §5.3 haben wir in Hilberträumen vermöge

$$K^* := R_{\mathcal{X}}^{-1} K' R_{\mathcal{X}}$$

auch die zu K adjungierte Abbildung definiert. Nach Satz 7.2.7 ist deshalb mit K' auch K^* kompakt, also $K^* \in \mathcal{KL}(\mathcal{X})$. Wir geben nun

Definition 7.6.1: *Es sei $\mathcal{X}$ ein Hilbertraum. Dann heißt $A \in \mathcal{CL}(\mathcal{X})$* normal, *wenn $AA^* = A^*A$ ist. $A \in \mathcal{CL}(\mathcal{X})$ heißt* hermitesch, *wenn $A = A^*$ ist.*

Eine hermitesche Abbildung ist selbstadjungiert, und mit A ist auch A_λ normal. Das folgt aus

$$A_\lambda^* = A^* - \bar{\lambda}\, id$$

durch Ausmultiplizieren. Ferner zeigen wir

Lemma 7.6.2: *Es sei $A \in \mathcal{CL}(\mathcal{X})$. Dann gilt:*

1. *A ist normal $\iff \forall\, x \in \mathcal{X} \quad \|Ax\| = \|A^*x\|$.*
2. *A ist hermitesch $\iff \forall\, x \in \mathcal{X} \quad (Ax, x) \in \mathbb{R}$.*

Beweis:

1. $\Rightarrow$: $(Ax, Ax) = (x, A^*Ax) = (x, AA^*x) = (A^*x, A^*x)$.
   $\Leftarrow$: Aus der Polarisationsformel folgt

$$(A^*Ax, y) = (Ax, Ay) = (A^*x, A^*y) = (AA^*x, y).$$

2. Das haben wir bereits in Satz 5.3.14 bewiesen. □

Folgerung 7.6.3: *Es sei $A \in \mathcal{CL}(\mathcal{X})$ normal. Dann gilt:*

1. $\mathcal{N}(A_\lambda) = \mathcal{N}(A_\lambda^*)$.

2. $\forall\, n \in \mathbb{N} \quad \|A^{2^n}\| = \|A\|^{2^n}$.

Beweis:

1. Das folgt aus $\|A_\lambda x\| = \|A_\lambda^* x\|$, weil A_λ normal ist.

2. Es ist $\|A\| = \|A^*\|$ und $\|A^*A\| = \|A\|^2$ nach Lemma 5.3.4. Wegen

$$(A^*)^2 A^2 = A^*AA^*A = AA^*AA^* = A^2(A^*)^2$$

ist mit A auch A^2 normal und

$$\left\|A^2\right\|^2 = \left\|(A^2)^*A^2\right\| = \left\|(A^*A)^*(A^*A)\right\| = \left\|A^*A\right\|^2 = \left\|A\right\|^4,$$

also

$$\left\|A^2\right\| = \left\|A\right\|^2.$$

Das beweist die Aussage für $n = 1$. Es sei $m = 2^n$. Dann ist auch A^m normal, und ein Induktionsschluß zeigt

$$\left\|(A^m)^2\right\| = \left\|A^m\right\|^2 = (\|A\|^m)^2 = \left\|A\right\|^{2m},$$

also

$$\left\|A^{2^{n+1}}\right\| = \left\|A\right\|^{2^{n+1}}.$$ □

Lemma 7.6.4: *Es sei $A \in \mathcal{CL}(\mathcal{X})$ normal. Dann gilt*

$$\forall\, n \in \mathbb{N} \qquad \left\|A^n\right\|^{1/n} = \|A\| = r_\sigma(A).$$

Beweis: In Satz 7.1.12 haben wir die Existenz von

$$r_\sigma(A) = \lim_{n\to\infty} \left\|A^n\right\|^{1/n}$$

bewiesen, und es war

$$r_\sigma(A) \le \left\|A^n\right\|^{1/n} \le \|A\|.$$

Folgerung 7.6.3 zeigt dann für $n = 2^k$

$$r_\sigma(A) = \|A\|,$$

also

$$\|A\| \le \left\|A^n\right\|^{1/n} \le \|A\|.$$ □

Wir geben noch

Definition 7.6.5: *Eine Abbildung $A \in \mathcal{CL}(\mathcal{X})$ heißt* nichtnegativ, *wenn*

$$\forall\, x \in \mathcal{X} \qquad (Ax, x) \ge 0$$

ist, und positiv, *wenn*

$$\forall\, x \in \mathcal{X} \qquad (Ax, x) > 0, \quad x \ne 0$$

ist.

Dafür schreiben wir auch $A \ge 0$, bzw. $A > 0$. Eine nichtnegative Abbildung ist also hermitesch. Die Abbildungen A^*A und AA^* sind stets nichtnegativ.

Als letztes vorbereitendes Resultat zeigen wir schließlich

Lemma 7.6.6: *Es sei $A \in \mathcal{CL}(\mathcal{X})$ hermitesch. Dann ist*

$$\|A\| = \sup_{\|x\|=1} |(Ax, x)|.$$

Allgemein würde nur

$$\|A\| = \sup_{\|x\|=1} \sup_{\|y\|=1} |(Ax, y)|$$

gelten.

Zum Beweis wählen wir $x \in \mathcal{X}$ mit $\|x\| \neq 0$ und $Ax \neq o$,

$$\alpha := \sqrt{\|Ax\|/\|x\|} > 0 \quad \text{und} \quad \lambda := \sup_{\|x\|=1} |(Ax, x)|.$$

Dann ist

$$\|Ax\|^2 = \frac{1}{4}\left\{ \left(A\Big(\alpha x + \frac{Ax}{\alpha}\Big), \alpha x + \frac{Ax}{\alpha}\right) - \left(A\Big(\alpha x - \frac{Ax}{\alpha}\Big), \alpha x - \frac{Ax}{\alpha}\right)\right\} \leq$$
$$\leq \frac{1}{4}\left\{ \lambda \Big\|\alpha x + \frac{Ax}{\alpha}\Big\|^2 + \lambda \Big\|\alpha x - \frac{Ax}{\alpha}\Big\|^2 \right\} = \frac{\lambda}{2}\left\{ \alpha^2 \|x\|^2 + \frac{\|Ax\|^2}{\alpha^2} \right\} = \lambda \|x\| \, \|Ax\|,$$

also

$$\|Ax\| \leq \lambda \|x\|$$

oder $\|A\| \leq \lambda$. Die Abschätzung $\lambda \leq \|A\|$ ist trivial. □

Wir beginnen nun mit der Diskussion kompakter normaler Abbildungen. Wesentlich ist das folgende

Lemma 7.6.7: *Es sei $K \in \mathcal{KL}(\mathcal{X})$ normal. Dann existiert ein $\lambda \in P\sigma(K)$ mit $|\lambda| = \|K\|$.*

Beweis: Für $K = O$ ist das Lemma richtig, es sei also $\|K\| > 0$. Dann ist

$$\sigma(K) \backslash \{o\} \subset P\sigma(K)$$

und

$$r_\sigma(K) = \sup_{\lambda \in \sigma(K)} |\lambda| = \|K\| > 0.$$

$P\sigma(K)$ besitzt keine von Null verschiedenen Häufungspunkte. Deshalb gibt es ein λ mit

$$\lambda \in P\sigma(K) \quad \text{und} \quad |\lambda| = \|K\|. \qquad \square$$

Es sei nun $\{\lambda_n\}_{n \in N}$, $N \subset \mathbb{N}_0$, die Menge der paarweise verschiedenen Eigenwerte von K. Dabei sei $\lambda_0 = 0$, falls $\mathcal{N}_0 := \mathcal{N}(K) \neq \{o\}$ ist. Es seien ferner $\mathcal{N}_n := \mathcal{N}(id - \lambda_n K)$ und P_n die orthogonale Projektion auf $\mathcal{N}_n$. Dann gilt der

Spektralsatz für kompakte normale Abbildungen: *Es sei $K \in \mathcal{KL}(\mathcal{X})$ normal. Dann gilt:*

1. *Es gibt einen Eigenwert λ mit $|\lambda| = \|K\| = r_\sigma(\lambda)$.*
2. $\forall\, k, l \in N \quad P_k P_l = \delta_{kl} P_k$.
3. $id = \sum_{i \in N} P_i$.
4. $K = \sum_{i \in N} \lambda_i P_i \quad$ *und* $\quad K^* = \sum_{i \in N} \bar{\lambda}_i P_i$.
5. $\forall\, \lambda \in \rho(K) \quad R(K, \lambda) = \sum_{i \in N} \frac{1}{\lambda_i - \lambda} P_i$.
6. $\forall\, k \in N \quad r(id - \lambda_k K) = 1$.
7. $\forall\, k \in N \quad \mathcal{X} = \mathcal{R}(id - \lambda_k K) \oplus \mathcal{N}(id - \lambda_k K)$.

Beweis:

1. Das wurde in Lemma 7.6.7 bewiesen.

2. Für $k,l \in N$ und $x_k \in \mathcal{N}_k$ ist wegen $\mathcal{N}(id-\lambda K) = \mathcal{N}\big((id-\lambda K)^*\big)$

$$\lambda_k(x_k,x_l) = (Kx_k,x_l) = (x_k,K^*x_l) = (x_k,\bar{\lambda}_l x_l) = \lambda_l(x_k,x_l).$$

3. Wir zeigen

$$\mathcal{X} = \overline{\bigoplus_{k\in N} \mathcal{N}_k}.$$

Es sei nämlich

$$y \in \mathcal{Y} := \Big(\overline{\bigoplus_{k\in N} \mathcal{N}_k}\Big)^{\perp}$$

und $y \neq o$. Dann folgt für alle $n \in N$

$$(Ky,x_n) = (y,K^*x_n) = \lambda_n(y,x_n) = 0$$

$$(x_n,K^*y) = (Kx_n,y) = \lambda_n(x_n,y) = 0,$$

also

$$K\mathcal{Y} \subset \mathcal{Y} \quad \text{und} \quad K^*\mathcal{Y} \subset \mathcal{Y}.$$

Es seien nun K_0 bzw. K_0^* die Einschränkungen von K bzw. K^* auf $\mathcal{Y}$. Dann gilt $K_0 \in \mathcal{KL}(\mathcal{Y})$ und

$$\forall\, y,z \in \mathcal{Y} \qquad (K_0y,z) = (Ky,z) = (y,K^*z) = (y,K_0^*z),$$

das heißt, K_0 und K_0^* sind in $\mathcal{Y}$ adjungiert. K_0 ist normal. Wegen $Ky \neq o$ ist $K_0 \neq O$, folglich existieren ein $\lambda \in \mathbb{C}$ und ein $z \in \mathcal{Y}$ mit $Kz = K_0z = \lambda z$ und $|\lambda| = \|K_0\| > 0$. Das ist ein Widerspruch. Die Behauptung folgt dann aus der Theorie der Orthonormalsysteme. Für alle $n > 0$ ist $\dim \mathcal{R}(P_n)$ endlich.

4. Aus

$$x = \sum_{i\in N} P_i x$$

folgt wegen $K \in \mathcal{CL}(\mathcal{X})$

$$Kx = \sum_{i\in N\setminus\{o\}} \lambda_i\, P_i x \quad \text{und} \quad K^*x = \sum_{i\in N\setminus\{o\}} \bar{\lambda}_i\, P_i x.$$

5. Es sei $y := (K-\lambda)^{-1}x$ für $\lambda \in \rho(K)$. Dann ist

$$\sum_{i\in N} P_i x = x = (K-\lambda)y = \sum_{i\in N} (\lambda_i-\lambda) P_i y,$$

also

$$P_i x = (\lambda_i-\lambda) P_i y$$

oder

$$y = \sum_{i\in N} P_i y = \sum_{i\in N} \frac{1}{\lambda_i-\lambda} P_i x.$$

Die Resolvente von K besitzt also Pole erster Ordnung in den λ_i.

6. Wir zeigen

$$\mathcal{N}_k = \mathcal{N}(id-\lambda_k K) = \mathcal{N}\big((id-\lambda_k K)^2\big).$$

Es sei nämlich $x \in \mathcal{N}\big((id-\lambda_k K)^2\big)$. Dann folgt

$$o = (id-\lambda_k K)^2 x = (id-\lambda_k K)\Big(\sum_{i\in N} (\lambda_i-\lambda_k) P_i x\Big) = \sum_{i\in N} (\lambda_i-\lambda_k)^2\, P_i x,$$

also

$$\forall\, i \neq k \quad P_i x = o.$$

Mithin ist $x \in \mathcal{N}_k$ und $r(id - \lambda_k K) = 1$.

7. Die Behauptung folgt aus $r(id - \lambda_k K) = 1$. □

Für die Anwendungen sind die hermiteschen Abbildungen besonders interessant. Es sei also $K = K^*$. Dann ist $\sigma(K) \subset \mathbb{R}$, genauer

$$\sigma(K) \subset \big[- \|K\|, \|K\| \big].$$

Ist speziell $K \geq 0$, dann gilt sogar

$$\sigma(K) \subset \big[0, \|K\| \big].$$

Für kompakte hermitesche Abbildungen soll noch ein zweiter Beweis des Spektralsatzes vorgestellt werden, der den Zusammenhang mit der Variationsrechnung besser erkennen läßt. Zur Vereinfachung sei $K \geq 0$. Wir zeigen zunächst

Lemma 7.6.8: *Es sei $K \in \mathcal{KL}(\mathcal{X})$ hermitesch mit*

$$\lambda := \sup_{\|x\|=1} (Kx, x) > 0.$$

Dann gilt

1. $\exists\, x \in \mathcal{X}$ *mit* $\|x\| = 1$ *und* $(Kx, x) = \lambda$.
2. $\forall\, x \in \mathcal{X}$ *mit* $\|x\| = 1$ *und* $(Kx, x) = \lambda$ *ist* $Kx = \lambda x$.

Beweis:

1. Es seien $x_n \in S_{\mathcal{X}}$ mit $(Kx_n, x_n) \to \lambda$. Der Hilbertraum $\mathcal{X}$ ist reflexiv. Mithin gibt es eine Teilfolge von (x_n), die wir wieder (x_n) nennen, und ein $x \in \mathcal{X}$ mit (vgl. Satz 4.1.4 und 4.0.6)

$$\text{w-}\lim_{n\to\infty} x_n = x \quad \text{und} \quad \|x\| \leq 1.$$

Weil K kompakt ist, folgt

$$Kx_n \to Kx,$$

also

$$\lambda \leftarrow (Kx_n, x_n) = (Kx_n - Kx, x_n) + (Kx, x_n) \to (Kx, x).$$

Aus $\lambda = \sup_{\|x\|=1}(Kx, x)$ folgt dann schließlich $\|x\| = 1$.

2. Es seien $x \in \mathcal{X}$ mit $\|x\| = 1$ und $(Kx, x) = \lambda$, $y \in \mathcal{X}$ mit $\|y\| = 1$ und $a \in \mathbb{C}$ beliebig. Dann ist

$$\big(K(x + ay), x + ay\big) = (Kx, x) + 2\,\mathrm{Re}\,\bar{a}(Kx, y) + |a|^2(Ky, y) \geq \lambda + 2\,\mathrm{Re}\,\bar{a}(Kx, y) - |a|^2\,\|K\|$$

und

$$\|x + ay\|^2 = 1 + 2\,\mathrm{Re}\,\bar{a}(x, y) + |a|^2.$$

Aus

$$\big(K(x + ay), x + ay\big) \leq \lambda\,\|x + ay\|^2$$

folgt also

$$\lambda + 2\,\mathrm{Re}\,\bar{a}(Kx, y) - |a|^2\|K\| \leq \lambda + 2\lambda\,\mathrm{Re}\,\bar{a}(x, y) + \lambda\,|a|^2$$

oder

$$2\,\mathrm{Re}\,\bar{a}(Kx - \lambda x, y) \leq |a|^2(\lambda + \|K\|).$$

Wählen wir nun

$$a := \frac{(Kx - \lambda x, y)}{\lambda + \|K\|},$$

dann erhalten wir

$$\big|(Kx - \lambda x, y)\big| \leq 0$$

und deshalb $Kx = \lambda x$. □

Damit haben wir einen ersten Eigenwert λ_1,

$$\lambda_1 := \sup_{\|x\|=1} (Kx, x),$$

und eine zugehörige Eigenfunktion x_1 mit $\|x_1\| = 1$ und

$$Kx_1 = \lambda_1 x_1,$$

gefunden.

Wir wollen nun den nächsten Eigenwert bestimmen. Dazu setzen wir $\mathcal{X}_1 := \mathcal{X}$, $K_1 := K$ und wählen

$$\mathcal{X}_2 := \left\{ x \in \mathcal{X} \;\middle|\; (x, x_1) = 0 \right\} \quad \text{und} \quad K_2 := K_1 \big| \mathcal{X}_2.$$

Wegen

$$\forall\, y \in \mathcal{X}_2 \qquad (K_2 y, x_1) = (y, Kx_1) = \lambda_1 (y, x_1) = 0$$

ist $K_2 \in \mathcal{L}(\mathcal{X}_2)$ und damit $K_2 \in \mathcal{KL}(\mathcal{X}_2)$ und nichtnegativ. Deshalb können wir das Konstruktionsverfahren wiederholen. Falls

$$\sup_{x \in S_{\mathcal{X}_2}} (K_2 x, x) = 0$$

ist, brechen wir ab; andernfalls existieren

$$\lambda_2 := \sup_{x \in S_{\mathcal{X}_2}} (K_2 x, x) > 0 \quad \text{und} \quad x_2 \in \mathcal{X}_2$$

mit

$$Kx_2 = \lambda_2 x_2, \;\; \|x_2\| = 1 \quad \text{und} \quad (x_1, x_2) = 0.$$

So fahren wir fort und erhalten Eigenwerte

$$\lambda_1 \geq \lambda_2 \geq \cdots$$

und Eigenfunktionen

$$x_n \in \mathcal{X}_n := \left\{ x \in \mathcal{X} \;\middle|\; \forall\, i = 1, \ldots, n-1 \quad (x, x_i) = 0 \right\}$$

mit $\|x_n\| = 1$ und

$$Kx_n = \lambda_n x_n.$$

Es sind zwei Fälle möglich:

1. Das Verfahren bricht nach N Schritten ab. Dann existieren

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N > 0 \quad \text{und} \quad x_1, \ldots, x_N$$

mit

$$Kx_i = \lambda_i x_i.$$

Die $\{x_1, \ldots, x_N\}$ bilden ein Orthonormalsystem, und es ist

$$\operatorname{span}(x_1, \ldots, x_N)^\perp = \mathcal{N}(K)$$

sowie

$$\forall\, x \in \mathcal{X} \qquad Kx = \sum_{i=1}^{N} \lambda_i \, (x, x_i) \, x_i.$$

Insbesondere ist $\dim \mathcal{R}(K) = N < \infty$.

2. Es ist $\dim \mathcal{R}(K) = \infty$. In diesem Falle bricht das Verfahren nicht ab, und es gilt

a) $\{\lambda_n\}$ ist eine Nullfolge. Das folgt aus $K \in \mathcal{KL}(\mathcal{X})$.

b) Es ist

$$\operatorname{span}(x_1, x_2, \ldots)^\perp = \mathcal{N}(K).$$

Es sei nämlich $x \in \operatorname{span}(x_1, x_2, \ldots)^\perp$ mit $\|x\| = 1$. Dann gilt erst recht

$$x \in \mathcal{X}_n = \operatorname{span}(x_1, x_2, \ldots, x_{n-1})^\perp$$

und aus

$$\lambda_n = \sup_{x \in S_{\mathcal{X}_n}} (K_n x, x) = \|K_n\|$$

folgt für alle n

$$\|Kx\| = \|K_n x\| \leq \lambda_n \to 0,$$

also $x \in \mathcal{N}(K)$. Die andere Richtung ist klar.

c) Es gilt

$$\forall\, x \in \mathcal{X} \qquad Kx = \sum_{i=1}^{\infty} \lambda_i (x, x_i) x_i.$$

Das folgt unmittelbar aus

$$x - \sum_{i=1}^{\infty} (x, x_i) x_i \in \operatorname{span}(x_1, x_2, \ldots)^{\perp} = \mathcal{N}(K).$$

d) $\{\lambda_1, \lambda_2, \ldots\}$ ist die Menge aller positiven Eigenwerte von K. Es sei nämlich $\lambda \in P\sigma(K)$ mit $\lambda > 0$ und

$$\lambda \notin \{\lambda_1, \lambda_2, \ldots\}.$$

Es sei ferner $x \in \mathcal{X}$ mit $\|x\| = 1$ und

$$Kx = \lambda x.$$

Dann folgt für alle $n \in \mathbb{N}$

$$\lambda (x, x_n) = (Kx, x_n) = (x, Kx_n) = \lambda_n (x, x_n),$$

also wegen $\lambda \neq \lambda_n$

$$(x, x_n) = 0.$$

Damit ist $x \in \mathcal{N}(K)$. Das ist ein Widerspruch. □

Bei manchen theoretischen und praktischen Anwendungen stört es, daß man bei diesem Verfahren zur Bestimmung von λ_n die Eigenfunktion $x_1, \ldots, x_{n-1}$ bereits kennen muß. Das macht Abschätzen schwierig. Das *Courantsche Maximum-Minimum Prinzip*, benannt nach Richard Courant, 1888–1972, vermeidet dies. Es soll noch kurz vorgestellt werden; es sei aber betont, daß zu seinem Nachweis der Spektralsatz benötigt wird. Das Prinzip enthält also keinen neuen Existenzsatz. Man findet es in *Courant [1920]*.

Das Courantsche Maximum-Minimum Prinzip: *Es seien $K \in \mathcal{KL}(\mathcal{X})$ hermitesch, $h_1, \ldots, h_{n-1} \in \mathcal{X}$ und*

$$\nu_n(h_1, \ldots, h_{n-1}) := \sup_{x \in S_{\mathcal{H}_{n-1}}} (Kx, x)$$

mit $\mathcal{H}_{n-1} := \operatorname{span}(h_1, \ldots, h_{n-1})^{\perp}$. Dann erhält man den n-ten Eigenwert λ_n von K durch

$$\lambda_n = \inf_{h_1, \ldots, h_{n-1}} \nu_n(h_1, \ldots, h_{n-1}).$$

Beweis: Aus dem soeben Bewiesenen folgt

$$\lambda_n = \nu_n(x_1, \ldots, x_{n-1}).$$

Wir haben also nur noch

$$\forall\, h_1, \ldots, h_{n-1} \in \mathcal{X} \qquad \nu_n(h_1, \ldots, h_{n-1}) \geq \lambda_n$$

zu zeigen. Es seien also $h_1, \ldots, h_{n-1}$ vorgegeben. Dann wählen wir

$$x = \sum_{i=1}^{n} \xi_i\, x_i$$

mit $\xi_i \in \mathbb{C}$, $\|x\| = 1$ und $x \in \mathcal{H}_{n-1}$ Eine solche Wahl ist möglich, weil das Gleichungssystem

$$0 = \sum_{i=1}^{n} \xi_i (x_i, h_k) \quad \text{für } k = 1, \ldots, n-1$$

$$1 = \sum_{i=1}^{n} |\xi_i|^2$$

stets nach den ξ_i aufgelöst werden kann. Dann folgt

$$(Kx,x) = \sum_{i,k=1}^{n} \xi_i\,\bar{\xi}_k (Kx_i,x_k) = \sum_{i=1}^{n} \lambda_i |\xi_i|^2 \geq \lambda_n \sum_{i=1}^{n} |\xi_i|^2 = \lambda_n,$$

also

$$\nu_n(h_1,\ldots,h_{n-1}) \geq \lambda_n. \qquad \square$$

Wir beschließen diesen Abschnitt mit dem folgenden Beispiel.

Beispiel 7.6.9: *Es seien $\Omega \subset \mathbb{R}^n$ eine offene beschränkte Menge, $\mathcal{X} = \mathcal{L}^2(\Omega)$, u^0, $u^1 \in \mathcal{X}$,*

$$\mathcal{D}(A) := \left\{ u \in \mathring{\mathcal{H}}_1(\Omega) \;\middle|\; \Delta u \in \mathcal{X} \right\}$$

und $Au := -\Delta u$. Es soll die Dirichletsche Anfangsrandwertaufgabe

$$u'' + Au = 0 \quad \textit{mit } u(0,\cdot) = u^0 \textit{ und } u'(0,\cdot) = u^1$$

gelöst werden. Dabei ist $u : \mathbb{R}_0^+ \times \Omega \to \mathbb{R}$, der genaue Lösungsbegriff wird im folgenden noch präzisiert.

Zur Lösung dieser Anfangsrandwertaufgabe macht man den klassischen Ansatz der *Separation der Variablen* (oder der Entwicklung nach *stehenden Wellen*), nämlich

$$u(t,x) \stackrel{!}{=} v(x)\,w(t).$$

Aus diesem Ansatz folgen mit $\lambda \in \mathbb{C}$

$$Av \stackrel{!}{=} \lambda v \tag{1}$$

$$-w'' \stackrel{!}{=} \lambda w. \tag{2}$$

Die zweite Gleichung ist leicht zu lösen, nämlich durch

$$w(t) = c_1 \cos\sqrt{\lambda}\,t + c_2 \sin\sqrt{\lambda}\,t.$$

Bei der ersten handelt es sich um eine *Eigenwertaufgabe*.

Wir wissen bereits aus §6.7.3, daß für alle $f \in \mathcal{X}$

$$Au \stackrel{!}{=} f$$

eindeutig lösbar ist. Ein Beweis sei noch einmal skizziert. Weil Ω beschränkt ist, gilt

$$\forall\, u \in \mathring{\mathcal{H}}_1(\Omega) \qquad \|u\| \leq d\,\|\nabla u\|$$

(Erste Poincarésche Abschätzung). Es sei

$$B(u,v) := (\nabla u, \nabla v)$$

und $|u| := \sqrt{B(u,u)}$. Dann folgt, etwa mit $p := 1/\sqrt{1+d^2}$,

$$p\,\|u\|_1 \leq |u| \leq \|u\|_1.$$

Deshalb können wir $\mathring{\mathcal{H}}_1(\Omega)$ mit dem Skalarprodukt $B(u,v)$ ausstatten, und der Rieszsche Darstellungssatz liefert die Existenz eines $u \in \mathring{\mathcal{H}}_1(\Omega)$ mit

$$\forall\, \varphi \in \mathring{\mathcal{H}}_1(\Omega) \qquad F\varphi := (f,\varphi) \stackrel{!}{=} B(u,\varphi),$$

$\|F\| = |u|$ und $\|F\| \leq d\,\|f\|$. Daraus folgt sogar $u \in \mathcal{D}(A) \subset \mathcal{X}$. Es sei

$$\begin{aligned} K \;:\; & \mathcal{X} \longrightarrow \mathring{\mathcal{H}}_1(\Omega), \\ & f \longmapsto u. \end{aligned}$$

Dann ist $K = A^{-1}$,

$$\|Kf\| \leq d\,|Kf| = d\,\|F\| \leq d^2\,\|f\|$$

und

$$\|Kf\|_1 \leq \frac{1}{p}\,|Kf| \leq \frac{d}{p}\,\|f\|.$$

Aufgrund des Rellichschen Auswahlsatzes ist deshalb $K \in \mathcal{KL}(\mathcal{X})$.

Gl. (1) ist damit äquivalent zu

$$v \stackrel{!}{=} \lambda\, K\, v. \tag{3}$$

Für $f, g \in \mathcal{X}$ seien $u := Kf$ und $v := Kg$. Dann gilt

$$(Kf, g) = (u, g) = \overline{(g, u)} = \overline{B(v, u)} = B(u, v) = (f, v) = (f, Kg).$$

Mithin ist $K = K^*$, und es folgt die Existenz von Eigenwerten μ_n und zugehörenden orthogonalen Eigenfunktionen mit

$$Kv_n = \mu_n v_n.$$

Wegen $\mathcal{R}(K) = \mathcal{D}(A)$ ist $\dim \mathcal{R}(K) = \infty$, also gibt es abzählbar unendlich viele Eigenwerte μ_n mit $|\mu_n| \to 0$. Wegen

$$\mu_n = (v_n, Kv_n) = B(Kv_n, Kv_n) > 0$$

sind alle μ_n positiv. Es sei $\lambda_n := 1/\mu_n$. Dann ist

$$0 < \lambda_1 \leq \lambda_2 \leq \cdots \to \infty$$

und

$$Av_n = \lambda_n v_n.$$

Die λ_n sind also Eigenwerte von A, insbesondere ist

$$\lambda_1 = \frac{1}{\|K\|} \geq \frac{1}{d^2},$$

und aus

$$v = Kf, \qquad (v, v_i) = \mu_i (f, v_i)$$

sowie $\mathcal{N}(K) = \{o\}$ folgt

$$\forall\, v \in \mathcal{D}(A) \qquad Av = f = \sum_{i=1}^{\infty} (f, v_i) v_i = \sum_{i=1}^{\infty} \lambda_i (v, v_i) v_i.$$

Damit läßt sich $\mathcal{D}(A)$ auch durch

$$\mathcal{D}(A) = \left\{ u \in \mathcal{X} \;\middle|\; \sum_{i=1}^{\infty} \lambda_i^2\, (u, v_i)^2 < \infty \right\}$$

charakterisieren.

Nach diesen Vorbereitungen können wir

$$\sqrt{A} \;:\; \mathcal{D}(\sqrt{A}) \subset \mathcal{X} \longrightarrow \mathcal{X}$$

definieren, und zwar durch

$$\mathcal{D}(\sqrt{A}) := \left\{ u \in \mathcal{X} \;\middle|\; \sum_{i=1}^{\infty} \lambda_i\, (u, v_i)^2 < \infty \right\}$$

$$\sqrt{A}\, u := \sum_{i=1}^{\infty} \sqrt{\lambda_i}\, (u, v_i)\, v_i.$$

Es läßt sich $\mathcal{D}(\sqrt{A}) = \mathring{\mathcal{H}}_1(\Omega)$ zeigen, und auch stetige Funktionen von A bzw. $\sqrt{A}$ lassen sich analog erklären. Damit erhalten wir formal die Lösung unserer Anfangsrandwertaufgabe durch

$$u(t) = \cos \sqrt{A}\, t \circ u^0 + \frac{\sin \sqrt{A}\, t}{\sqrt{A}} \circ u^1. \tag{$*$}$$

Diese Abbildung müssen wir natürlich interpretieren. Es gibt verschiedene Lösungsbegriffe, zum Beispiel *Lösungen im Distributionensinn, schwache Lösungen, starke Lösungen* oder *klassische Lösungen.*

Hier soll nur ein (physikalisch sinnvoller) Lösungsbegriff angegeben werden, nämlich der einer Lösung mit *endlicher Energie.* Die Energie $E(t)$ ist als Summe von *kinetischer* und *potentieller Energie* gegeben durch

$$E(t) := \Big(\big\| u' \big\|^2 + \big\| \sqrt{A}\, u \big\|^2 \Big)(t) < \infty .$$

Wegen

$$E(0) = \big\| u^1 \big\|^2 + \big\| \sqrt{A}\, u^0 \big\|^2$$

muß man dafür aber $u^0 \in \mathcal{D}(\sqrt{A}) = \mathring{\mathcal{H}}_1(\Omega)$ voraussetzen und wählt $u \in \mathcal{C}\big(\mathbb{R}_0^+, \mathcal{D}(A)\big) \cap \mathcal{C}_1(\mathbb{R}_0^+, \mathcal{X})$. Man rechnet dann nach, daß $E(t)$ konstant ist und daß durch Gl. $(*)$ die Lösung unserer Aufgabe in der Form

$$u(t) = \sum_{i=1}^{\infty} \left\{ \cos \sqrt{\lambda_i} t \, (u^0, v_i)\, v_i + \frac{\sin \sqrt{\lambda_i} t}{\sqrt{\lambda_i}} \, (u^1, v_i)\, v_i \right\} .$$

gegeben wird.

7.7 Klassifikation kompakter linearer Abbildungen

Wir beschließen dieses Kapitel mit einer Klassifikation kompakter Abbildungen, die für manche Anwendungen interessant ist.

Für kompakte normale Abbildungen haben wir eine handliche Spektraldarstellung angeben können. Leider fehlt eine solche Darstellung für allgemeinere kompakte Abbildungen. Um auch hier etwas Vergleichbares zu erhalten, betrachtet man mit $K \in \mathcal{KL}(\mathcal{X})$, wobei $\mathcal{X}$ ein Hilbertraum sei, die Abbildung

$$K^* K \in \mathcal{KL}(\mathcal{X}).$$

Offenbar ist $(K^*K)^* = K^*K$. Die Abbildung K^*K ist also selbstadjungiert und wegen

$$(K^* K x, x) = (Kx, Kx) \geq 0$$

nichtnegativ. Es seien

$$\kappa_1 = \|K^* K\| = \|K\|^2 \geq \kappa_2 \geq \kappa_3 \geq \cdots > 0$$

die positiven Eigenwerte von K^*K. Dann gibt es also ein Orthonormalsystem $\{u_i\}$ mit

$$\forall\, u \in \mathcal{X} \qquad K^* K u = \sum_{i \geq 1} \kappa_i \, (u, u_i)\, u_i .$$

Definition 7.7.1: *Die Zahlen*

$$\mu_i := \sqrt{\kappa_i} > 0$$

heißen singuläre Werte *von K.*

Damit können wir das erste Resultat dieses Abschnitts formulieren:

Satz 7.7.2: *Es gibt zwei orthonormale Folgen (u_n) und (v_n) mit*

$$K u_j = \mu_j\, v_j \quad \textit{und} \quad K^* v_j = \mu_j\, u_j$$

sowie

$$K = \sum_{j \geq 1} \mu_j \, (\,\cdot\,, u_j)\, v_j . \tag{$*$}$$

Gl. $(*)$ ist die gesuchte Darstellung von K.

Beweis: Es ist $\mu_j^2 = \kappa_j$ Eigenwert zu K^*K, die zugehörende Eigenfunktion ist u_j. Es sei P_0 der Projektor auf $\mathcal{N}(K^*K)$. Dann ist

$$id = P_0 + \sum_{j \geq 1} (\,\cdot\,, u_j)\, u_j .$$

Es seien ferner

$$v_j := \frac{1}{\mu_j} K u_j .$$

Dann ist

$$(v_j, v_k) = \frac{1}{\mu_j \mu_k}(Ku_j, Ku_k) = \frac{1}{\mu_j \mu_k}(u_j, K^*Ku_k) = \frac{\mu_k}{\mu_j}(u_j, u_k) = \delta_{jk}$$

und

$$K^*v_j = \frac{1}{\mu_j} K^*Ku_j = \mu_j u_j$$

sowie

$$KK^*v_j = \mu_j Ku_j = \kappa_j\, v_j.$$

Es sei nun $x \in \mathcal{X}$. Dann gibt es ein $x_0 \in \mathcal{N}(K^*K) = \mathcal{N}(K)$ mit

$$x = x_0 + \sum_{j \geq 1} (x, u_j)\, u_j.$$

Daraus folgt schließlich

$$Kx = \sum_{j \geq 1} \mu_j\, (x, u_j)\, v_j. \qquad \square$$

Diese Darstellung wollen wir im folgenden zur Charakterisierung spezieller K benutzen. Zunächst zeichnen wir die beiden Extremfälle aus:

$$\mathcal{K}_\infty(\mathcal{X}) := \mathcal{KL}(\mathcal{X}) \quad \text{und} \quad \mathcal{K}_0(\mathcal{X}) := \Big\{ K \in \mathcal{KL}(\mathcal{X}) \;\Big|\; \dim \mathcal{R}(K) < \infty \Big\}.$$

Für $0 < \alpha < \infty$ sei nun

$$\|K\|_\alpha := \Big(\sum_{j \geq 1} \mu_j^\alpha \Big)^{1/\alpha}$$

und

$$\mathcal{K}_\alpha(\mathcal{X}) := \Big\{ K \in \mathcal{KL}(\mathcal{X}) \;\Big|\; \|K\|_\alpha < \infty \Big\}.$$

Mit K gehört auch K^* zu $\mathcal{K}_\alpha(\mathcal{X})$, weil beide Abbildungen dieselben singulären Werte besitzen, und für $0 \leq \alpha \leq \beta \leq \infty$ kann man

$$\mathcal{K}_0 \subset \mathcal{K}_\alpha \subset \mathcal{K}_\beta \subset \mathcal{K}_\infty$$

zeigen. Auf Beweiseinzelheiten wollen wir hier verzichten.

Von $\mathcal{K}_0$ abgesehen, sind die beiden folgenden Klassen wichtig:

1. Die Klasse $\mathcal{K}_2$ der *Hilbert-Schmidt Operatoren*. Es sei zum Beispiel mit $k \in \mathcal{L}^2(G \times G)$ und $f \in \mathcal{L}^2(G)$

$$Kf := \int_G f(y)\, k(\,\cdot\,, y)\, dy.$$

Dann ist $K \in \mathcal{K}_2$, und für alle vollständigen Orthonormalsysteme (x_j) in $\mathcal{L}^2(G)$ gilt

$$\|K\|_2 = \sqrt{\iint |k(x,y)|^2 dx\, dy} = \sqrt{\sum_{i=1}^{\infty} \|Kx_i\|^2}.$$

Mit $K_1, K_2 \in \mathcal{K}_2$ ist $K_1 \circ K_2 \in \mathcal{K}_1$.

2. Die Klasse $\mathcal{K}_1$ der *nuklearen Operatoren* oder der *Spurklasse-Operatoren*. Es sei $K \in \mathcal{K}_1$. Dann gibt es $K_1, K_2 \in \mathcal{K}_2$ mit

$$K = K_1 \circ K_2.$$

Es seien (x_j) ein vollständiges Orthonormalsystem und $K \in \mathcal{KL}(\mathcal{X})$. Dann ist

$$(Kx_j, x_j) = \sum_{i \geq 1} \mu_i\, (x_j, u_i)\, (v_i, x_j) = \sum_{i \geq 1} (x_j, u_i)\, (Ku_i, x_j),$$

also

$$\sum_{j \geq 1} (Kx_j, x_j) = \sum_{i \geq 1} (Ku_i, u_i).$$

Deshalb ist

$$\operatorname{spur} K := \sum_{j \geq 1} (Kx_j, x_j)$$

von (x_j) unabhängig. Besitzt K ein vollständiges System von Eigenfunktionen (x_j) mit $Kx_j = \lambda_j x_j$, dann ist

$$\operatorname{spur} K = \sum_{j \geq 1} \lambda_j .$$

Für $K \in \mathcal{K}_2$ ist

$$\|K\|_2^2 = \|K^*\|_2^2 = \sum_{i \geq 1} \mu_i^2 = \operatorname{spur} K^* K,$$

und für $K \in \mathcal{K}_1$ kann

$$\sqrt{K^* K} := \sum_{j \geq 1} \mu_j (\cdot, u_j) u_j \in \mathcal{K}_1$$

definiert werden. Dann ist

$$\|K\|_1 = \sum_{j \geq 1} \mu_j = \operatorname{spur} \sqrt{K^* K}.$$

Wir beschließen dieses Kapitel mit dem Beispiel einer Abbildung $K \in \mathcal{KL}(\mathcal{X})$ mit

$$\forall\, \alpha \in [0, \infty) \quad K \notin \mathcal{K}_\alpha(\mathcal{X}).$$

Beispiel 7.7.3: *Es seien $(u_n)_{n \in \mathbb{N}}$ und $(v_n)_{n \in \mathbb{N}}$ zwei Orthonormalsysteme in $\mathcal{X}$, $\mu_n := 1/\ln(1+n)$ und*

$$K := \sum_{j=1}^{\infty} \mu_j (\,\cdot\,, u_j) v_j .$$

Dann ist $K \in \mathcal{K}_\infty$, aber $K \notin \mathcal{K}_\alpha$ für alle $\alpha \in \mathbb{R}_0^+$.

Beweis: Es gilt $\mu_n \to 0$, und es sei

$$K_n := \sum_{j=1}^{n} \mu_j (\,\cdot\,, u_j)\, v_j \in \mathcal{K}_\alpha .$$

Dann ist $m < n$

$$\|(K_n - K_m)x\|^2 = \Big\| \sum_{j=m+1}^{n} \mu_j \,(x, u_j)\, v_j \Big\|^2 = \sum_{j=m+1}^{n} \mu_j^2 \,(x, u_j)^2 \leq \mu_{m+1}^2 \|x\|^2,$$

also

$$\|K_n - K_m\| \to 0.$$

Das beweist $K \in \mathcal{K}_\infty$. Es ist aber für $\alpha \in \mathbb{R}_0^+$.

$$\sum_{j=1}^{\infty} \mu_j^\alpha \geq \int_1^\infty \Big(\frac{1}{\ln(1+x)} \Big)^\alpha dx = \int_2^\infty \frac{dt}{(\ln t)^\alpha} = \int_{\ln 2}^\infty \frac{e^x}{x^\alpha}\, dx = \infty . \qquad \square$$

8 Selbstadjungierte Fortsetzungen symmetrischer Abbildungen

Bei der Beschreibung vieler physikalischer Vorgänge erwartet man selbstadjungierte Operatoren. Das Spektrum soll reell sein. Nun ist es aber bei der ersten Formulierung oft noch nicht klar, wie der Definitionsbereich des zugrundeliegenden Operators gewählt werden muß, und man startet deshalb zunächst mit einem geeignet erscheinenden Definitionskern, etwa mit den Testfunktionen. Dann erhält man aber nur einen symmetrischen Operator S und im allgemeinen noch keinen selbstadjungierten. Es ist deshalb sehr wichtig, die Frage zu klären, ob sich S selbstadjungiert fortsetzen läßt, ob es gegebenenfalls mehrere selbstadjungierte Fortsetzungen gibt und welche davon in Frage kommt. Dadurch wird dann auch das physikalische Problem erst richtig formuliert.

Auch aus mathematischer Sicht ist das eine wichtige Frage, weil man für selbstadjungierte Abbildungen besonders schöne Darstellungen kennt (Spektralsätze). Das werden wir im nächsten Kapitel vertiefen.

Wir beginnen mit dem Beispiel eines *Schrödingeroperators für das Wasserstoffatom.*

Ein Schrödingeroperator: *Es seien* $\mathcal{X} = \mathcal{L}^2(\mathbb{R}^3)$, $q > 0$, $\mathcal{D}(A) = \mathring{\mathcal{C}}_\infty(\mathbb{R}^3)$ *und*

$$\begin{aligned} A \quad : \quad \mathcal{D}(A) \subset \mathcal{X} &\longrightarrow \mathcal{X}, \\ u &\longmapsto (-\Delta + V)\, u \end{aligned}$$

mit

$$(V\,u)(x) := -\frac{q}{|x|}\, u(x).$$

Aus der Zweiten Poincaréschen Abschätzung folgt unmittelbar

$$\forall\, \varphi \in \mathring{\mathcal{C}}_\infty(\mathbb{R}^3) \qquad \left\| \frac{\varphi(\cdot)}{|\cdot|} \right\| \leq 2\, \|\nabla \varphi\|$$

oder

$$\|V u\| \leq 2q\, |u|_1.$$

Deshalb ist

$$(Au, u) = |u|_1^2 + (Vu, u) \geq |u|_1^2 - 2q\, |u|_1\, \|u\| \geq -q^2\, \|u\|^2.$$

Die Abbildung A ist also halbbeschränkt. Nach Satz 5.3.14 ist sie daher symmetrisch und deshalb abschließbar (Satz 5.3.10).

Mit etwas Rechnung findet man auch

$$\exists\, p_1, p_2 > 0 \quad \forall\, \varphi \in \mathring{\mathcal{C}}_\infty(\mathbb{R}^3) \qquad p_1\, \|\varphi\|_2 \leq \|\varphi\|_A \leq p_2\, \|\varphi\|_2,$$

also

$$\mathcal{D}\big(\bar{A}\big) = \mathcal{H}_2(\mathbb{R}^3).$$

Damit haben wir den Abschluß $\bar{A}$ von A gefunden. Es sei $S := \bar{A}$ der *Schrödingeroperator.* Wir möchten gerne wissen, ob S selbstadjungiert ist.

8.1 Die Cayleytransformation

Nach Satz 5.3.10 ist eine symmetrische Abbildung A stets abschließbar. Der Abschluß $\bar{A}$, die *triviale Fortsetzung*, ist ebenfalls symmetrisch. Im folgenden wollen wir deshalb von der Situation

$$S \quad : \quad \mathcal{D}(S) \subset \mathcal{X} \longrightarrow \mathcal{X}$$

ausgehen. Dabei sollen $\mathcal{X}$ ein Hilbertraum und S eine symmetrische abgeschlossene Abbildung sein.

Wir wollen untersuchen, ob bzw. wie sich S zu einer selbstadjungierten Abbildung fortsetzen läßt. Dazu benutzen wir in diesem Abschnitt die *Cayleytransformation.* Die Cayleytransformation wurde von John von Neumann, 1903–1957, zur Diskussion symmetrischer Abbildungen herangezogen. Sie ist das Analogon zur *Möbiustransformation*

$$w = \frac{z - \mathrm{i}}{z + \mathrm{i}}$$

in der Funktionentheorie; die Variable $z \in \mathbb{C}$ wird durch S ersetzt. Man bildet damit das Spektrum auf der reellen Achse auf den Rand des Einheitskreises ab und erhält so eine beschränkte Abbildung, die sich leichter diskutieren läßt. Benannt wird diese Abbildung nach Arthur Cayley, 1821–1895, der sie in der Gruppentheorie verwandt hat.

Wir beginnen mit

$$\|(S-\lambda)x\|^2 = \|(S-\alpha)x\|^2 + \beta^2\|x\|^2. \qquad (*)$$

Diese Aussage folgt aus der Symmetrie von S mit $\lambda = \alpha + \mathrm{i}\,\beta$. Insbesondere ist also für $\beta \neq 0$

$$\|x\| \leq \frac{1}{|\mathrm{Im}\,\lambda|}\|(S-\lambda)x\|.$$

Folgerung 8.1.1: *Es sei* $\mathrm{Im}\,\lambda \neq 0$. *Dann gilt:*

1. $\mathcal{R}(S-\lambda) = \overline{\mathcal{R}(S-\lambda)}$.
2. *Die Abbildung* $(S-\lambda)^{-1} : \mathcal{R}(S-\lambda) \subset \mathcal{X} \to \mathcal{D}(S) \subset \mathcal{X}$ *existiert, ist beschränkt und surjektiv.*

Wir charakterisieren nun die adjungierte Abbildung S^*:

Satz 8.1.2: *Es gilt:*

1. $\mathcal{N}(S^*-\lambda) = \overline{\mathcal{R}(S-\bar{\lambda})}^{\perp}$
2. $\mathcal{D}(S^*) = \mathcal{D}(S) \oplus \mathcal{R}(S+\mathrm{i})^{\perp} \oplus \mathcal{R}(S-\mathrm{i})^{\perp}$.

Beweis: Es gilt

$$\Big(u \in \overline{\mathcal{R}(S-\bar{\lambda})}^{\perp}\Big) \Longleftrightarrow \Big(\forall\, x \in \mathcal{D}(S) \quad 0 = \big(u, (S-\bar{\lambda})x\big) = \big((S^*-\lambda)u, x\big)\Big) \Longleftrightarrow \Big((S^*-\lambda)\,u = o\Big).$$

Das beweist die erste Aussage.

Zum Nachweis der zweiten sei $x \in \mathcal{D}(S^*)$. Dann folgt aus

$$\mathcal{X} = \mathcal{R}(S+\mathrm{i}) \oplus \mathcal{N}(S^*-\mathrm{i}), \quad \text{orthogonal},$$

die Zerlegung

$$(S^*+\mathrm{i})x = (S+\mathrm{i})x_0 + x'$$

mit $x_0 \in \mathcal{D}(S)$ und $x' \in \mathcal{N}(S^*-\mathrm{i})$. Es sei

$$x_1 := \frac{x'}{2\mathrm{i}} \in \mathcal{N}(S^*-\mathrm{i}) = \mathcal{R}(S+\mathrm{i})^{\perp}.$$

Dann ist $(S^*+\mathrm{i})x_1 = x'$ und wegen $S \subset S^*$

$$(S^*+\mathrm{i})x = (S+\mathrm{i})x_0 + (S^*+\mathrm{i})x_1 = (S^*+\mathrm{i})(x_0+x_1)$$

oder

$$(S^*+\mathrm{i})(x-x_0-x_1) = o.$$

Nun gilt aber auch

$$\mathcal{X} = \mathcal{R}(S-\mathrm{i}) \oplus \mathcal{N}(S^*+\mathrm{i}), \quad \text{orthogonal},$$

also

$$x_2 := x - x_0 - x_1 \in \mathcal{N}(S^*+\mathrm{i}) = \mathcal{R}(S-\mathrm{i})^{\perp}.$$

Damit wird

$$x = x_0 + x_1 + x_2$$

in der geforderten Weise zerlegt. Diese Zerlegung ist eindeutig. Es sei nämlich

$$o = y_0 + y_1 + y_2$$

mit y_0, y_1, y_2 aus $\mathcal{D}(S)$, $\mathcal{R}(S+\mathrm{i})^{\perp}$, $\mathcal{R}(S-\mathrm{i})^{\perp}$. Dann folgt

$$o = (S^*+\mathrm{i})y_0 + (S^*+\mathrm{i})y_1 = \underbrace{(S+\mathrm{i})y_0}_{\in \mathcal{R}(S+\mathrm{i})} + \underbrace{2\mathrm{i}y_1}_{\in \mathcal{N}(S^*-\mathrm{i})},$$

also $y_1 = o$, $(S+\mathrm{i})y_0 = o$ und damit $y_0 = o$. Dann ist aber auch $y_2 = o$. □

Nach diesen Vorbereitungen können wir die Cayleytransformation einführen.

Definition 8.1.3: *Die Abbildung*

$$\begin{aligned} U \;:\; \mathcal{D}(U) = \mathcal{R}(S+\mathrm{i}) \subset \mathcal{X} &\longrightarrow \mathcal{R}(U) = \mathcal{R}(S-\mathrm{i}) \subset \mathcal{X}, \\ u &\longmapsto (S-\mathrm{i})(S+\mathrm{i})^{-1}u \end{aligned}$$

heißt Cayleytransformierte *von S.*

Offenbar ist die Cayleytransformierte wohldefiniert. Wir zeigen

Satz 8.1.4: *Es gilt:*

1. *U ist abgeschlossen, surjektiv und isometrisch.*
2. *$(id-U)^{-1} : \mathcal{D}(S) \to \mathcal{D}(U)$ existiert und ist surjektiv.*
3. *Es ist*

$$S = \mathrm{i}(id+U)(id-U)^{-1},$$

und insbesondere ist $\mathcal{R}(id-U) = \mathcal{D}(S)$ in $\mathcal{X}$ dicht.

Beweis:

1. $\mathcal{D}(U) = \mathcal{R}(S+\mathrm{i})$ und $\mathcal{R}(U) = \mathcal{R}(S-i)$ sind abgeschlossen, und U ist surjektiv.

   a) Wir zeigen, daß U abgeschlossen ist: Es sei (x_n) eine Folge mit $x_n \to x$ und $Ux_n \to u$. Dann ist $z_n := (S+\mathrm{i})^{-1}x_n$ eine Cauchyfolge, etwa mit $z_n \to z$. Weil S abgeschlossen ist, folgt dann aus $x_n = (S+\mathrm{i})z_n$

   $$x = (S+\mathrm{i})z \in \mathcal{D}(U).$$

   Nun ist

   $$(S-\mathrm{i})z_n = Ux_n \to u,$$

   also $(S-\mathrm{i})z = u$. Das heißt, es ist

   $$u = (S-\mathrm{i})(S+\mathrm{i})^{-1}x = Ux.$$

   b) Wir zeigen, daß U isometrisch ist: Weil $\mathcal{D}(U)$ abgeschlossen ist, folgt aus dem Satz vom abgeschlossenen Graphen die Beschränktheit von U, also $U \in \mathcal{CL}\big(\mathcal{R}(S+\mathrm{i}),\, \mathcal{R}(S-\mathrm{i})\big)$. Nun ist für $u \in \mathcal{R}(S+\mathrm{i})$ nach Gl. $(*)$

   $$\|Uu\|^2 = \big\|(S-\mathrm{i})(S+\mathrm{i})^{-1}u\big\|^2 = \big\|(S+\mathrm{i})(S+\mathrm{i})^{-1}u\big\|^2 = \|u\|^2.$$

2. Für $u \in \mathcal{D}(U)$ ist

   $$(id-U)u = \big((S+\mathrm{i}) - (S-\mathrm{i})\big)(S+\mathrm{i})^{-1}u = 2\mathrm{i}\,(S+\mathrm{i})^{-1}u.$$

   Nun ist

   $$(S+\mathrm{i})^{-1} \;:\; \mathcal{R}(S+\mathrm{i}) = \mathcal{D}(U) \longrightarrow \mathcal{D}(S)$$

   surjektiv. Mithin existiert

   $$\begin{aligned} (id-U)^{-1} \;:\; \mathcal{D}(S) &\longrightarrow \mathcal{R}(S+\mathrm{i}) = \mathcal{D}(U), \\ x &\longmapsto \tfrac{1}{2\mathrm{i}}\,(S+\mathrm{i})x \end{aligned}$$

   und ist surjektiv.

3. Analog erhalten wir für $u \in \mathcal{D}(U)$

   $$(id+U)u = \big((S+\mathrm{i}) + (S-\mathrm{i})\big)(S+\mathrm{i})^{-1}u = 2S(S+\mathrm{i})^{-1}u.$$

   Für alle $x \in \mathcal{D}(S)$ gilt also

   $$(id+U)(id-U)^{-1}x = \tfrac{1}{2\mathrm{i}}\,(id+U)(S+\mathrm{i})x = \tfrac{1}{\mathrm{i}}\,Sx. \qquad \square$$

Als nächstes zeigen wir einen Umkehrsatz, nämlich

Satz 8.1.5: *Es sei*

$$V \;:\; \mathcal{D}(V) \subset \mathcal{X} \longrightarrow \mathcal{X}$$

eine abgeschlossene isometrische Abbildung mit $\overline{\mathcal{R}(id-V)} = \mathcal{X}$. *Dann gibt es genau eine abgeschlossene symmetrische Abbildung*

$$H \;:\; \mathcal{D}(H) \subset \mathcal{X} \longrightarrow \mathcal{X}$$

mit $\mathcal{D}(H) = \mathcal{R}(id-V)$, *deren* Cayley-Transformierte V *ist.*

Wir führen den Beweis in fünf Schritten:

1. Wir zeigen, daß

$$(id-V)^{-1} \;:\; \mathcal{R}(id-V) \longrightarrow \mathcal{D}(V)$$

existiert. Es sei nämlich $(id-V)x = o$. Dann folgt für alle $y \in \mathcal{R}(id-V)$ mit $y = (id-V)z$

$$(x,y) = (x,z) - (x,Vz) = (Vx - x, Vz) = 0,$$

also $x = o$.

2. Es sei

$$\begin{aligned} H \;:\; \mathcal{R}(id-V) &\longrightarrow \mathcal{R}(id+V), \\ x &\longmapsto \mathrm{i}\,(id+V)(id-V)^{-1}x. \end{aligned}$$

Dann ist $\mathcal{D}(H) = \mathcal{R}(id-V)$ in $\mathcal{X}$ dicht, und für alle $x, y \in \mathcal{D}(H)$ folgt mit $x = (id-V)u$, $y = (id-V)v$

$$(Hx,y) = \mathrm{i}\big((id+V)u, (id-V)v\big) = \mathrm{i}\big((Vu,v) - (u,Vv)\big) = \big((id-V)u, \mathrm{i}(id+V)v\big) = (x,Hy).$$

H ist also symmetrisch.

3. Wir zeigen, daß H abgeschlossen ist. Es seien $x_n \in \mathcal{D}(H)$, $x_n \to x$ und $Hx_n \to y$ sowie $x_n = (id-V)u_n$. Dann gilt

$$\begin{aligned} x_n &= (id-V)u_n, \\ -\mathrm{i}\,Hx_n &= (id+V)u_n, \end{aligned}$$

also

$$\begin{aligned} 2u_n &= x_n - \mathrm{i}\,Hx_n \to x - \mathrm{i}\,y, \\ 2Vu_n &= -x_n - \mathrm{i}\,Hx_n \to -x - \mathrm{i}\,y. \end{aligned}$$

Weil V abgeschlossen ist, folgt mit $u_n \to u$

$$Vu_n \to Vu$$

oder

$$x - \mathrm{i}\,y = 2u \quad \text{und} \quad x + \mathrm{i}\,y = -2Vu,$$

also

$$x = (id-V)u \quad \text{und} \quad -\mathrm{i}\,y = (id+V)u.$$

Somit ist $y = Hx$.

4. Wir zeigen, daß V die Cayley-Transformierte von H ist. Es sei $x \in \mathcal{D}(H)$. Dann existiert ein $u \in \mathcal{D}(V)$ mit

$$x = (id-V)u \quad \text{und} \quad Hx = \mathrm{i}(id+V)u,$$

und zu jedem $u \in \mathcal{D}(V)$ gibt es ein solches x. Daraus folgen

$$(H+\mathrm{i})x = 2\mathrm{i}\,u \quad \text{und} \quad (H-\mathrm{i})x = 2\mathrm{i}\,Vu,$$

also $\mathcal{D}(V) = \mathcal{R}(H+\mathrm{i})$, $\mathcal{R}(V) = \mathcal{R}(H-\mathrm{i})$ und

$$V = (H-\mathrm{i})(H+\mathrm{i})^{-1}.$$

5. Als letztes zeigen wir die Eindeutigkeit. Es sei J eine abgeschlossene symmetrische Abbildung, deren Cayley-Transformierte V ist. Dann ist nach Satz 8.1.4

$$J = \mathrm{i}(id+V)(id-V)^{-1}$$

und damit $J = H$. □

Folgerung 8.1.6: *Es sind äquivalent:*

(1) $S = S^*$.

(2) U *ist unitär (das heißt,* U *ist isometrisch mit* $\mathcal{D}(U) = \mathcal{R}(U) = \mathcal{X}$*).*

(3) $\mathcal{R}(S \pm \mathrm{i}) = \mathcal{X}$.

(4) $\mathcal{N}(S^* \pm \mathrm{i}) = \{o\}$.

Beweis: (1) $\Leftrightarrow$ (3) und (1) $\Leftrightarrow$ (4) folgen aus Satz 8.1.2. Die Äquivalenz (2) $\Leftrightarrow$ (3) folgt aus

$$\mathcal{D}(U) = \mathcal{R}(S + \mathrm{i}) \quad \text{und} \quad \mathcal{R}(U) = \mathcal{R}(S - \mathrm{i}). \qquad \square$$

Nach diesen Vorbereitungen wollen wir die Frage nach der selbstadjungierten Fortsetzbarkeit von S behandeln. Wir beginnen mit

Definition 8.1.7: *Die Zahlen* $m, n \in \mathbb{N}_0 \cup \{\infty\}$

$$m := \dim \mathcal{R}(S + \mathrm{i})^\perp \quad \textit{und} \quad n := \dim \mathcal{R}(S - \mathrm{i})^\perp$$

heißen Defektindizes *von* S.

Es gilt

$$m = \dim \mathcal{D}(U)^\perp = \dim \mathcal{N}(S^* - \mathrm{i}) = \operatorname{codim} \mathcal{R}(S + \mathrm{i}),$$
$$n = \dim \mathcal{R}(U)^\perp = \dim \mathcal{N}(S^* + \mathrm{i}) = \operatorname{codim} \mathcal{R}(S - \mathrm{i})$$

und

Folgerung 8.1.8: *Es gilt*

$$S = S^* \iff m = n = 0.$$

S heißt *maximal symmetrisch*, wenn S keine echte symmetrische Fortsetzung besitzt.

Es ist nun einfacher, statt S die Cayley-Transformierte U von S fortzusetzen. Wir zeigen

Satz 8.1.9: *Es seien* $p > 0$, $m = m' + p$ *und* $n = n' + p$ *mit* $m', n' \in \mathbb{N}$. *Dann gibt es eine abgeschlossene symmetrische Fortsetzung* S' *von* S *mit den Defektindizes* (m', n').

Folgerung 8.1.10: *Es gilt*

$$\Big(S \textit{ ist maximal symmetrisch}\Big) \iff \Big(m = 0 \textit{ oder } n = 0\Big).$$

Beweis des Satzes: Es seien

$$\{\varphi_1, \ldots, \varphi_p, \varphi_{p+1}, \ldots, \varphi_{p+m'}\}, \quad \varphi_\mathrm{i} \in \mathcal{D}(U)^\perp,$$
$$\{\psi_1, \ldots, \psi_p, \psi_{p+1}, \ldots, \psi_{p+n'}\}, \quad \psi_\mathrm{i} \in \mathcal{R}(U)^\perp$$

Orthonormalsysteme. Wir definieren eine isometrische Fortsetzung V von U durch

$$Vx := \begin{cases} Ux & \text{für } x \in \mathcal{D}(U) \\ \displaystyle\sum_{\mathrm{i}=1}^{p} \xi_\mathrm{i} \psi_\mathrm{i} & \text{für } x = \displaystyle\sum_{\mathrm{i}=1}^{p} \xi_\mathrm{i} \varphi_\mathrm{i},\ \xi_i \in \mathbb{R}. \end{cases}$$

Dann ist $\|Vx\| = \|x\|$ mit

$$\mathcal{D}(V) = \mathcal{D}(U) \oplus \operatorname{span}(\varphi_1, \ldots, \varphi_p),$$
$$\mathcal{R}(V) = \mathcal{R}(U) \oplus \operatorname{span}(\psi_1, \ldots, \psi_p).$$

Es sei

$$S' \ : \ \mathcal{R}(id - V) \subset \mathcal{X} \longrightarrow \mathcal{R}(id + V) \subset \mathcal{X},$$
$$x \longmapsto \mathrm{i}(id + V)(id - V)^{-1} x.$$

Dann ist S' symmetrisch, $S' \supset S$ und hat die Defektindizes (m', n'). $\square$

Setzt man S also symmetrisch fort, dann nehmen die Defektindizes um die gleiche Zahl ab. Die Fortsetzung S' ist maximal, wenn einer der Defektindizes verschwindet, S' ist selbstadjungiert, wenn beide verschwinden.

Wir konstruieren schließlich noch eine maximal symmetrische Abbildung mit den Defektindizes $(0,1)$. Es seien $\mathcal{X}$ ein separabler Hilbertraum, $\{\varphi_j\}$ ein vollständiges Orthonormalsystem in $\mathcal{X}$ und U durch

$$\begin{aligned} U \;&:\; \mathcal{X} \longrightarrow \mathcal{X}, \\ &\quad \varphi_j \longmapsto \varphi_{j+1} \end{aligned}$$

definiert. Dann ist $\mathcal{D}(U) = \mathcal{X}$, aber

$$\mathcal{R}(U)^{\perp} = \operatorname{span}(\varphi_1).$$

Die Abbildung U ist abgeschlossen und isometrisch mit $\overline{\mathcal{R}(id-U)} = \mathcal{X}$. Letzteres folgt aus

$$\begin{aligned} x \in \mathcal{R}(id-U)^{\perp} &\Longleftrightarrow \forall\, y \in \mathcal{X} \quad 0 = (x, (id-U)y) \\ &\Longrightarrow 0 = \|x\|^2 - (x, Ux) \Longrightarrow \|(id-U)x\|^2 = 0, \end{aligned}$$

also mit $x = \Sigma_{i=1}^{\infty} \xi_i \, \varphi_i$

$$o = (id-U)x = \xi_1 \, \varphi_1 + (\xi_2 - \xi_1)\varphi_2 + \cdots$$

oder $x = o$. Mithin existiert

$$S := \mathrm{i}(id+U)(id-U)^{-1},$$

und es ist $(m,n) = (0,1)$. Die Abbildung S kann also nicht selbstadjungiert fortgesetzt werden.

Wir wollen unsere Ergebnisse noch einmal in einer etwas anderen Form zusammenfassen, nämlich im

Satz 8.1.11: *Es sei S eine abgeschlossene symmetrische Abbildung. Dann gilt:*

1. *Sowohl in der oberen komplexen Halbebene als auch in der unteren ist*

$$\dim \mathcal{R}(S-\bar{\lambda})^{\perp} = \dim \mathcal{N}(S^* - \lambda)$$

   jeweils konstant.

2. *Das Spektrum $\sigma(S)$ besteht*
   (1) entweder aus der abgeschlossenen oberen Halbebene
   (2) oder aus der abgeschlossenen unteren Halbebene
   (3) oder aus der ganzen komplexen Ebene
   (4) oder aus einer abgeschlossenen Teilmenge der reellen Achse.

3. *S ist selbstadjungiert $\Longleftrightarrow$ Aussage 2(4) gilt.*

4. *S ist selbstadjungiert $\Longleftrightarrow$ sowohl in der oberen Halbebene als auch in der unteren ist*

$$\dim \mathcal{R}(S-\bar{\lambda})^{\perp} = \dim \mathcal{N}(S^* - \lambda) = 0.$$

Zum Beweis bemerken wir folgendes:

1. Es seien M, N abgeschlossene Teilräume von $\mathcal{X}$ mit

$$\dim M < \dim N.$$

   Dann ist $N \cap M^{\perp} \neq \{o\}$.

2. Es sei $u \in \mathcal{N}\big(S^* - (\lambda+\mu)\big) \cap \mathcal{N}(S^* - \lambda)^{\perp}$ mit $\|u\| = 1$ und $\mu \in \mathbb{C}$. Dann ist $u \in \mathcal{R}(S-\bar{\lambda})$, und es gibt ein $x \in \mathcal{D}(S)$ mit $u = (S-\bar{\lambda})x$. Daraus folgt

$$0 = ((S^* - (\lambda+\mu))u, x) = \|u\|^2 - \mu(u,x),$$

   also

$$\|u\| \leq |\mu| \, \|x\|.$$

   Andererseits aber für $0 < |\mu| < |\operatorname{Im}\lambda|$ nach Gl. $(*)$

$$\|u\| \geq |\beta| \, \|x\| > |\mu| \, \|x\|.$$

Für kleine $|\mu|$ gibt es also kein solches u.

3. Für $|\mu| < |\operatorname{Im}\lambda|$ folgt also

$$\dim\mathcal{N}\big(S^* - (\lambda+\mu)\big) \le \dim\mathcal{N}(S^* - \lambda).$$

Die umgekehrte Abschätzung gilt für kleine $|\mu|$ entsprechend. Für $\operatorname{Im}\lambda \neq 0$ und kleine $|\mu|$ ist also

$$\dim\mathcal{N}\big(S^* - (\lambda+\mu)\big) = \dim\mathcal{N}(S^* - \lambda).$$

Das beweist die erste Aussage des Satzes.

4. Die übrigen Aussagen sind in den bereits bewiesenen Resultaten enthalten. □

Wir beschließen diesen Abschnitt mit einem Beispiel aus der Theorie der gewöhnlichen Differentialgleichungen.

Beispiel 8.1.12: *Es seien $G \subset \mathbb{R}^1$ ein Gebiet, $\mathcal{X} = \mathcal{L}^2(G)$, $\mathcal{D}(A) = \mathring{\mathcal{H}}_2(G) \subset \mathcal{X}$ und $A := -\bar{\Delta}$.*

Dabei sei Δ der in §5.4 eingeführte Δ-Operator. Es war

$$\forall\,\varphi \in \mathring{\mathcal{C}}_\infty(G) \qquad \sqrt{\tfrac{2}{3}}\,\|\varphi\|_2 \;\le\; \|\varphi\|_\Delta \;\le\; \|\varphi\|_2.$$

Die Abbildung A ist symmetrisch, und es gilt

$$\mathcal{D}(A^*) = \left\{u \in \mathcal{X} \;\middle|\; \Delta u \in \mathcal{X}\right\}.$$

Wegen $G \subset \mathbb{R}^1$ ist nach Folgerung 6.3.1 $\mathcal{D}(A^*) = \mathcal{H}_2(G)$. Aus

$$(A^* - \mathrm{i})u = o \iff (A^* + \mathrm{i})\bar{u} = o$$

folgt dann die Gleichheit der Defektindizes von A. Die Abbildung A ist also selbstadjungiert fortsetzbar. Insbesondere besitzt

$$(-\Delta - \mathrm{i})u = o$$

zwei linear unabhängige Lösungen, nämlich

$$u_1(x) = \mathrm{e}^{\mathrm{i}\sqrt{\mathrm{i}}x} = \mathrm{e}^{\alpha(-1+\mathrm{i})x} \quad \text{und} \quad u_2(x) = \mathrm{e}^{-\mathrm{i}\sqrt{\mathrm{i}}x} = \mathrm{e}^{\alpha(1-\mathrm{i})x}$$

mit $\alpha = \sqrt{2}/2$. Es sind nun drei typische Fälle möglich:

1. Es sei $G = \mathbb{R}^1$. In diesem Falle sind $u_1, u_2 \notin \mathcal{X}$. Es ist also $\mathcal{N}(A^* \pm \mathrm{i}) = \{o\}$, und beide Defektindizes verschwinden. Mithin ist A selbstadjungiert. Das folgt auch aus $\mathring{\mathcal{H}}_2(\mathbb{R}^1) = \mathcal{H}_2(\mathbb{R}^1)$.
2. Es sei $G = (0,\infty)$. In diesem Falle ist nur $u_1 \in \mathcal{X}$, also $m = n = 1$. Wir wollen diesen Fall etwas genauer diskutieren. Es ist

$$\mathcal{D}(U)^\perp = \mathcal{N}(A^* - \mathrm{i}) = \operatorname{span}(u_1)$$
$$\mathcal{R}(U)^\perp = \mathcal{N}(A^* + \mathrm{i}) = \operatorname{span}(\bar{u}_1).$$

Es sei U_c eine isometrische Fortsetzung von U, definiert durch

$$U_c x := \begin{cases} Ux & \text{für } x \in \mathcal{D}(U) \\ c\bar{x} & \text{für } x = u_1 \end{cases}$$

mit $c \in \mathbb{C}$, $|c| = 1$. Es ist also

$$\mathcal{D}(U_c) = \mathcal{D}(U) \oplus \operatorname{span}(u_1) = \mathcal{X}.$$

Die Abbildung U_c ist unitär, und mit $c \in \mathbb{C}$, $|c| = 1$, erhält man alle unitären Fortsetzungen von U. Damit sind

$$A_c := \mathrm{i}(id + U_c)(id - U_c)^{-1}$$

die selbstadjungierten Fortsetzungen von A. Es ist

$$\mathcal{D}(A_c) = \mathcal{R}(id - U_c) = \mathcal{R}(id - U) \cup \operatorname{span}\big((id - U_c)u_1\big) = \mathcal{D}(A) \cup \operatorname{span}(u_c)$$

mit $u_c := u_1 - c\bar{u}_1$.

Um zu verstehen, welche Randbedingung die Aussage $u \in \mathcal{D}(A_c)$ enthält, entwickeln wir $u_c(x)$ in einer Umgebung von Null, $U(0)$, in eine Potenzreihe. Dort ist

$$\begin{aligned} u_c(x) &= u_1(x) - c\bar{u}_1(x) \\ &= \left[1 + \alpha(-1+\mathrm{i})x + \frac{\alpha^2}{2!}(-1+\mathrm{i})^2 x^2 + \cdots\right] - c\left[1 + \alpha(-1-\mathrm{i})x + \frac{\alpha^2}{2!}(-1-\mathrm{i})^2 x^2 + \cdots\right] \\ &= (1-c) + x\,\alpha(1+\mathrm{i})(c+\mathrm{i}) - x^2\,\alpha^2\mathrm{i}(1+c) + \cdots . \end{aligned}$$

Es seien $\varphi \in \mathring{\mathcal{C}}_\infty(\mathbb{R}^1)$ mit $\varphi = 1$ in $U(0)$ und

$$v_c(x) := \left\{(1-c) + x\,\alpha(1+\mathrm{i})(c+\mathrm{i})\right\}\varphi(x).$$

Dann ist

$$u_c(x) =: v_c(x) + w_c(x)$$

mit $u_c, v_c, w_c \in \mathcal{X}$ und $w_c \in \mathring{\mathcal{H}}_2(G) = \mathcal{D}(A)$. Deshalb erhalten wir

$$\begin{aligned} u_c(0) &= v_c(0) = 1 - c \\ u_c'(0) &= v_c'(0) = \alpha(1+\mathrm{i})(c+\mathrm{i}). \end{aligned}$$

Wir unterscheiden nun zwei Fälle:

a) Es sei $c = 1$. Das ist die Dirichletsche Randwertaufgabe. Es ist

$$\mathcal{D}(A_1) = \mathring{\mathcal{H}}_2(G) \cup \operatorname{span}(v_1)$$

mit

$$v_1(x) = \alpha(1+\mathrm{i})^2 x\varphi(x) \quad \text{und} \quad v_1'(0) = \alpha(1+\mathrm{i})^2 \neq 0.$$

Daraus folgt

$$\mathcal{D}(A_1) = \mathring{\mathcal{H}}_1(G) \cap \left\{u \in \mathcal{H} \;\middle|\; \Delta u \in \mathcal{X}\right\} = \mathring{\mathcal{H}}_1(G) \cap \mathcal{H}_2(G).$$

b) Es sei $c \neq 1$. Hierin sind die Neumannsche und die dritte Randwertaufgabe enthalten. Jetzt ist für $c \neq -\mathrm{i}$

$$v_c(0) = 1 - c = \frac{1-c}{\alpha(1+\mathrm{i})(c+\mathrm{i})} v_c'(0).$$

u_c genügt also der Randbedingung

$$u_c'(0) + z_c u_c(0) = 0$$

mit

$$z_c := -\alpha\frac{(1+\mathrm{i})(c+\mathrm{i})}{1-c} = \alpha\frac{1 - \cos\varphi + \sin\varphi}{1-\cos\varphi} =: f(\varphi)$$

für $c =: \mathrm{e}^{\mathrm{i}\varphi}$. Es ist

$$f \;:\; (0, 2\pi) \longrightarrow (-\infty, \infty)$$

eine bijektive, monoton fallende Abbildung.

Für $c = -\mathrm{i}$ erhält man die Neumannsche Randbedingung.

3. Es sei $G = (0,1)$. In diesem Falle sind beide Defektindizes gleich Zwei, und an beiden Intervallenden sind Randwertaufgaben möglich. Wir verzichten auf Einzelheiten.

8.2 Halbbeschränkte Abbildungen

Es sei S nun eine halbbeschränkte, symmetrische Abbildung. Indem wir von S zu $S + c$ oder $-S + c$ übergehen, können wir o.B.d.A. von

$$\forall\, x \in \mathcal{D}(S) \qquad (Sx, x) \geq \|x\|^2$$

ausgehen. Die halbbeschränkten bilden eine wichtige Teilklasse der symmetrischen Abbildungen. Der Schrödingeroperator gehört zu ihnen.

Für halbbeschränkte Abbildungen gibt es eine besonders einfache selbstadjungierte Fortsetzung, die *Friedrichssche Fortsetzung*, benannt nach Kurt Otto Friedrichs, 1901–1982. Es gilt

Satz 8.2.1: *Es sei S symmetrisch und halbbeschränkt. Dann gibt es eine selbstadjungierte halbbeschränkte Fortsetzung A von S mit derselben Schranke.*

Beweis: O.B.d.A. sei also $(Sx,x) \geq \|x\|^2$. Wir definieren auf $\mathcal{D}(S)$ das Skalarprodukt

$$(u,v)_{\mathcal{F}} := (Su,v)$$

mit

$$\|u\|_{\mathcal{F}}^2 := (u,u)_{\mathcal{F}} = (Su,u) \geq \|u\|^2.$$

Dann gilt

$$\forall\, u \in \mathcal{D}(S) \qquad \|u\| \leq \|u\|_{\mathcal{F}} \leq \|u\|_S.$$

Es sei nun

$$\mathcal{F} := \Big\{ u \in \mathcal{D}(S);\ \|\cdot\|_{\mathcal{F}} \Big\}^{\sim},$$

ausgestattet mit dem Skalarprodukt $(u,v)_{\mathcal{F}}$. Dann gilt

$$\mathcal{D}(S) \subset \mathcal{D}(\bar{S}) \subset \mathcal{F} \subset \mathcal{X}.$$

Wir wollen nun S auf eine Teilmenge von $\mathcal{F}$ fortsetzen, und zwar so, daß der Wertebereich dieser Fortsetzung ganz $\mathcal{X}$ wird. Eine solche Fortsetzung läßt sich mit Hilfe des Rieszschen Darstellungssatzes leicht angeben. Es sei $x \in \mathcal{X}$. Dann lösen wir in $\mathcal{F}$ die Gleichung

$$\forall\, v \in \mathcal{F} \qquad (u,v)_{\mathcal{F}} \stackrel{!}{=} (x,v). \tag{$*$}$$

Das ist eindeutig möglich, weil (x,v) wegen

$$|(x,v)| \leq \|x\| \cdot \|v\| \leq \|x\| \cdot \|v\|_{\mathcal{F}}$$

ein beschränktes lineares Funktional auf $\mathcal{F}$ ist. Es existiert also die injektive Abbildung

$$\begin{aligned} R \ :\ & \mathcal{X} \longrightarrow \mathcal{F}, \\ & x \longmapsto u \end{aligned}$$

mit

$$\|u\|_{\mathcal{F}} = \|Rx\|_{\mathcal{F}} \leq \|x\|.$$

Dabei ist u die Lösung von Gl. $(*)$. Nun gilt für alle $x, y \in \mathcal{X}$

$$(x,Ry) = (Rx,Ry)_{\mathcal{F}} = \overline{(Ry,Rx)}_{\mathcal{F}} = \overline{(y,Rx)} = (Rx,y).$$

Mithin ist $R : \mathcal{X} \to \mathcal{X}$ eine beschränkte symmetrische Abbildung, folglich selbstadjungiert. Damit ist auch R^{-1} selbstadjungiert. Es sei

$$\begin{aligned} F \ :\ & \mathcal{D}(F) = \mathcal{R}(R) \subset \mathcal{F} \subset \mathcal{X} \longrightarrow \mathcal{X}, \\ & u \longmapsto R^{-1}u \end{aligned}$$

diese selbstadjungierte Abbildung. Für ein $s \in \mathcal{D}(S)$ ist das Funktional (Ss,v) über $\mathcal{X}$ beschränkt, also gibt es ein $x \in \mathcal{X}$ mit

$$\forall\, v \in \mathcal{X} \qquad (Ss,v) = (x,v).$$

Dann gilt erst recht

$$\forall\, v \in \mathcal{F} \qquad (s,v)_{\mathcal{F}} = (Ss,v) = (x,v),$$

das heißt, es ist

$$\mathcal{D}(S) \subset \mathcal{D}(F) \subset \mathcal{F} \subset \mathcal{X}$$

und $Rx = s$, also $Ss = x = Fs$.

Damit ist

$$S \subset F,$$

und F ist die gesuchte selbstadjungierte *Friedrichssche Fortsetzung* von S. Es gilt

$$\mathcal{D}(F) = \Big\{ u \in \mathcal{F} \ \Big|\ \exists\, x \in \mathcal{X} \quad \forall\, v \in \mathcal{D}(S) \quad (u,Sv) = (x,v) \Big\},$$

und für alle $u \in \mathcal{D}(F)$ ist mit $x := R^{-1}u$

$$(Fu,u) = (x,u) = (Rx,u)_{\mathcal{F}} = \|u\|_{\mathcal{F}}^2 \geq \|u\|^2. \qquad \square$$

Es seien zum Beispiel $G \subset \mathbb{R}^n$ ein Gebiet, $\mathcal{D}(S) = \mathring{\mathcal{C}}_\infty(G)$, $\mathcal{X} = \mathcal{L}^2(G)$ und

$$Su := (-\Delta + 1)u.$$

Dann ist

$$(Su,u) = \|u\|_1^2,$$

also $\mathcal{F} = \mathring{\mathcal{H}}_1(G)$, und die Friedrichssche Fortsetzung von S erfüllt gerade die Dirichletsche Randbedingung.

Auch der eingangs eingeführte Schrödingeroperator ist halbbeschränkt, es war ja

$$\forall\, \varphi \in \mathring{\mathcal{C}}_\infty(\mathbb{R}^3) \qquad (A\varphi,\varphi) \geq -q^2\|\varphi\|^2.$$

Damit besitzt A eine selbstadjungierte Friedrichssche Fortsetzung B, denn man nehme statt A zunächst $\tilde{A} := A + q^2 + 1$. Dann ist $(\tilde{A}\varphi,\varphi) \geq \|\varphi\|^2$. Es gilt auch

$$\exists\, p_1,p_2 > 0 \quad \forall\, \varphi \in \mathring{\mathcal{C}}_\infty(\mathbb{R}^3) \qquad p_1\|\varphi\|_2 \leq \|\varphi\|_A \leq p_2\|\varphi\|_2,$$

also

$$\mathcal{D}(A) = \mathcal{H}_2(\mathbb{R}^3) \subset \mathcal{X} = \mathcal{L}^2(\mathbb{R}^3).$$

Wählt man $\mathcal{F}$ bezüglich $\tilde{A}$, dann folgt ebenfalls

$$\exists\, p_1,p_2 > 0 \quad \forall\, \varphi \in \mathring{\mathcal{C}}_\infty(\mathbb{R}^3) \qquad p_1\|\varphi\|_1 \leq \|\varphi\|_{\mathcal{F}} \leq p_2\|\varphi\|_1.$$

Damit erhalten wir

$$\mathcal{D}(B) = \left\{ u \in \mathring{\mathcal{H}}_1(\mathbb{R}^3) \;\middle|\; \Delta u \in \mathcal{X} \right\} = \mathcal{H}_2(\mathbb{R}^3).$$

Der Abschluß $S = \bar{A}$ ist also bereits selbstadjungiert.

Im allgemeinen ist es natürlich möglich, daß eine symmetrische Abbildung S noch andere selbstadjungierte Fortsetzungen besitzt. Aber unter diesen Fortsetzungen gibt es nur eine einzige, nämlich die soeben konstruierte Friedrichssche Fortsetzung F, deren Definitionsbereich in $\mathcal{F}$ enthalten ist. Es gilt nämlich

Satz 8.2.2: *Es sei S' eine beliebige symmetrische Fortsetzung von S mit $\mathcal{D}(S') \subset \mathcal{F}$. Dann ist $S' \subset F$.*

Beweis: Es seien $s \in \mathcal{D}(S)$ und $s' \in \mathcal{D}(S')$. Dann ist

$$(RS's',s)_{\mathcal{F}} = (S's',s) = (s',S's) = (s',Ss) = (s',s)_{\mathcal{F}},$$

also

$$(RS's',s)_{\mathcal{F}} = (s',s)_{\mathcal{F}},$$

oder

$$RS's' = s'.$$

Damit ist $s' \in \mathcal{D}(F)$ und

$$Fs' = FRS's' = S's'. \qquad \square$$

In diesem Sinne spricht man auch von der *kleinsten selbstadjungierten Fortsetzung* von S.

9 Der Spektralsatz für selbstadjungierte Abbildungen

In diesem Kapitel seien $\mathcal{H}$ ein Hilbertraum und

$$A \;:\; \mathcal{D}(A) \subset \mathcal{H} \longrightarrow \mathcal{H}$$

eine selbstadjungierte Abbildung. Wir wollen für A eine *Spektraldarstellung* angeben. Solche Darstellungssätze gehen schon auf Hilbert, ab 1906, zurück. Der nunmehr klassische *Spektralsatz für selbstadjungierte Abbildungen* stammt von John von Neumann, 1903–1957, und wurde [1929] veröffentlicht. Wenig später hat auch Marshall Stone, 1903–1989, [1932] einen Beweis gegeben. Er hat auch die nach ihm benannte Umkehrformel angegeben. Inzwischen sind viele Beweise veröffentlicht worden. Es gibt völlig elementare, etwa der von *Leinfelder [1979]*, und auch recht abstrakte. Wir wollen hier einen Mittelweg wählen und einen Beweis geben, der etwas Funktionentheorie und etwas Maßtheorie verwendet.

9.1 Spektralscharen

Wir beginnen mit folgender

Definition 9.1.1: *Eine Familie von orthogonalen Projektoren*

$$\begin{aligned} F \;:\; \mathbb{R} &\longrightarrow \mathcal{CL}(\mathcal{H},\mathcal{H}),\\ \lambda &\longmapsto P(\lambda) \end{aligned}$$

heißt Spektralschar, *wenn folgendes gilt:*

1. $P(\lambda)\,P(\mu) = P\big(\min(\lambda,\mu)\big)$.
2. *Die Schar ist stark stetig von rechts, es gilt also* $P(\lambda+0) = P(\lambda)$.
3. *Die starken Limites* $P(-\infty) = O$ *und* $P(\infty) = id$ *existieren.*

Dabei dient die zweite Forderung nur zur Normierung, die Existenz der starken Limites $P(\lambda \pm 0)$ ist beweisbar (vgl. Lemma 9.1.4).

Folgende Bezeichnungen werden ebenfalls verwendet

$$R := \big\{(\alpha,\beta] \subset \mathbb{R} \;\big|\; \alpha < \beta\big\}$$

und

$$P(I) := P(\beta) - P(\alpha) \quad \text{für } I = (\alpha,\beta] \in R.$$

Unmittelbar folgt

$$P(I)P(I) = P(\beta) - P(\alpha) \quad \text{und} \quad \big(P(I)x, y\big) = \big(x, P(I)y\big).$$

Mithin ist $P(I)$ ein orthogonaler Projektor.

Die Spektralschar $P(\lambda)$ heißt von *unten beschränkt*, wenn

$$\exists\,\lambda \in \mathbb{R} \qquad P(\lambda) = O$$

gilt, sie heißt von *oben beschränkt*, wenn

$$\exists\,\lambda \in \mathbb{R} \qquad P(\lambda) = id$$

gilt, und sie heißt in $\lambda \in \mathbb{R}$ *konstant*, wenn

$$\exists\,\varepsilon > 0 \quad \forall\,\mu \in (\lambda-\varepsilon,\,\lambda+\varepsilon) \qquad P(\mu) = P(\lambda)$$

ist. Mit

$$\operatorname{supp} P := \overline{\Big\{\lambda \in \mathbb{R} \;\Big|\; P(\lambda) \text{ ist nicht konstant}\Big\}}$$

wird der *Träger* von P bezeichnet.

Bevor wir nun den eigentlichen *Spektralsatz* formulieren und beweisen, geben wir eine Reihe von Hilfssätzen, Bemerkungen und Beispielen.

Lemma 9.1.2: *Für alle* $x \in \mathcal{H}$ *wächst* $\|P(\lambda)x\|^2$ *monoton in* λ.

Beweis: Mit $\varepsilon > 0$ sei $Q := P(\lambda+\varepsilon) - P(\lambda)$. Dann ist

$$\|P(\lambda+\varepsilon)x\|^2 - \|P(\lambda)x\|^2 = \|(Q+P(\lambda))x\|^2 - \|P(\lambda)x\|^2 = \|Qx\|^2 \geq 0. \qquad \square$$

Lemma 9.1.3: *Für alle $x, y \in \mathcal{H}$ ist*

$$\begin{aligned} g \;:\; \mathbb{R} &\longrightarrow \mathbb{C}, \\ \lambda &\longmapsto \big(P(\lambda)x, y\big) \end{aligned}$$

von beschränkter Variation mit

$$\sum_{j=1}^{n} \big|g(\lambda_j) - g(\lambda_{j-1})\big| \leq \|x\| \cdot \|y\| \quad \text{für alle } \alpha = \lambda_0 < \lambda_1 < \cdots < \lambda_n = \beta.$$

Beweis: Es sei $\Lambda_j := (\lambda_{j-1}, \lambda_j]$. Dann ist

$$\begin{aligned} \sum_{j=1}^{n} \big|g(\lambda_j) - g(\lambda_{j-1})\big| &= \sum_{j=1}^{n} \big|\big(P(\Lambda_j)x, y\big)\big| = \sum_{j=1}^{n} \big|\big(P(\Lambda_j)x,\, P(\Lambda_j)y\big)\big| \leq \\ &\leq \sqrt{\sum_{j=1}^{n} \big\|P(\Lambda_j)x\big\|^2} \sqrt{\sum_{j=1}^{n} \big\|P(\Lambda_j)y\big\|^2} = \sqrt{\big\|P\big((\alpha,\beta]\big)x\big\|^2}\, \sqrt{\big\|P\big((\alpha,\beta]\big)y\big\|^2} \leq \|x\| \cdot \|y\|. \end{aligned}$$

$\square$

Lemma 9.1.4: *Für alle $\lambda \in \mathbb{R}$ existieren die Limites*

$$P(\lambda \pm 0) := \operatorname*{s\text{-}lim}_{\varepsilon \downarrow 0} P(\lambda \pm \varepsilon).$$

Zum Beweis wählen wir eine Folge $\varepsilon_n \downarrow 0$. Dann folgt aus dem vorangegangenen Lemma

$$\|P(\lambda+\varepsilon_m)x - P(\lambda+\varepsilon_n)x\|^2 = |g(\lambda+\varepsilon_m) - g(\lambda+\varepsilon_n)| \to 0. \qquad \square$$

Beispiel 9.1.5: *Es seien $G \subset \mathbb{R}^n$ beschränkt, $\mathcal{H} = \mathcal{L}^2(G)$ sowie*

$$\mathcal{D}(A) = \Big\{ u \in \mathring{\mathcal{H}}_1(G) \;\Big|\; \Delta u \in \mathcal{H} \Big\} \quad \text{und} \quad Au = -\Delta u.$$

Wir wissen bereits, daß A abzählbar unendlich viele Eigenwerte λ_n,

$$0 < \lambda_1 \leq \lambda_2 \leq \cdots,$$

und zugehörige orthonormale Eigenfunktionen u_n mit

$$id = \sum_{n=1}^{\infty} (\,\cdot\,, u_n) u_n$$

besitzt und daß

$$A = \sum_{n=1}^{\infty} \lambda_n (\,\cdot\,, u_n) u_n$$

ist. Es sei

$$P(\lambda) := \sum_{\lambda_n \leq \lambda} (\,\cdot\,, u_n) u_n.$$

Dann gilt

$$\forall\, x \in \mathcal{D}(A) \quad \forall\, y \in \mathcal{H} \qquad (Ax, y) = \int_{\mathbb{R}} \lambda \, d\big(P(\lambda)x, y\big). \tag{$*$}$$

Man nennt deshalb $P(\lambda)$ auch die *zu A gehörende Spektralschar* und schreibt kurz

$$A = \int_{\mathbb{R}} \lambda \, dP(\lambda).$$

Das ist die *Spektraldarstellung* von A. $\square$

Im folgenden wollen wir eine solche Darstellung für beliebige selbstadjungierte Abbildungen beweisen. Für $x \in \mathcal{H}$ wächst

$$f(\lambda) := \big(P(\lambda)x, x\big)$$

monoton und erzeugt vermöge

$$\mu_x\big((\alpha,\beta]\big) := f(\beta) - f(\alpha)$$

das Lebesgue-Stieltjessche Maß μ_x. Es sei $\mathcal{L}^2_{\mu_x}(\mathbb{R})$ der Raum der bezüglich μ_x quadratintegrablen Funktionen. In Beispiel 9.1.5 folgt damit aus Gl. $(*)$

$$\forall\, x \in \mathcal{D}(A) \qquad \|Ax\|^2 = \int_{\mathbb{R}} \lambda^2 \, d\mu_x(\lambda).$$

Für $x \in \mathcal{D}(A)$ ist also $id \in \mathcal{L}^2_{\mu_x}(\mathbb{R})$. Umgekehrt erhält man aus $id \in \mathcal{L}^2_{\mu_x}(\mathbb{R})$ die Existenz des Integrals auf der rechten Seite von Gl. $(*)$ durch Polarisieren.

Beispiel 9.1.6: *Es seien $\mathcal{H} = \mathcal{L}^2(\mathbb{R})$ und A der Abschluß von*

$$\begin{aligned} B \;:\; \mathring{\mathcal{C}}_\infty(\mathbb{R}) \subset \mathcal{H} &\longrightarrow \mathcal{H}, \\ x &\longmapsto -\mathrm{i}\, x'. \end{aligned}$$

Dann ist A eine symmetrische Abbildung mit $\mathcal{D}(A) = \mathcal{H}_1(\mathbb{R})$ und (vgl. Satz 5.3.10)

$$\mathcal{D}(A^*) = \Big\{ u \in \mathcal{H} \;\Big|\; \exists\, f \in \mathcal{H} \quad \forall\, v \in \mathcal{D}(B) \quad (u, Bv) = (f, v) \Big\} = \mathcal{W}_1(\mathbb{R}).$$

Wegen $\mathcal{W}_1(\mathbb{R}) = \mathcal{H}_1(\mathbb{R})$ ist A selbstadjungiert. Aus

$$Au = f$$

folgt durch Fouriertransformation

$$\xi\, \hat{u}(\xi) = \hat{f}(\xi),$$

also

$$\big(Au\big)(x) = F^{-1}\big(\xi\, \hat{u}(\xi)\big)(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \xi e^{ix\xi} \hat{u}(\xi)\, d\xi.$$

Unsere Abbildung A besitzt kein Punktspektrum, sondern ein stetiges Spektrum, und die letzte Formel liefert die Spektraldarstellung von A. Die zugehörende Spektralschar wird durch

$$\big(P(\lambda)u\big)(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\lambda} e^{ix\xi} \hat{u}(\xi)\, d\xi$$

gegeben. Der Beweis wird in Beispiel 9.3.8 ausgeführt. Man findet hier also den Zusammenhang mit der Fouriertransformation. Einzelheiten wollen wir jetzt nicht nachrechnen. Ist das zugrundeliegende Gebiet beschränkt, etwa $\mathcal{H} = \mathcal{L}^2([-\pi,\pi])$, dann erhält man entsprechend eine Darstellung durch die Fourierreihe. □

Es seien nun $\Lambda := (\alpha,\beta] \in R$ ein Intervall, $x \in \mathcal{H}$ fest und μ_x das von $(P(\lambda)x, x)$ erzeugte Maß auf $\mathbb{R}$; wir schreiben $\mu_x(\lambda) := \big(P(\lambda)x, x\big) = \|P(\lambda)x\|^2$. Dann kann man auch das $\mathcal{H}$-wertige *Lebesgue-Stieltjessche Integral* einführen. Es seien

$$\alpha = \lambda_0 < \lambda_1 < \cdots < \lambda_n = \beta$$

und $\Lambda_j := (\lambda_{j-1}, \lambda_j] \in R$ für $1 \le j \le n$. Es sei ferner

$$t := \sum_{j=1}^{n} c_j \chi_{\Lambda_j}$$

eine Treppenfunktion mit Werten in $\mathbb{C}$. Dann definieren wir

$$\int_\Lambda t(\lambda)\, dP(\lambda)x := \sum_{j=1}^{n} c_j\, P(\Lambda_j)x.$$

Offenbar ist

$$\Big\| \int_\Lambda t(\lambda)\; dP(\lambda)x \Big\| \le \|t\|_{\mathcal{L}^\infty(\Lambda)} \cdot \|x\|$$

und

$$\Big\| \int_\Lambda t(\lambda)\ dP(\lambda)x \Big\|^2 = \sum_{j=1}^n |c_j|^2 \big(P(\Lambda_j)x, x\big) = \int_\Lambda |t|^2\ d\mu_x.$$

Durch Vervollständigen erhält man daraus die Existenz der beschränkten Abbildung

$$\begin{aligned} J \ : \ \mathcal{L}^2_{\mu_x}(\Lambda, \mathbb{C}) &\longrightarrow \mathcal{H}, \\ f &\longmapsto \int_\Lambda f(\lambda)\, dP(\lambda)x \end{aligned}$$

mit $\|Jf\|_{\mathcal{H}} = \|f\|_{\mathcal{L}^2_{\mu_x}(\Lambda)}$.

Auch uneigentliche Integrale lassen sich definieren, zum Beispiel

$$\int_{\mathbb{R}} f(\lambda)\ dP(\lambda)x := \operatorname*{s-lim}_{\substack{\alpha \to -\infty \\ \beta \to \infty}} \int_\alpha^\beta f(\lambda)\ dP(\lambda)x,$$

falls diese Limites existieren. Wir stellen einige einfache Eigenschaften des Integrals zusammen:

Satz 9.1.7: *Es seien $x \in \mathcal{H}$ und $f \in \mathcal{C}(\mathbb{R}, \mathbb{C})$. Dann sind folgende Aussagen äquivalent:*

1. *$\int_{\mathbb{R}} f(\lambda)\, dP(\lambda)x$ existiert.*
2. *$f \in \mathcal{L}^2_{\mu_x}(\mathbb{R}, \mathbb{C})$.*
3. *Die Abbildung*

$$\begin{aligned} F \ : \ \mathcal{H} &\longrightarrow \mathbb{C}, \\ h &\longmapsto \int_{\mathbb{R}} f(\lambda)\, d\big(P(\lambda)h, x\big) \end{aligned}$$

ist ein beschränktes lineares Funktional.

Dabei wird das letzte Integral wieder durch Polarisieren erklärt.

Beweis: Zunächst bemerken wir, daß wegen der Stetigkeit des Skalarprodukts für $f \in \mathcal{C}(\mathbb{R}, \mathbb{C})$

$$\int_\alpha^\beta f(\lambda)\, d\big(P(\lambda)h, x\big) = \Big(h, \int_\alpha^\beta \overline{f(\lambda)}\, dP(\lambda)x\Big)$$

ist. Wir beweisen nun im einzelnen

(1) $\Rightarrow$ (3): Für alle $y \in \mathcal{H}$ ist

$$\Big(y, \int_{-\infty}^\infty \overline{f(\lambda)}\, dP(\lambda)x\Big) = \lim_{\substack{\alpha \to -\infty \\ \beta \to \infty}} \Big(y, \int_\alpha^\beta \overline{f(\lambda)}\, dP(\lambda)x\Big) = \lim_{\substack{\alpha \to -\infty \\ \beta \to \infty}} F_{\alpha\beta}\, y = Fy$$

mit

$$F_{\alpha\beta}\, y := \int_\alpha^\beta f(\lambda)\, d\big(P(\lambda)y, x\big).$$

Nun folgt aus

$$|F_{\alpha\beta}\, y| = \Big|\Big(y, \int_{-\infty}^\infty \cdots\Big) - \Big(y, \int_{-\infty}^\alpha \cdots\Big) - \Big(y, \int_\beta^\infty \cdots\Big)\Big| \le \|y\| \Big\{ \Big\| \int_{-\infty}^\infty \cdots \Big\| + \cdots \Big\}$$

mit

$$k := \Big\| \int_{-\infty}^\infty \overline{f(\lambda)}\, dP(\lambda)x \Big\|$$

die Abschätzung

$$\forall\, \varepsilon > 0 \quad \exists\, N > 0 \quad \forall\, -\alpha, \beta > N \qquad |F_{\alpha\beta}\, y| \le \|y\|\, \{k + \varepsilon\}$$

oder

$$\forall\, \varepsilon > 0 \quad \exists\, N > 0 \quad \forall\, -\alpha, \beta > N \qquad \|F_{\alpha\beta}\| \le k + \varepsilon.$$

Daraus erhalten wir

$$\|F\| = \sup_{y \in S_{\mathcal{H}}} |Fy| = \sup_{y \in S_{\mathcal{H}}} \left\{ \lim_{\substack{\alpha \to -\infty \\ \beta \to \infty}} |F_{\alpha\beta}\, y| \right\} \leq k.$$

(3) $\Rightarrow$ (2): Es sei

$$y := \int_\alpha^\beta \overline{f(\lambda)}\, dP(\lambda)x.$$

Dann folgt aus $P\big((\alpha,\beta]\big)y = y$

$$\overline{Fy} = \int_{\mathbb{R}} \overline{f(\lambda)}\, d\big(P(\lambda)x, y\big) = \int_\alpha^\beta \overline{f(\lambda)}\, d\big(P(\lambda)x, y\big) = \|y\|^2.$$

Deshalb gilt die Abschätzung $\|y\| \leq \|F\|$, also

$$\|y\|^2 = \int_\alpha^\beta |f(\lambda)|^2\, d\mu_x(\lambda) \leq \|F\|^2.$$

(2) $\Rightarrow$ (1): Für $\alpha' < \alpha < \beta < \beta'$ ist für $\alpha, \alpha' \to -\infty$ und $\beta, \beta' \to \infty$

$$\Big\| \int_{\alpha'}^{\beta'} f(\lambda)\, dP(\lambda)x - \int_\alpha^\beta f(\lambda)\, dP(\lambda)x \Big\|^2 = \int_{\alpha'}^{\alpha} |f(\lambda)|^2\, d\mu_x(\lambda) + \int_\beta^{\beta'} |f(\lambda)|^2\, d\mu_x(\lambda) \to 0. \qquad \square$$

Satz 9.1.8: *Es seien $f \in \mathcal{C}(\mathbb{R},\mathbb{R})$ und $H : \mathcal{D}(H) \subset \mathcal{H} \to \mathcal{H}$ mit*

$$\mathcal{D}(H) := \Big\{ x \in \mathcal{H} \;\Big|\; f \in \mathcal{L}^2_{\mu_x}(\mathbb{R}) \Big\} \quad \textit{und} \quad Hx := \int_{\mathbb{R}} f(\lambda)\, dP(\lambda)x.$$

Dann gilt:

1. *H ist selbstadjungiert.*
2. *Für alle $\Lambda \in R$ ist $\mathcal{R}(P(\Lambda)) \subset \mathcal{D}(H)$.*
3. *$P(\lambda)H \subset HP(\lambda)$.*

Beweis: Als erstes zeigen wir $H \subset H^*$: Es ist

$$\forall\, x \in \mathcal{D}(H) \qquad \|Hx\|_{\mathcal{H}} = \|f\|_{\mathcal{L}^2_{\mu_x}(\mathbb{R})}$$

und

$$\forall\, x \in \mathcal{H} \quad \forall\, \varepsilon > 0 \quad \exists\, \Lambda = (\alpha,\beta] \in R \qquad \|x - P(\Lambda)x\| < \varepsilon.$$

Ferner gilt

$$\int_{\mathbb{R}} \big|f(\lambda)\big|^2\, d\big\|P(\lambda)P(\Lambda)x\big\|^2 = \int_\alpha^\beta \big|f(\lambda)\big|^2\, d\mu_x(\lambda) < \infty,$$

also

$$\forall\, x \in \mathcal{H} \quad \forall\, \Lambda \in R \qquad P(\Lambda)x \in \mathcal{D}(H).$$

Deshalb ist $\mathcal{D}(H)$ in $\mathcal{H}$ dicht. Weil f reellwertig ist, folgt

$$(Hx, x) = \int_{\mathbb{R}} f(\lambda)\, d\mu_x(\lambda) \in \mathbb{R}.$$

Mithin ist H symmetrisch, also $H \subset H^*$.

Es sei nun $y \in \mathcal{D}(H^*)$. Dann gilt für alle $x \in \mathcal{H}$ und $\Lambda = (\alpha,\beta]$

$$\big(x, P(\Lambda)H^*y\big) = \big(HP(\Lambda)x, y\big) = \int_\alpha^\beta f(\lambda)\, d\big(P(\lambda)x, y\big) =: F_{\alpha\beta}\, x,$$

und es existiert

$$Fx := \lim_{\substack{\alpha \to -\infty \\ \beta \to \infty}} F_{\alpha\beta}\, x.$$

Nun ist $|F_{\alpha\beta}x| \leq \|x\| \cdot \|H^*y\|$, also

$$\|F\| = \sup_{x \in S_{\mathcal{H}}} |Fx| = \sup_{x \in S_{\mathcal{H}}} \left\{ \lim_{\substack{\alpha \to -\infty \\ \beta \to \infty}} |F_{\alpha\beta}\, x| \right\} \leq \|H^*y\|.$$

Deshalb ist

$$F \cdot = \int_{\mathbb{R}} f(\lambda)\, d\big(P(\lambda) \cdot, y\big)$$

ein beschränktes lineares Funktional. Aus dem vorangegangenen Satz folgt daher

$$f \in \mathcal{L}^2_{\mu_y}(\mathbb{R})$$

oder $y \in \mathcal{D}(H)$. Damit ist $H = H^*$. Das beweist die beiden ersten Behauptungen.

Zum Nachweis der dritten Behauptung wählen wir $x \in \mathcal{D}(H)$ und erhalten

$$P(\mu)Hx = \int_{\mathbb{R}} f(\lambda)\, dP(\lambda)P(\mu)x = HP(\mu)\, x. \qquad \square$$

Es sei H nun wie im letzten Satz definiert. Dann können wir auch Potenzen von H erklären, nämlich durch

$$\mathcal{D}(H^n) := \left\{ x \in \mathcal{H} \;\middle|\; f^n \in \mathcal{L}^2_{\mu_x}(\mathbb{R}) \right\}$$

und

$$H^n x := \int_{\mathbb{R}} f^n(\lambda)\, dP(\lambda)x.$$

Das folgt aus

$$(Hx, Hy) = \cdots = \int_{\mathbb{R}} f^2(\lambda)\, d\big(P(\lambda)x, y\big).$$

Auch stetige Funktionen von H lassen sich auf diese Weise definieren, zum Beispiel $\cos H$.

Beispiel 9.1.9: *Wir betrachten den speziellen Fall $f(\lambda) = \lambda$, also*

$$\mathcal{D}(H) := \left\{ x \in \mathcal{H} \;\middle|\; \int_{\mathbb{R}} \lambda^2\, d\mu_x(\lambda) < \infty \right\} \quad \textit{und} \quad H := \int_{\mathbb{R}} \lambda\, dP(\lambda).$$

Diese selbstadjungierte Abbildung H nennt man die *zur Spektralschar $P(\lambda)$ gehörende Abbildung.* $\square$

Im folgenden möchten wir umgekehrt zu einer vorgegebenen selbstadjungierten Abbildung A die *zu A gehörende Spektralschar $P(\lambda)$* bestimmen. Dann soll also

$$\forall\, i \in \mathbb{N}_0 \qquad \big(A^i x, y\big) = \int_{\mathbb{R}} \lambda^i\, d\big(P(\lambda)x, y\big)$$

gelten. Für beschränkte Abbildungen haben wir im Anschluß an die Behandlung des Hausdorffschen Momentenproblems in §6.7.1 die Existenz von $P(\lambda)$ bereits sehen können. Im nächsten Abschnitt wird eine solche Aussage allgemein formuliert und bewiesen.

Wir beschließen diesen Abschnitt mit einigen weiteren Beispielen. Dabei sei stets H die zu $P(\lambda)$ gehörende Abbildung.

Beispiel 9.1.10: *Es sei*

$$P(\lambda) := \begin{cases} O & \textit{für } \lambda < 0 \\ id & \textit{für } \lambda \geq 0. \end{cases}$$

Dann ist $P(\lambda)$ beschränkt, also $\mathcal{D}(H) = \mathcal{H}$ und

$$\forall\, x \in \mathcal{H} \qquad \|Hx\|^2 = \int_{\mathbb{R}} \lambda^2\, d\mu_x(\lambda) = 0.$$

Mithin ist $H = O$. $\square$

Beispiel 9.1.11: *Es sei Q eine orthogonale Projektion und*

$$P(\lambda) := \begin{cases} O & \textit{für } \lambda < 0 \\ id - Q & \textit{für } 0 \leq \lambda < 1 \\ id & \textit{für } 1 \leq \lambda. \end{cases}$$

Auch jetzt ist $P(\lambda)$ beschränkt, also $\mathcal{D}(H) = \mathcal{H}$, und

$$Hx = \int_{\mathbb{R}} \lambda \, dP(\lambda)x = \lambda \, P(\lambda)x \Big|_{1-0}^{1+0} = \big(id - (id - Q)\big)x = Qx. \qquad \square$$

Beispiel 9.1.12: *Es seien $\mathcal{H} = \mathcal{L}^2(\mathbb{R})$ und fast überall*

$$\big(P(\lambda)x\big)(t) := \begin{cases} x(t) & \text{für } t \le \lambda \\ 0 & \text{für } t > \lambda. \end{cases}$$

Dann ist

$$\|Hx\|^2 = \int_{\mathbb{R}} \lambda^2 \, d\mu_x(\lambda) = \int_{\mathbb{R}} \lambda^2 \, d\Big(\int_{-\infty}^{\lambda} |x(t)|^2 dt \Big) = \int_{\mathbb{R}} \lambda^2 \, |x(\lambda)|^2 \, d\lambda,$$

also

$$\mathcal{D}(H) = \Big\{ x \in \mathcal{H} \;\Big|\; \int_{\mathbb{R}} t^2 \, |x(t)|^2 \, dt < \infty \Big\},$$

und es folgt für $x \in \mathcal{D}(H)$ und $y \in \mathcal{H}$

$$\big(Hx, y\big) = \int_{\mathbb{R}} \lambda \, d\big(P(\lambda)x, y\big) = \int_{\mathbb{R}} \lambda \, d\Big(\int_{-\infty}^{\lambda} x(t) \, \overline{y(t)} \, dt \Big) = \int_{\mathbb{R}} \lambda \, x(\lambda) \, \overline{y(\lambda)} \, d\lambda$$

oder fast überall

$$\big(Hx\big)(t) = t \, x(t). \qquad \square$$

9.2 Der Spektralsatz

Bisher haben wir positive Maße μ_x betrachtet, die von $\big(P(\lambda)x, x\big)$ erzeugt waren, also für $\Lambda \in R$

$$\mu_x(\Lambda) = \big(P(\Lambda)x, x\big).$$

Im folgenden wollen wir auch komplexe Maße zulassen. Zur Klarstellung wiederholen wir zugrundeliegende Begriffe:

Es sei $\mathcal{B}$ die von R erzeugte *σ-Algebra*; die Elemente von $\mathcal{B}$ heißen *Borelmengen*. Dann gilt

1. $\emptyset \in \mathcal{B}$.
2. $M \in \mathcal{B} \Longrightarrow \mathbb{R} \backslash M \in \mathcal{B}$.
3. $M_i \in \mathcal{B}, i \in \mathbb{N}, \Longrightarrow \cup_{i \in \mathbb{N}} M_i \in \mathcal{B}$.

Eine Abbildung

$$\mu \;:\; \mathcal{B} \longrightarrow [0, \infty]$$

heißt *positives Borelmaß*, wenn folgendes gilt:

1. $\mu \not\equiv \infty$.
2. Für alle paarweise disjunkten $M_k \in \mathcal{B}$ ist

$$\mu\Big(\bigcup_{k \in \mathbb{N}} M_k \Big) = \sum_{k \in \mathbb{N}} \mu(M_k).$$

Ein positives Borelmaß mit

$$\mu(\mathbb{R}) < \infty$$

heißt *beschränkt*. Eine Abbildung

$$\mu \;:\; \mathcal{B} \longrightarrow \mathbb{C}$$

heißt *komplexes Borelmaß*, wenn für alle paarweise disjunkten $M_k \in \mathcal{B}$

$$\mu\Big(\bigcup_{k \in \mathbb{N}} M_k \Big) = \sum_{k \in \mathbb{N}} \mu(M_k)$$

gilt. Ein komplexes Borelmaß ist also stets beschränkt. Es ist durch seine Werte auf R eindeutig bestimmt. Man nennt

$$|\mu|(M) := \sup \sum_{i=1}^{\infty} |\mu(M_i)|$$

Totalvariation von μ. Dabei wird das Supremum über alle Partitionen M_i von M genommen. $|\mu|$ ist ein beschränktes positives Maß. Ist μ reellwertig, dann definiert man

$$\mu^+ := \tfrac{1}{2}\big(|\mu| + \mu\big) \quad \text{und} \quad \mu^- := \tfrac{1}{2}\big(|\mu| - \mu\big).$$

Beide Maße, μ^+ und μ^-, sind positiv und beschränkt. Es gilt

$$\mu = \mu^+ - \mu^- \quad \text{und} \quad |\mu| = \mu^+ + \mu^-.$$

Damit läßt sich auch ein komplexes Maß in positive zerlegen, nämlich

$$\mu = \mu_1 - \mu_2 + \mathrm{i}\,\mu_3 - \mathrm{i}\,\mu_4$$

mit

$$\mu_1 := \tfrac{1}{2}\big(|\mathrm{Re}\,\mu| + \mathrm{Re}\,\mu\big), \quad \mu_2 := \tfrac{1}{2}\big(|\mathrm{Re}\,\mu| - \mathrm{Re}\,\mu\big), \quad \mu_3 := \tfrac{1}{2}\big(|\mathrm{Im}\,\mu| + \mathrm{Im}\,\mu\big), \quad \mu_4 := \tfrac{1}{2}\big(|\mathrm{Im}\,\mu| - \mathrm{Im}\,\mu\big).$$

Es seien nun $P(\lambda)$ wieder eine Spektralschar, $x \in \mathcal{H}$ fest und μ_x das von $\big(P(\,\cdot\,)x, x\big)$ erzeugte Maß. Dann ist μ_x ein beschränktes positives Maß, denn es gilt

$$\mu_x(\mathbb{R}) = \|x\|^2 < \infty.$$

Für $x, y \in \mathcal{H}$ sind auch die Borelmaße

$$\mu_{x+y}, \quad \mu_{x-y}, \quad \mu_{x+\mathrm{i}y}, \quad \mu_{x-\mathrm{i}y}$$

positiv und beschränkt. Wir definieren das zugehörende komplexe Borelmaß durch Polarisieren, nämlich

$$\mu_{x,y}(M) := \tfrac{1}{4}\big\{\mu_{x+y}(M) - \mu_{x-y}(M) + \mathrm{i}\,\mu_{x+\mathrm{i}y}(M) - \mathrm{i}\,\mu_{x-\mathrm{i}y}(M)\big\}.$$

Dann gelten

$$\mu_{x,y}(\Lambda) = \big(P(\Lambda)x, y\big),$$

$\mu_{x,x} = \mu_x$ und für $M \in \mathcal{B}$

$$|\mu_{x,y}(M)| \le \sqrt{\mu_x(M) \cdot \mu_y(M)}.$$

Insbesondere ist also

$$|\mu_{x,y}(\mathbb{R})| \le \|x\| \cdot \|y\|.$$

Man vergleiche Lemma 9.1.3.

Es seien nun μ ein komplexes Borelmaß und für $z \in \mathbb{C}\backslash\mathbb{R}$

$$f(z) := \int_{\mathbb{R}} \frac{d\mu(t)}{t - z}.$$

Dann zeigen wir als erstes den nach Thomas Stieltjes, 1856–94, benannten

Satz 9.2.1 (Die Stieltjessche Umkehrformel): *Es gilt*

1.

$$\forall\, t \in \mathbb{R} \qquad \mu\big((-\infty, t]\big) = \lim_{\delta \downarrow 0} \lim_{\varepsilon \downarrow 0} \frac{1}{2\pi\mathrm{i}} \int_{-\infty}^{t+\delta} \Big(f(s + \mathrm{i}\varepsilon) - f(s - \mathrm{i}\varepsilon)\Big)\, ds.$$

2. Das Maß μ ist durch die Werte von f auf $\mathbb{C}\backslash\mathbb{R}$ eindeutig bestimmt.

Beweis:

1. Es seien μ beschränkt und positiv, $r \in \mathbb{R}$ und $\varepsilon > 0$. Dann ist mit $\mu(t) := \mu\big((-\infty, t]\big)$, μ rechtsstetig,

$$\frac{1}{2\pi\mathrm{i}}\Big(f(s + \mathrm{i}\varepsilon) - f(s - \mathrm{i}\varepsilon)\Big) = \frac{1}{\pi} \int_{\mathbb{R}} \frac{\varepsilon}{(s-t)^2 + \varepsilon^2}\, d\mu(t).$$

Nun gilt

$$\frac{1}{\pi} \int_{-\infty}^{r} \frac{\varepsilon}{(s-t)^2 + \varepsilon^2}\, ds = \frac{1}{\pi} \arctan \frac{r-t}{\varepsilon} + \frac{1}{2}.$$

Aus dem Satz von Fubini folgt deshalb

$$\frac{1}{2\pi\mathrm{i}}\int_{-\infty}^{r}\big(f(s+\mathrm{i}\varepsilon)-f(s-\mathrm{i}\varepsilon)\big)ds=\int_{\mathbb{R}}\Big(\frac{1}{\pi}\arctan\frac{r-t}{\varepsilon}+\frac{1}{2}\Big)d\mu(t).$$

Wegen

$$\lim_{\varepsilon\downarrow 0}\frac{1}{\pi}\arctan\frac{r-t}{\varepsilon}+\frac{1}{2}=\begin{cases}1 & \text{für } t<r\\ 1/2 & \text{für } t=r\\ 0 & \text{für } t>r\end{cases}$$

erhalten wir daher aus dem Satz von Lebesgue

$$\lim_{\varepsilon\downarrow 0}\frac{1}{2\pi\mathrm{i}}\int_{-\infty}^{r}\big(f(s+\mathrm{i}\varepsilon)-f(s-\mathrm{i}\varepsilon)\big)ds=\mu(r-0)+\frac{1}{2}\big(\mu(r)-\mu(r-0)\big)=\frac{1}{2}\big(\mu(r)+\mu(r-0)\big).$$

Setzt man nun $r=t+\delta$, $\delta>0$, dann folgt in diesem Fall die Behauptung.

2. Es sei μ nun komplex. Dann kann man μ wie besprochen in positive Maße zerlegen, nämlich

$$\mu=\mu_1-\mu_2+\mathrm{i}\,\mu_3-\mathrm{i}\,\mu_4.$$

Mit

$$f_j(z):=\int_{\mathbb{R}}\frac{d\mu_j(t)}{t-z}$$

folgt deshalb aus dem Vorhergehenden die Umkehrung $\mu_j=\cdots$. Das beweist die erste Aussage.

3. Es sei nun $f|_{\mathbb{C}\setminus\mathbb{R}}=o$. Dann folgt aus der ersten Aussage

$$\forall\, t\in\mathbb{R}\qquad \mu((-\infty,t])=0.$$

Das Maß μ ist aber nach Konstruktion durch seine Werte auf R eindeutig festgelegt; es verschwindet also. □

Wir wählen nun den separablen Raum

$$\mathcal{C}^0(\mathbb{R}):=\Big\{f\in\mathcal{C}(\mathbb{R})\;\Big|\;\lim_{x\to\pm\infty}f(x)=0\Big\}.$$

Im folgenden sind wir an den positiven stetigen Funktionalen über $\mathcal{C}^0(\mathbb{R})$ interessiert.

Definition 9.2.2: *Das Funktional $\Lambda\in\mathcal{C}^0(\mathbb{R})'$ heißt* positiv, *wenn*

$$\forall\,\varphi\in\mathcal{C}^0(\mathbb{R}),\ \varphi\geq 0,\qquad \Lambda\varphi\geq 0$$

gilt.

Es sei an die vergleichbare Situation der Funktionale vom positiven Typ beim Hausdorffschen Momentenproblem erinnert. Es ist $\mathcal{C}([a,b])'$ zum Raum der normalisierten Funktionen von beschränkter Variation, $\mathcal{NBV}([a,b])$, kongruent (vgl. den Beweis zu Satz 4.2.1). Funktionen von beschränkter Variation lassen sich immer als Differenz zweier monoton wachsender Funktionen schreiben. So erhält man den

Satz 9.2.3 (Darstellungssatz für positive lineare Funktionale aus $\mathcal{C}^0(\mathbb{R})'$): *Zu jedem positiven $\Lambda\in\mathcal{C}^0(\mathbb{R})'$ gibt es genau ein beschränktes positives Borelmaß μ mit $\mu(\mathbb{R})=\|\Lambda\|$ und*

$$\forall\,\varphi\in\mathcal{C}^0(\mathbb{R})\qquad \Lambda\varphi=\int_{\mathbb{R}}\varphi\,d\mu.$$

Einen Beweis dieses Darstellungssatzes findet man in *Rudin [1987], S. 130.* Wir skizzieren ihn in den Ergänzungen §10.2.

Es sei nun

$$\mathbb{H}:=\big\{z\in\mathbb{C}\;\big|\;\mathrm{Im}\,z>0\big\}$$

die obere komplexe Halbebene. Wir wollen das zweite wichtige Resultat zeigen, nämlich

Satz 9.2.4: *Es seien $f : \mathbb{H} \to \mathbb{C}$ holomorph und*

1. $\forall\, z \in \mathbb{H} \quad \operatorname{Im} f(z) \geq 0.$

2. $\exists\, p \geq 0 \quad \forall\, z \in \mathbb{H} \qquad |f(z)| \operatorname{Im} z \leq p.$

Dann gibt es genau ein beschränktes positives Borelmaß μ mit $\mu(\mathbb{R}) \leq p$ und

$$\forall\, z \in \mathbb{H} \qquad f(z) = \int_{\mathbb{R}} \frac{d\mu(t)}{t - z}.$$

Den Beweis führen wir in sechs Schritten:

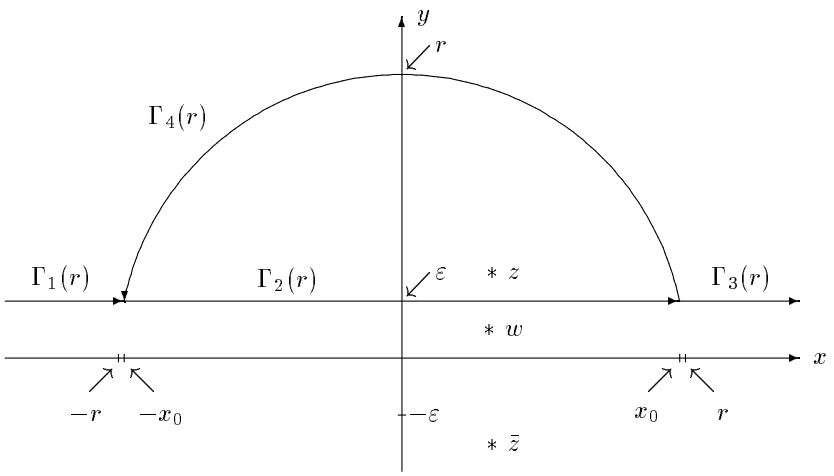
1. Es seien x, y, ε fest mit $z := x + \mathrm{i}y \in \mathbb{H}$, $\varepsilon < y$ und

$$w := \bar{z} + 2\mathrm{i}\varepsilon = x + \mathrm{i}(2\varepsilon - y).$$

Wir wählen mit $r \geq 2|z|$ und $x_0 := \sqrt{r^2 - \varepsilon^2}$ folgende Wege in $\mathbb{H}$ (man vergleiche die folgende Skizze):

$$\Gamma_1(r) := \left\{ (s, \mathrm{i}\varepsilon) \;\middle|\; -\infty < s < -x_0 \right\}, \qquad \Gamma_2(r) := \left\{ (s, \mathrm{i}\varepsilon) \;\middle|\; -x_0 < s < x_0 \right\},$$

$$\Gamma_3(r) := \left\{ (s, \mathrm{i}\varepsilon) \;\middle|\; x_0 < s < \infty \right\}, \qquad \Gamma_4(r) := \left\{ (s,t) \in \partial B(0,r) \;\middle|\; t > \varepsilon \right\}.$$



Dann folgt aus der Cauchyschen Integralformel

$$f(z) = \frac{1}{2\pi\mathrm{i}} \int_{\Gamma_2 + \Gamma_4} \frac{f(\zeta)}{\zeta - z}\, d\zeta$$

und

$$0 = \frac{1}{2\pi\mathrm{i}} \int_{\Gamma_2 + \Gamma_4} \frac{f(\zeta)}{\zeta - w}\, d\zeta,$$

zusammen also

$$f(z) = \frac{1}{2\pi\mathrm{i}} \int_{\Gamma_2 + \Gamma_4} \frac{z - w}{(\zeta - z)(\zeta - w)} f(\zeta)\, d\zeta.$$

Nun ist für $\zeta \in \Gamma_1 \cup \Gamma_3 \cup \Gamma_4$

$$|\zeta| \geq r \quad \text{und} \quad \operatorname{Im} \zeta \geq \varepsilon,$$

also wegen $|\zeta - z| \geq r - r/2 \geq r/2$

$$\left| \frac{z - w}{(\zeta - z)(\zeta - w)} \right| \leq \frac{4|z - w|}{r^2} \quad \text{und} \quad |f(\zeta)| \leq \frac{p}{\varepsilon}.$$

Mithin gilt für $j = 1, 3, 4$

$$\lim_{r \to \infty} \int_{\Gamma_j(r)} \left| \frac{z - w}{(\zeta - z)(\zeta - w)} \right| |f(\zeta)|\, d\zeta = 0.$$

Deshalb ist

$$\int_{\operatorname{Im} \zeta = \varepsilon} \frac{z - w}{(\zeta - z)(\zeta - w)} f(\zeta)\, d\zeta = \int_{-\infty}^{\infty} \frac{2\mathrm{i}(y - \varepsilon)}{(x - t)^2 + (y - \varepsilon)^2} f(t + \mathrm{i}\varepsilon)\, dt$$

wohldefiniert, und es folgt

$$f(z) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{y-\varepsilon}{(x-t)^2+(y-\varepsilon)^2}\, f(t+\mathrm{i}\varepsilon)\, dt.$$

Für $v := \operatorname{Im} f$ gilt entsprechend

$$v(x+\mathrm{i}y) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{y-\varepsilon}{(x-t)^2+(y-\varepsilon)^2}\, v(t+\mathrm{i}\varepsilon)\, dt.$$

2. Wir zeigen

$$\forall\, \varepsilon > 0 \qquad \frac{1}{\pi}\int_{-\infty}^{\infty} v(t+\mathrm{i}\varepsilon)\, dt \le p.$$

Für alle $n > \varepsilon$ gilt nämlich

$$0 \le \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{(n-\varepsilon)^2}{t^2+(n-\varepsilon)^2} v(t+\mathrm{i}\varepsilon)\, dt = (n-\varepsilon)v(\mathrm{i}n) \le n v(\mathrm{i}n) \le p,$$

und für $t \in \mathbb{R}$ strebt

$$\left(\frac{(n-\varepsilon)^2}{t^2+(n-\varepsilon)^2}\, v(t+\mathrm{i}\varepsilon)\right)_{n>\varepsilon}$$

mit $n \to \infty$ monoton gegen $v(t+\mathrm{i}\varepsilon)$. Aus dem Satz von B. Levi über die monotone Konvergenz folgt deshalb die Behauptung.

3. Wir zeigen

$$\forall\, x, y \in \mathbb{R},\, y > 0, \qquad v(x+\mathrm{i}y) = \lim_{\varepsilon \downarrow 0} \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{y}{(x-t)^2+y^2} v(t+\mathrm{i}\varepsilon)\, dt.$$

Man ersetze nämlich im ersten Resultat y durch $y+\varepsilon$, $\varepsilon > 0$. Dann folgt für alle $y > 0$

$$v\big(x+\mathrm{i}(y+\varepsilon)\big) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{y}{(x-t)^2+y^2} v(t+\mathrm{i}\varepsilon)\, dt,$$

und links existiert der Limes.

4. Es seien nun $\varepsilon > 0$, $\varphi \in \mathcal{C}^0(\mathbb{R})$ und

$$\Lambda_\varepsilon \varphi := \frac{1}{\pi}\int_{\mathbb{R}} \varphi(t) v(t+\mathrm{i}\varepsilon)\, dt.$$

Dann ist $\Lambda_\varepsilon \in \mathcal{C}^0(\mathbb{R})'$, und aus dem zweiten Resultat folgt $\|\Lambda_\varepsilon\| \le p$. Außerdem ist Λ_ε wegen $v \ge 0$ ein positives Funktional. Nun ist $\mathcal{C}^0(\mathbb{R})$ separabel, nach Satz 4.1.2 ist die Kugel in $\mathcal{C}^0(\mathbb{R})'$ deshalb schwach-stern folgenkompakt. Mithin gibt es eine Nullfolge (ε_n) und ein $\Lambda \in \mathcal{C}^0(\mathbb{R})'$ mit

$$\forall\, \varphi \in \mathcal{C}^0(\mathbb{R}) \qquad \lim_{n\to\infty} \Lambda_{\varepsilon_n}\varphi = \Lambda\varphi.$$

Das Funktional Λ ist ebenfalls positiv mit $\|\Lambda\| \le p$, denn die Norm ist schwach-stern halbstetig nach unten (Satz 4.0.6). Aus dem Darstellungssatz folgt deshalb die Existenz eines positiven Borelmaßes μ mit $\mu(\mathbb{R}) = \|\Lambda\|$ und

$$\forall\, \varphi \in \mathcal{C}^0(\mathbb{R}) \qquad \Lambda\varphi = \int_{\mathbb{R}} \varphi\, d\mu.$$

Es seien nun $x, y \in \mathbb{R}$, $y > 0$, und

$$\varphi(t) = \varphi(t; x, y) := \frac{y}{(x-t)^2+y^2}.$$

Dann ist $\varphi \in \mathcal{C}^0(\mathbb{R})$, und es folgt aus der dritten Aussage

$$v(x+\mathrm{i}y) = \lim_{n\to\infty} \frac{1}{\pi}\int_{\mathbb{R}} \varphi(t; x, y)\, v(t+\mathrm{i}\,\varepsilon_n)\, dt = \Lambda\varphi(\,\cdot\,; x, y) = \int_{\mathbb{R}} \varphi(\,\cdot\,; x, y)\, d\mu.$$

5. Wir zeigen

$$\forall\, z \in \mathbb{H} \qquad f(z) = \int_{\mathbb{R}} \frac{d\mu(t)}{t-z}.$$

Es sei nämlich

$$\tilde{f} \;:\; \mathbb{H} \longrightarrow \mathbb{C}$$

holomorph mit

$$\tilde{f}(z) := \int_{\mathbb{R}} \frac{d\mu(t)}{t-z}.$$

Wegen $\mu(\mathbb{R}) \leq p$ ist dann

$$\left|\tilde{f}(z)\right| \leq \frac{p}{\operatorname{Im} z},$$

und für alle $x, y, \in \mathbb{R}$ mit $y > 0$ gilt

$$\operatorname{Im} \tilde{f}(x+\mathrm{i}y) = \int_{\mathbb{R}} \operatorname{Im} \frac{1}{t-x-\mathrm{i}y}\, d\mu(t) = \int_{\mathbb{R}} \frac{y}{(x-t)^2+y^2}\, d\mu(t)$$

sowie

$$\operatorname{Im} f(x+\mathrm{i}y) = v(x+\mathrm{i}y) = \int_{\mathbb{R}} \frac{y}{(x-t)^2+y^2}\, d\mu(t).$$

Mit $g := f - \tilde{f}$ ist also $\operatorname{Im} g = o$. Aus den Cauchy-Riemannschen Differentialgleichungen folgt deshalb $\operatorname{Re} g = k \in \mathbb{R}$ und aus

$$|g(z)| \leq 2\,\frac{p}{\operatorname{Im} z} \underset{y\to\infty}{\longrightarrow} 0$$

schließlich $g = o$.

6. Als letztes zeigen wir, daß es höchstens ein beschränktes positives Maß μ mit

$$\forall\, z \in \mathbb{H} \qquad f(z) = \int_{\mathbb{R}} \frac{d\mu(t)}{t-z}$$

gibt. Es sei nämlich

$$\forall\, z \in \mathbb{C}\setminus\mathbb{R} \qquad \tilde{f}(z) := \int_{\mathbb{R}} \frac{d\mu(t)}{t-z}.$$

Dann ist aufgrund der Stieltjesschen Umkehrformel das Maß μ durch $\tilde{f}$ eindeutig festgelegt. Weil μ positiv ist, gilt $\tilde{f}(z) = \overline{f(\bar{z})}$ für alle $z \in \mathbb{C}$ mit $\operatorname{Im} z < 0$. Deshalb ist $\tilde{f}$ bereits durch f festgelegt und damit μ nach Satz 9.2.1.2. □

Die soeben bewiesenen Resultate enthalten bereits den Spektralsatz. Zur Formulierung verwenden wir noch die lineare Abbildung

$$\begin{aligned} \hat{P}(f) \;:\; \mathcal{D}\big(\hat{P}(f)\big) \subset \mathcal{H} &\longrightarrow \mathcal{H}, \\ x &\longmapsto \int_{\mathbb{R}} f(\lambda)\, dP(\lambda)x \end{aligned}$$

mit

$$\mathcal{D}\big(\hat{P}(f)\big) := \left\{ x \in \mathcal{H} \;\middle|\; f \in \mathcal{L}^2_{\mu_x} \right\}.$$

Dabei seien $P(\lambda)$ wieder eine Spektralschar, μ_x das von $\big(P(\lambda)x, x\big)$ erzeugte Maß und $f : \mathbb{R} \to \mathbb{C}$ eine μ_x-meßbare Funktion. Durch Polarisieren folgt

$$\forall\, x \in \mathcal{D}\big(\hat{P}(f)\big) \quad \forall\, y \in \mathcal{H} \qquad \big(\hat{P}(f)x, y\big) = \int_{\mathbb{R}} f(\lambda)\, d\big(P(\lambda)x, y\big).$$

Wir zeigen nun den

Spektralsatz für selbstadjungierte Abbildungen: *Es sei $A : \mathcal{D}(A) \subset \mathcal{H} \to \mathcal{H}$ eine selbstadjungierte Abbildung. Dann gibt es genau eine Spektralschar P mit*

$$A = \hat{P}(id),$$

und es gilt die Stonesche Formel

$$\forall\, x, y \in \mathcal{H} \quad \forall\, t \in \mathbb{R} \qquad \big(P(t)x, y\big) = \lim_{\delta\downarrow 0} \lim_{\varepsilon\downarrow 0} \frac{1}{2\pi\mathrm{i}} \int_{-\infty}^{t+\delta} \Big(\Big(\big[A - (s+\mathrm{i}\varepsilon)\big]^{-1} - \big[A - (s-\mathrm{i}\varepsilon)\big]^{-1} \Big) x, y \Big)\, ds.$$

Wir skizzieren den Beweis:

1. Es sei P eine beliebige Spektralschar. Dann gilt

$$\Big(A = \hat{P}(id)\Big) \Longleftrightarrow \Big(\forall\, z \in \mathbb{C}\backslash\mathbb{R} \quad (A - z)^{-1} = \hat{P}\big((id - z)^{-1}\big)\Big).$$

Es ist nämlich

$$\hat{P}(id - z) = \hat{P}(id) - z\, id,$$

sowie

$$\hat{P}(id - z) \circ \hat{P}\big((id - z)^{-1}\big) = id_{\mathcal{D}\left(\hat{P}((id-z)^{-1})\right)} = id_{\mathcal{H}}$$

und

$$\hat{P}\big((id - z)^{-1}\big) \circ \hat{P}(id - z) = id_{\mathcal{D}\left(\hat{P}(id-z)\right)} = id_{\mathcal{D}\left(\hat{P}(id)\right)}.$$

Es folgt also

$$\hat{P}\big((id - z)^{-1}\big) = \Big(\hat{P}(id - z)\Big)^{-1} = \Big(\hat{P}(id) - z\, id\Big)^{-1}.$$

2. Es sei nun $A := \hat{P}(id)$. Dann gilt für alle $x, y \in \mathcal{H}$ und $z \in \mathbb{C}\backslash\mathbb{R}$

$$\big((A - z\, id)^{-1}x, y\big) = \big(\hat{P}\big((id - z)^{-1}\big)x, y\big) = \int_{\mathbb{R}} \frac{1}{t - z}\, d\mu_{x,y}(t) =: f(z).$$

Deshalb folgt aus der Stieltjesschen Umkehrformel

$$\begin{aligned}(P(t)x, y) = \mu_{x,y}\big((-\infty, t]\big) &= \lim_{\delta \downarrow 0} \lim_{\varepsilon \downarrow 0} \frac{1}{2\pi \mathrm{i}} \int_{-\infty}^{t+\delta} \Big(f(s + \mathrm{i}\varepsilon) - f(s - \mathrm{i}\varepsilon)\Big)\, ds = \\ &= \lim_{\delta \downarrow 0} \lim_{\varepsilon \downarrow 0} \frac{1}{2\pi \mathrm{i}} \int_{-\infty}^{t+s} \left(\Big(\big[A - (s + \mathrm{i}\varepsilon)id\big]^{-1} - \big[A - (s - \mathrm{i}\varepsilon)id\big]^{-1}\Big)x, y\right) ds.\end{aligned}$$

Das ist die *Stonesche Formel*, und P ist durch A eindeutig festgelegt.

3. Als letztes zeigen wir die Existenz der Spektralschar $P(\lambda)$. Es seien wieder $x \in \mathcal{H}$ und $z \in \mathbb{H}$. Dann ist

$$f(z) := \big((A - z)^{-1}x, x\big)$$

in $\mathbb{H}$ holomorph mit

$$|f(z)|\, \mathrm{Im}\, z \leq \|x\|^2.$$

Aus $(A - z)y = x$ folgt nämlich

$$-\mathrm{Im}\, z\, \|y\|^2 = \mathrm{Im}\,(x, y),$$

oder

$$\|y\| \leq \|x\| / \mathrm{Im}\, z,$$

also

$$\big|\big((A - z)^{-1}x, x\big)\big| = \big|(y, x)\big| \leq \|x\|^2 / \mathrm{Im}\, z.$$

Es ist auch

$$\begin{aligned}\mathrm{Im}\, f(z) = \mathrm{Im}\Big\{\big((A - z)^{-1}x, A(A - z)^{-1}x\big) - \bar{z}\big((A - z)^{-1}x, (A - z)^{-1}x\big)\Big\} &= \\ = -\mathrm{Im}\, \bar{z}\, \big\|(A - z)^{-1}x\big\|^2 = \mathrm{Im}\, z\, \big\|(A - z)^{-1}x\big\|^2 &\geq 0.\end{aligned}$$

Nach Satz 9.2.4 gibt es deshalb genau ein beschränktes positives Maß μ_x mit

$$\forall\, z \in \mathbb{H} \qquad \big((A - z)^{-1}x, x\big) = \int_{\mathbb{R}} \frac{d\mu_x(t)}{t - z}$$

und $\mu_x(\mathbb{R}) \leq \|x\|^2$. Diese Darstellung gilt sogar für alle $z \in \mathbb{C}\backslash\mathbb{R}$. Es sei nämlich $\mathrm{Im}\, z < 0$. Dann ist

$$\big((A - z)^{-1}x, x\big) = \big(x, (A - \bar{z})^{-1}x\big) = \overline{\big((A - \bar{z})^{-1}x, x\big)} = \overline{\int_{\mathbb{R}} \frac{d\mu_x(t)}{t - \bar{z}}} = \int_{\mathbb{R}} \frac{d\mu_x(t)}{t - z}.$$

Für beliebiges $x, y \in \mathcal{H}$ definieren wir nun durch Polarisieren das komplexe Maß $\mu_{x,y}$ und erhalten

$$\forall\, z \in \mathbb{C}\backslash\mathbb{R} \qquad \big((A-z)^{-1}x, y\big) = \int_{\mathbb{R}} \frac{d\mu_{x,y}(t)}{t-z}.$$

Dabei gilt für meßbare Mengen M

$$\mu_{x,x}(M) = \mu_x(M) \geq 0$$

und

$$\big|\mu_{x,y}(M)\big| \leq \sqrt{\mu_x(M)\cdot\mu_y(M)} \leq \|x\|\cdot\|y\|.$$

$\mu_{x,y}$ ist also eine beschränkte Sesquilinearform. Aufgrund des Rieszschen Darstellungssatzes existiert deshalb zu $y \in \mathcal{H}$ genau ein $P(M)y$ mit

$$\forall\, x \in \mathcal{H} \qquad \mu_{x,y}(M) = \big(x, P(M)y\big)$$

und $\|P(M)y\| \leq \|y\|$. Wegen

$$\big(x, P(M)x\big) = \mu_x(M) \geq 0$$

ist $P(M)$ hermitesch. P ist die gesuchte Spektralschar, zum Nachweis ihrer Eigenschaften zeigen wir:

a) Für alle $x, y \in \mathcal{H}$, $z \in \mathbb{C}\backslash\mathbb{R}$ und $M \in \mathcal{B}$ gilt

$$\mu_{(A-z)^{-1}x,y}(M) = \int_{\mathbb{R}} \chi_M(t)\frac{1}{t-z}d\mu_{x,y}(t).$$

Durch

$$\mu(M) := \int_{\mathbb{R}} \chi_M(t)\frac{1}{t-z}d\mu_{x,y}(t)$$

wird nämlich ein komplexes Maß mit

$$\int_{\mathbb{R}} \frac{d\mu(t)}{t-w} = \int_{\mathbb{R}} \frac{1}{t-w}\cdot\frac{1}{t-z}d\mu_{x,y}(t)$$

für $w \in \mathbb{C}\backslash\mathbb{R}$ definiert. Nun folgt aus der Resolventenformel

$$\int_{\mathbb{R}} \frac{1}{t-w}d\mu_{(A-z)^{-1}x,y}(t) = \big((A-w)^{-1}(A-z)^{-1}x, y\big) = \Big(\frac{1}{w-z}\big[(A-w)^{-1}-(A-z)^{-1}\big]x, y\Big) =$$
$$= \frac{1}{w-z}\Big[\int_{\mathbb{R}} \frac{d\mu_{x,y}(t)}{t-w} - \int_{\mathbb{R}} \frac{d\mu_{x,y}(t)}{t-z}\Big] = \int_{\mathbb{R}} \frac{d\mu_{x,y}(t)}{(t-w)(t-z)}.$$

Daraus erhält man die Behauptung.

b) Es ist $P(\mathbb{R}) = id$: Für alle $x \in \mathcal{D}(A)$, $y \in \mathcal{H}$ ist nämlich

$$\big(x, P(\mathbb{R})y\big) = \mu_{x,y}(\mathbb{R}) = \mu_{(A-z)^{-1}(A-z)x,y}(\mathbb{R}) =$$
$$= \int_{\mathbb{R}} \frac{1}{t-z}d\mu_{(A-z)x,y}(t) = \big((A-z)^{-1}(A-z)y, x\big) = (x, y).$$

Nun ist $\mathcal{D}(A)$ in $\mathcal{H}$ dicht. Daraus folgt die Behauptung und $\mu_x(\mathbb{R}) = \|x\|^2$.

c) Für alle $x, y \in \mathcal{H}$ und $M, N \in \mathcal{B}$ gilt

$$\mu_{x,P(M)y}(N) = \mu_{x,y}(M \cap N).$$

Durch

$$\mu(N) := \mu_{x,x}(M \cap N)$$

wird nämlich für festes M ein komplexes Borelmaß definiert mit

$$\int_{\mathbb{R}} \frac{d\mu(t)}{t-z} = \int_{\mathbb{R}} \chi_M(t)\frac{1}{t-z}d\mu_{x,y}(t)$$

für $z \in \mathbb{C}\backslash\mathbb{R}$. Es folgt

$$\int_{\mathbb{R}} \frac{1}{t-z} d\mu_{x,P(M)y}(t) = \big((A-z)^{-1}x, P(M)y\big) = \mu_{(A-z)^{-1}x,y}(M) =$$
$$= \int_{\mathbb{R}} \chi_M(t) \frac{1}{t-z} d\mu_{x,y}(t) = \int_{\mathbb{R}} \frac{1}{t-z} d\mu(t),$$

also

$$\forall\, N \in \mathcal{B} \qquad \mu(N) = \mu_{x,y}(M \cap N) = \mu_{x,P(M)y}(N).$$

d) Es folgt

$$\forall\, M, N \in \mathcal{B} \qquad P(M \cap N) = P(N) \circ P(M),$$

denn es ist

$$\big(x, P(N)P(M)y\big) = \mu_{x,P(M)y}(N) = \mu_{x,y}(M \cap N) = \big(x, P(M \cap N)y\big).$$

e) Damit besitzt

$$\forall\, t \in \mathbb{R} \qquad P(t) := P\big((-\infty, t]\big)$$

alle Eigenschaften einer Spektralschar. Insbesondere gilt

$$\|P(t)x\|^2 = \mu_x\big((-\infty, t]\big), \quad \|P(t)x - P(s)x\|^2 = \mu_x\big((s-t]\big), \quad \|x - P(s)x\|^2 = \mu_x\big((s, \infty]\big),$$

sowie

$$\big((A-z)^{-1}x, y\big) = \int_{\mathbb{R}} \frac{1}{t-z} d\mu_{x,y}(t).$$

Aus der ersten Aussage folgt dann $A = \hat{P}(id)$, also

$$(Ax, y) = \int_{\mathbb{R}} t\, d\mu_{x,y}(t)$$

für $x \in \mathcal{D}(A)$ und $y \in \mathcal{H}$. □

9.3 Folgerungen

In diesem Abschnitt wollen wir einige Folgerungen aus dem soeben bewiesenen Spektralsatz skizzieren. Es seien stets $\mathcal{H}$ ein Hilbertraum,

$$A \;:\; \mathcal{D}(A) \subset \mathcal{H} \longrightarrow \mathcal{H}$$

eine selbstadjungierte Abbildung und $P(\lambda)$ die zu A gehörende Spektralschar. Dann wissen wir bereits $\sigma(A) \subset \mathbb{R}$ und $R\sigma(A) = \emptyset$. Es sei

$$P\big(\{\lambda\}\big) := P(\lambda) - P(\lambda - 0).$$

Dann ist $P\big(\{\lambda\}\big)$ ein orthogonaler Projektor auf $\mathcal{N}(A - \lambda\, id)$. Es gilt nämlich

$$x \in \mathcal{N}(A-\lambda) \iff 0 = \int_{\mathbb{R}} (t-\lambda)^2\, d\big\|P(t)x\big\|^2 \iff x = P\big(\{\lambda\}\big)x.$$

Wir zeigen nun

Satz 9.3.1: *Für $\lambda \in \mathbb{R}$ gilt:*

1. *Es ist $\lambda \in P\sigma(A) \iff P\big(\{\lambda\}\big) \neq O$.*
2. *Es ist $\lambda \in \rho(A) \iff P$ ist in einer Umgebung $U(\lambda)$ konstant.*

Beweis: Wir müssen nur noch die zweite Aussage zeigen.

$\Rightarrow$: Es sei $\lambda \in \rho(A)$. Dann gilt

$$\exists\, p > 0 \quad \forall\, x \in \mathcal{D}(A) \qquad \|(A-\lambda)x\| \geq p\, \|x\|.$$

Wir wählen $c \in \mathbb{R}^+$ mit $0 < c < p$. Dann ist

$$P\big((\lambda - c,\, \lambda + c]\big) = O.$$

Es sei nämlich $x = P\big((\lambda - c,\, \lambda + c]\big)x \neq o$. Dann ist $x \in \mathcal{D}(A)$, und es folgt der Widerspruch

$$\big\|(A-\lambda)x\big\|^2 = \int_{\mathbb{R}} \big|(t-\lambda)\chi_{(\lambda-c,\lambda+c)}\big|^2 \, d\mu_x(t) \leq c^2 \, \|x\|^2.$$

$\Leftarrow$: Es seien P in $(\lambda - c,\, \lambda + c)$ konstant und

$$f(t) := \begin{cases} 0 & \text{für } t \in (\lambda - c, \lambda + c] \\ \frac{1}{t-\lambda} & \text{für } t \notin (\lambda - c, \lambda + c]. \end{cases}$$

Dann ist $\mathcal{D}\big(\hat{P}(f)\big) = \mathcal{H}$ und

$$\big\|\hat{P}(f)\big\| \leq \tfrac{1}{c}.$$

Aus $(id - \lambda)f = \chi_{\mathbb{R}\setminus(\lambda-c,\lambda+c]}$ folgt

$$\hat{P}\big((id-\lambda)f\big) = id - P\big((\lambda - c, \lambda + c]\big) = id$$

sowie

$$\mathcal{D}\big(\hat{P}(f) \circ (A-\lambda)\big) = \mathcal{D}(A) \quad \text{und} \quad \hat{P}(f) \circ (A-\lambda) \subset id$$

und

$$\mathcal{D}\big((A-\lambda) \circ \hat{P}(f)\big) = \mathcal{H} \quad \text{und} \quad (A-\lambda) \circ \hat{P}(f) = id,$$

also ist $\lambda \in \rho(A)$. □

Definition 9.3.2: *Man nennt*

$$\mathcal{H}_p := \overline{\Big\{ x \in \mathcal{H} \;\Big|\; x \text{ ist Eigenfunktion von } A \Big\}}$$

den unstetigen Teilraum *von* $\mathcal{H}$ *und*

$$\mathcal{H}_c := \mathcal{H}_p^{\perp}$$

den stetigen Teilraum *von* $\mathcal{H}$. *Im Englischen findet man auch* $\mathcal{H}_{pp}$ *statt* $\mathcal{H}_p$ *(pure point).*

Damit wird also

$$\mathcal{H} = \mathcal{H}_p \oplus \mathcal{H}_c$$

orthogonal zerlegt. Es gilt

Satz 9.3.3: *Für alle* $x \in \mathcal{H}$ *sind äquivalent:*

1. $x \in \mathcal{H}_p$.
2. *Es gibt eine abzählbare Menge* $M \subset \mathbb{R}$, $M \neq \emptyset$ *mit* $P(M)x = x$.
3. *Es gibt eine abzählbare Menge* $M \subset \mathbb{R}$, $M \neq \emptyset$ *mit* $\mu_x(\mathbb{R}\setminus M) = 0$.

Beweis: Es seien $\mathcal{N}_t := \mathcal{N}(A - t\,id)$ und $P\big(\{t\}\big)$ der Projektor auf $\mathcal{N}_t$. Dann ist

$$\mathcal{N}_t \perp \mathcal{N}_s \quad \text{für } t \neq s$$

und

$$\mathcal{H}_p = \overline{\bigoplus_{t \in \mathbb{R}} \mathcal{N}_t}.$$

Es sei Q der orthogonale Projektor auf $\mathcal{H}_p$,

$$Qx = \sum_{t \in \mathbb{R}} P\big(\{t\}\big)x.$$

In diese Summe gehen nämlich bei festem x höchstens abzählbar unendlich viele t ein. Man vergleiche Folgerung 1.4.8. Daraus ergeben sich die Behauptungen. □

Analog gilt

Satz 9.3.4: *Für alle $x \in \mathcal{H}$ sind äquivalent:*

1. $x \in \mathcal{H}_c$.
2. *Es ist* $\mu_x(\{t\}) = 0$ *für alle* $t \in \mathbb{R}$.
3. *Die Abbildung* $\mathbb{R} \ni t \mapsto P(t)x \in \mathcal{H}$ *ist stetig.*

Beweis:

1. $\Rightarrow$ 2.: Es seien $x \in \mathcal{H}_c$ und $t \in \mathbb{R}$. Dann ist $P(\{t\})x \in \mathcal{H}_p$ und

$$\mu_x(\{t\}) = \big\|P(\{t\})x\big\|^2 = \big(P(\{t\})x, x\big) = 0.$$

2. $\Rightarrow$ 1.: Es seien $x \in \mathcal{H}$ mit $\mu_x(\{t\}) = 0$ für alle $t \in \mathbb{R}$ und $y \in \mathcal{H}_p$. Dann gibt es eine abzählbare Menge $M \subset \mathbb{R}$ mit $P(M)y = y$, und es folgt

$$P(M)x = \sum_{t \in M} P(\{t\})x = o,$$

also

$$\forall\, y \in \mathcal{H}_p \qquad (x, y) = \big(x, P(M)y\big) = \big(P(M)x, y\big) = 0.$$

2. $\Leftrightarrow$ 3.: Es seien $x \in \mathcal{H}$, $t \in \mathbb{R}$ und $\delta > 0$. Dann folgt aus

$$\big\|P(t)x - P(t-\delta)x\big\|^2 = \mu_x\big((t-\delta, t]\big)$$

die Aussage

$$\lim_{\delta \downarrow 0} \big\|P(t)x - P(t-\delta)x\big\|^2 = \mu_x(\{t\}). \qquad \square$$

Damit haben wir das Spektrum $\sigma(A)$ in seinen stetigen und nicht stetigen Anteil zerlegt. Um nun aber zu physikalisch interessanten Anwendungen zu gelangen, zum Beispiel in der Streutheorie, muß man das stetige Spektrum weiter unterteilen. Zunächst sei an einige Begriffe aus der Maßtheorie erinnert:

Definition 9.3.5: *Es seien μ ein beschränktes positives Maß und λ das Lebesguesche Maß.*

(1) μ *heißt* absolut stetig *bezüglich* λ, $\mu \ll \lambda$, *wenn eine der folgenden äquivalenten Bedingungen erfüllt ist:*

1. $\forall\, M \in \mathcal{B} \quad \lambda(M) = 0 \implies \mu(M) = 0.$
2. $\forall\, \varepsilon > 0 \quad \exists\, \delta > 0 \quad \forall\, M \in \mathcal{B} \quad \lambda(M) < \delta \implies \mu(M) < \varepsilon.$
3. *Es gibt eine Lebesgueintegrable Borelfunktion* $h \in \mathcal{L}^1(\mathbb{R})$ *mit*

$$\forall\, M \in \mathcal{B} \qquad \mu(M) = \int_{\mathbb{R}} \chi_M h \, d\lambda.$$

4. *Die Abbildung*

$$\mathbb{R} \ni t \longmapsto \mu\big((-\infty, t]\big) \in \mathbb{R}$$

ist absolut stetig.

(2) μ *heißt* singulär stetig *bezüglich* λ, $\mu \perp \lambda$, *wenn*

$$\exists\, N \in \mathcal{B},\ \lambda(N) = 0, \qquad \mu(\mathbb{R} \backslash N) = 0$$

gilt. Das Maß μ „lebt“ also auf einer Lebesgue-Nullmenge.

Erläuterung: Aus der Lebesgueschen Integrationstheorie ist bekannt:

1. Es sei μ monoton. Dann ist μ fast überall differenzierbar mit $\mu' \in \mathcal{L}^1$.
2. Es sei $f \in \mathcal{L}^1$. Dann gilt fast überall

$$\frac{d}{dx} \int_0^x f(t)\ dt = f(x).$$

3. μ ist absolut stetig $:\Longleftrightarrow$

$$\exists\, f \in \mathcal{L}^1 \qquad \mu(x) = \mu(0) + \int_0^x f(t)\, dt$$

oder äquivalent

$$\forall\, \varepsilon > 0 \quad \exists\, \delta > 0 \quad \forall\, (x_1, y_1), \ldots, (x_n, y_n) \text{ disjunkt und mit } \sum_{i=1}^{n} |y_i - x_i| < \delta$$

gilt

$$\sum_{i=1}^{n} \big|\mu(y_i) - \mu(x_i)\big| < \varepsilon.$$

Eine absolut stetige Funktion f ist von beschränkter Variation. Sie ist fast überall differenzierbar mit $f' \in \mathcal{L}^1$, und aus $f' = o$ fast überall folgt $f = const$.

Es sei μ monoton und stetig. Dann existiert ein $f \in \mathcal{L}^1_{loc}(\mathbb{R})$ mit $\mu' = f$ fast überall. Es seien

$$a(x) := \int_a^x f(t)\, dt$$

und

$$s(x) := \mu(x) - a(x).$$

Dann ist also

$$\mu(x) = a(x) + s(x)$$

mit einer *absolut stetigen* Funktion a und einer *singulär stetigen* Funktion s. Es ist also fast überall $s' = o$. Eine solche Zerlegung positiver Maße spielt im folgenden eine große Rolle. Ein positives singulär stetiges Maß, das auf der Cantormenge – einer Lebesgue-Nullmenge – lebt, läßt sich leicht angeben.

Wir bezeichnen nun weiter

Definition 9.3.6: *Es seien*

$$\mathcal{H}_s := \Big\{ x \in \mathcal{H} \;\Big|\; \exists\, N \in \mathcal{B},\, \lambda(N) = 0, \quad P(\mathbb{R}\backslash N)x = o \Big\},$$

sowie

$$\mathcal{H}_{ac} := \mathcal{H}_s^{\perp} \quad \textit{und} \quad \mathcal{H}_{sc} := \mathcal{H}_s \cap \mathcal{H}_c.$$

$\mathcal{H}_s$ *ist der* singuläre Teilraum *von* $\mathcal{H}$, $\mathcal{H}_{ac}$ *ist der* absolut stetige Teilraum *von* $\mathcal{H}$, *und* $\mathcal{H}_{sc}$ *ist der* singulär stetige Teilraum *von* $\mathcal{H}$.

Es gilt also jeweils orthogonal

$$\mathcal{H} = \underbrace{\mathcal{H}_p \oplus \mathcal{H}_{sc}}_{=\mathcal{H}_s} \oplus \mathcal{H}_{ac} = \mathcal{H}_p \oplus \underbrace{\mathcal{H}_{sc} \oplus \mathcal{H}_{ac}}_{=\mathcal{H}_c}$$

und insbesondere

Satz 9.3.7:

1. $\mathcal{H}_s$ *ist abgeschlossen, und es gilt* $\mathcal{H} = \mathcal{H}_{ac} \oplus \mathcal{H}_s$.
2. *Für alle* $x \in \mathcal{H}$ *sind*

$$x \in \mathcal{H}_s \quad \textit{und} \quad \mu_x \perp \lambda$$

äquivalent.

3. *Für alle* $x \in \mathcal{H}$ *sind die Aussagen*

$$x \in \mathcal{H}_{ac} \textit{ mit } \mu_x \ll \lambda \quad \textit{und} \quad \mathbb{R} \ni t \mapsto \|P(t)x\|^2 \textit{ ist absolut stetig}$$

äquivalent.

4. *Es seien* $x = x_{ac} + x_s$ *mit* $x_{ac} \in \mathcal{H}_{ac}$, $x_s \in \mathcal{H}_s$. *Dann gilt*

$$\mu_x = \mu_{x_{ac}} + \mu_{x_s}, \quad \mu_{x_{ac}} \ll \lambda \quad \textit{und} \quad \mu_{x_s} \perp \lambda.$$

Das stetige Spektrum der meisten in der Physik auftretenden Operatoren ist absolut stetig. Es gibt jedoch Beispiele stark oszillierender Potentiale, die auf singulär stetige Spektren führen.

Beispiel 9.3.8: *Es seien* $\mathcal{H} := \mathcal{L}^2(\mathbb{R})$ *und wie in Beispiel 9.1.6*

$$\begin{aligned} A \;:\; \mathcal{H}_1(\mathbb{R}) \subset \mathcal{H} &\longrightarrow \mathcal{H}, \\ x &\longmapsto -\mathrm{i}\, x'. \end{aligned}$$

Die Abbildung A ist selbstadjungiert. Aus $Au = \lambda u$ folgt $u(x) = \mathrm{e}^{\mathrm{i}\lambda x}$. Wegen $u \notin \mathcal{H}$ wird u auch *verallgemeinerte Eigenfunktion* genannt, und es ist $\sigma(A) = C\sigma(A)$.

Als erstes zeigen wir

$$\sigma(A) = C\sigma(A) = \mathbb{R}.$$

Es seien nämlich $\lambda \in \mathbb{R}, \sigma > 0$ und

$$u_\sigma(t) := \mathrm{e}^{\mathrm{i}\lambda t}\mathrm{e}^{-\sigma|t|} \in \mathcal{D}(A).$$

Dann ist

$$(A - \lambda\, id)\, u_\sigma(t) = f_\sigma(t)$$

mit

$$f_\sigma(t) := \begin{cases} \mathrm{i}\,\sigma u_\sigma(t) & \text{für } t > 0, \\ -\mathrm{i}\,\sigma u_\sigma(t) & \text{für } t < 0, \end{cases}$$

und es folgt für $\sigma \to 0$

$$\|f_\sigma\|^2 = \sigma \to 0 \quad \text{und} \quad \|u_\sigma\|^2 = \frac{1}{\sigma} \to \infty.$$

Damit ist $\lambda \in C\sigma(A)$.

Als nächstes konstruieren wir die Spektralschar $P(\lambda)$ von A: Es seien $z \in \mathbb{C}\backslash\mathbb{R}$, $\varphi \in \mathring{\mathcal{C}}_\infty(\mathbb{R})$ und

$$v_1(r) := \mathrm{i}\,\mathrm{e}^{\mathrm{i}zr} \int_{-\infty}^{r} \varphi(t)\,\mathrm{e}^{-\mathrm{i}zt} dt \quad \text{für } \operatorname{Im} z > 0,$$

$$v_2(r) := -\mathrm{i}\,\mathrm{e}^{\mathrm{i}zr} \int_{r}^{\infty} \varphi(t)\,\mathrm{e}^{-\mathrm{i}zt} dt \quad \text{für } \operatorname{Im} z < 0.$$

Dann ist für $j = 1, 2$

$$Av_j - zv_j = \varphi,$$

also

$$\big((A - z)^{-1}\varphi\big)(r) = \begin{cases} \mathrm{i}\,\mathrm{e}^{\mathrm{i}zr} \int_{-\infty}^{r} \varphi(t)\,\mathrm{e}^{-\mathrm{i}zt} dt & \text{für } \operatorname{Im} z > 0 \\ -\mathrm{i}\,\mathrm{e}^{\mathrm{i}zr} \int_{r}^{\infty} \varphi(t)\,\mathrm{e}^{-\mathrm{i}zt} dt & \text{für } \operatorname{Im} z < 0. \end{cases}$$

Daraus folgt für $x, y \in \mathring{\mathcal{C}}_\infty(\mathbb{R})$, $\varepsilon > 0$ und $s \in \mathbb{R}$

$$\begin{aligned} \big(\big[\big(A - (s + \mathrm{i}\varepsilon)\big)^{-1} - \big(A - (s - \mathrm{i}\varepsilon)\big)^{-1}\big]x\big)(r) = \\ = \mathrm{i}\,\mathrm{e}^{\mathrm{i}sr} \int_{-\infty}^{r} x(t)\,\mathrm{e}^{-\mathrm{i}st}\mathrm{e}^{\varepsilon(t-r)}\, dt + \mathrm{i}\,\mathrm{e}^{\mathrm{i}sr} \int_{r}^{\infty} x(t)\,\mathrm{e}^{-\mathrm{i}st}\mathrm{e}^{\varepsilon(r-t)}\, dt = \mathrm{i}\,\mathrm{e}^{\mathrm{i}sr} \int_{\mathbb{R}} x(t)\,\mathrm{e}^{-\mathrm{i}st}\mathrm{e}^{-\varepsilon|r-t|}\, dt \end{aligned}$$

oder

$$\big(\big[\,\cdots\,\big]x, y\big)(s) = \mathrm{i} \iint_{\mathbb{R}\times\mathbb{R}} \underbrace{x(t)\,\mathrm{e}^{-\mathrm{i}st}\mathrm{e}^{-\varepsilon|r-t|}\,\overline{y(r)\,\mathrm{e}^{-\mathrm{i}sr}}}_{=:I(\varepsilon,s,t,r)}\, dt\, dr.$$

Nun gilt für $a \le b$, $(s, t, r) \in [a, b] \times \mathbb{R} \times \mathbb{R}$ und $\varepsilon \ge 0$

$$\big|I(\varepsilon, s, t, r)\big| \le |x(t)| \cdot |y(r)|$$

sowie

$$\lim_{\varepsilon \downarrow 0} I(\varepsilon, s, t, r) = x(t)\,\mathrm{e}^{-\mathrm{i}st}\overline{y(r)\,\mathrm{e}^{-\mathrm{i}sr}}.$$

Damit folgt aus den Sätzen von Lebesgue und Fubini für alle $x, y \in \mathring{\mathcal{C}}_\infty(\mathbb{R})$

$$\lim_{\varepsilon \downarrow 0} \int_a^b \big([\,\cdots\,]x, y\big)(s)\, ds = \mathrm{i} \int_a^b \Big(\int_{\mathbb{R}} x(t) \mathrm{e}^{-\mathrm{i}st} dt \Big) \overline{\Big(\int_{\mathbb{R}} y(t) \mathrm{e}^{-\mathrm{i}sr} dr \Big)} ds$$

oder

$$\Big(\big[P(\lambda) - P(\lambda_0)\big]x, y \Big) = \int_{\lambda_0}^{\lambda} \hat{x}(s)\, \overline{\hat{y}(s)}\, ds$$

mit

$$\begin{aligned} F_0 \;:\; & \mathring{\mathcal{C}}_\infty(\mathbb{R}) \longrightarrow \mathcal{C}(\mathbb{R}), \\ & v \longmapsto \hat{v}(s) := \big(F_0 v\big)(s) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} v(t)\, \mathrm{e}^{-\mathrm{i}st} dt. \end{aligned}$$

Aus dem Satz von der monotonen Konvergenz erhalten wir daher für $x = y$, $\lambda \to \infty$ und $\lambda_0 \to -\infty$

$$F_0 x \in \mathcal{H},$$

also

$$F_0 \;:\; \mathring{\mathcal{C}}_\infty(\mathbb{R}) \longrightarrow \mathcal{H}$$

mit

$$\|F_0 x\| = \|x\|$$

und

$$(x, y) = \big(F_0 x, F_0 y\big),$$

es gilt also die Parsevalsche Gleichung. Es sei nun

$$\begin{aligned} G_0 \;:\; & \mathring{\mathcal{C}}_\infty(\mathbb{R}) \longrightarrow \mathcal{C}(\mathbb{R}), \\ & x \longmapsto \overline{F_0 \bar{x}}. \end{aligned}$$

Dann ist

$$\forall\, x, y \in \mathring{\mathcal{C}}_\infty(\mathbb{R}) \qquad (G_0 x, G_0 y) = (x, y),$$

und aus dem Satz von Fubini folgt

$$(F_0 x, y) = (x, G_0 y)\,,$$

also

$$\begin{aligned} F_0 \;:\; & \mathring{\mathcal{C}}_\infty(\mathbb{R}) \longrightarrow \mathcal{H}, \qquad \|F_0 x\| = \|x\|, \\ G_0 \;:\; & \mathring{\mathcal{C}}_\infty(\mathbb{R}) \longrightarrow \mathcal{H}, \qquad \|G_0 x\| = \|x\|. \end{aligned}$$

Es seien nun

$$F, G \;:\; \mathcal{H} \longrightarrow \mathcal{H}$$

die eindeutig definierten stetigen Fortsetzungen von F_0, G_0 auf $\mathcal{H}$. Es übertragen sich die Abbildungseigenschaften

$$Gx = \overline{F\bar{x}}, \qquad (Fx, Fy) = (Gx, Gy) = (x, y), \qquad (Fx, y) = (x, Gy)$$

sowie

$$\Big(\big(P(b) - P(a)\big)x, y \Big) = \int_a^b (Fx)\, \overline{(Fy)}$$

und insbesondere

$$\begin{aligned} (FGx, y) &= (Gx, Gy) = (x, y) \\ (GFx, y) &= (Fx, Fy) = (x, y). \end{aligned}$$

Damit ist $G = F^{-1}$, also

$$G = F^* = F^{-1}.$$

Die Abbildung F ist unitär, und mit

$$\big(P(\lambda)x\big)(s) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\lambda} \mathrm{e}^{\mathrm{i}st} \big(F(x)\big)(t)\, dt$$

gilt fast überall

$$x(s) = \big(P(\infty)x\big)(s) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}st} \big(F(x)\big)(t)\, dt := \mathcal{L}^2(\mathbb{R})\text{-}\lim_{r\to\infty} \frac{1}{\sqrt{2\pi}} \int_{-r}^{r} \mathrm{e}^{\mathrm{i}st} \big(F(x)\big)(t)\, dt. \qquad (*)$$

Wegen

$$\big(P(\lambda)x\,,x\big) = \int_{-\infty}^{\lambda} \big|\hat{x}(s)\big|^2\, ds$$

ist das Spektrum absolut stetig. Es folgt fast überall

$$\big(Ax\big)(s) = \Big(\int_{-\infty}^{\infty} \lambda\, dP(\lambda)x\Big)(s) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t\, \mathrm{e}^{\mathrm{i}st} \big(F(x)\big)(t)\, dt. \qquad (**)$$

Damit ist auch

$$\mathcal{H}_1(\mathbb{R}) = \Big\{ x \in \mathcal{H} \;\Big|\; \int_{\mathbb{R}} s^2 \big|\big(F(x)\big)(s)\big|^2\, ds < \infty \Big\}$$

und fast überall

$$\big(FAx\big)(t) = t\big(Fx\big)(t).$$

Die *Fouriertransformation* liefert also die Entwicklung einer Funktion $x \in \mathcal{H}$ nach den verallgemeinerten Eigenfunktionen von A. Damit haben wir eine zweite Herleitung der Fouriertransformation gefunden. Man vergleiche mit §6.4.4. □

Beispiel 9.3.9: *Es seien $G \subset \mathbb{R}^3$ ein* Außengebiet, $\mathcal{H} := \mathcal{L}^2(G)$, $u^0, u^1 \in \mathcal{H}$,

$$\mathcal{D}(A) := \Big\{ u \in \mathring{\mathcal{H}}_1(G) \;\Big|\; \Delta u \in \mathcal{H} \Big\} \quad \textit{und} \quad Au := -\Delta u.$$

Wir betrachten die Anfangswertaufgabe zur Wellengleichung

$$u_{tt} + Au = o \quad \textit{mit } u(0) = u^0 \;\textit{ und } u_t(0) = u^1.$$

Dabei ist ein Außengebiet ein Gebiet mit beschränktem Komplement. Wir fragen zunächst nach *Lösungen mit endlicher Energie*

$$E(t) := \big\|u_t(t)\big\|^2 + \big\|A^{1/2}u(t)\big\|^2.$$

Dabei wird $A^{1/2}$ über die Spektraldarstellung von A definiert, und man nennt $u \in \mathcal{C}\big(\mathbb{R}_0^+, \mathcal{D}(A^{1/2})\big) \cap \mathcal{C}_1\big(\mathbb{R}_0^+, \mathcal{H}\big)$ eine Lösung der Wellengleichung mit endlicher Energie, wenn für $u^0 \in \mathcal{D}(A^{1/2})$ und $u^1 \in \mathcal{H}$

$$\forall\, \varphi \in \mathring{\mathcal{C}}(\mathbb{R}, \mathcal{D}(A)) \cap \mathcal{C}_2(\mathbb{R}, \mathcal{H}) \qquad 0 = \int_{\mathbb{R}^+ \times G} u\, \overline{(\varphi_{tt} + A\varphi)} + \big(u^0, \varphi(0,\cdot)\big) - \big(u^1, \varphi(0,\cdot)\big)$$

gilt. Man zeigt $\mathcal{D}(A^{1/2}) = \mathring{\mathcal{H}}_1(G)$ und erhält die eindeutige Lösung der Aufgabe in der Form

$$u(t) = \int_0^{\infty} \cos\sqrt{\lambda}t\, dP(\lambda)u^0 + \int_0^{\infty} \frac{\sin\sqrt{\lambda}t}{\sqrt{\lambda}}\, dP(\lambda)u^1.$$

Es ist

$$E(t) = E(0) = \big\|u^1\big\|^2 + \big\|A^{1/2}u^0\big\|^2.$$

Auf Einzelheiten soll hier verzichtet werden. Es sei nun $K \subset G$ mit $K \Subset \mathbb{R}^3$. Dann gilt *das Prinzip vom lokalen Abklingen der Energie*, nämlich

$$\lim_{t\to\infty} \Big\{ \big\|u_t(t)\big\|_{\mathcal{L}^2(K)} + \big\|A^{1/2}u(t)\big\|_{\mathcal{L}^2(K)} \Big\} = 0.$$

Man erwartet deshalb, daß sich die Lösung u asymptotisch wie eine freie Welle verhält. Genauere Untersuchungen darüber erfolgen in der *Streutheorie*. Wir skizzieren hier den Beweis des Prinzips vom lokalen Abklingen der Energie und beschränken uns dabei auf die Diskussion des Terms

$$w(t) := \cos A^{1/2}t \circ f = \int_0^{\infty} \cos\sqrt{\lambda}t\, dP(\lambda) f$$

für $f \in \mathcal{H}$.

1. Wir zeigen die schwache Konvergenz: Weil das Spektrum von A absolut stetig ist, folgt für alle $v \in \mathcal{H}$

$$F(\lambda) := \frac{d}{d\lambda}\big(P(\lambda)f, v\big) \in \mathcal{L}^1(\mathbb{R}^+),$$

also

$$\big(w(t), v\big) = \int_0^\infty \cos\sqrt{\lambda}t\, F(\lambda)\, d\lambda.$$

Nun erhält man aus dem Riemann-Lebesgueschen Lemma (durch partielles Integrieren)

$$\forall\, F \in \mathcal{C}_1(\mathbb{R}_0^+) \cap \mathcal{L}^1(\mathbb{R}^+) \qquad \big(w(t), v\big) \to 0.$$

Approximieren wir $F \in \mathcal{L}^1(\mathbb{R}^+)$ durch $g \in \mathcal{C}_1(\mathbb{R}_0^+) \cap \mathcal{L}^1(\mathbb{R}^+)$, dann folgt also für $t \to \infty$

$$\forall\, v \in \mathcal{H} \qquad \big(w(t), v\big) \to 0.$$

2. Wir zeigen die behauptete starke Konvergenz: Es sei zunächst $f \in \mathcal{D}(A^{1/2}) = \mathring{\mathcal{H}}_1(G)$. Dann ist

$$\big\|\cos A^{1/2}t \circ f\big\|_1 \le c\,\big\|\cos A^{1/2}t \circ f\big\|_{A^{1/2}} \le c\,\big\{\|f\| + \big\|A^{1/2}f\big\|\big\}.$$

Jede Folge $(\cos A^{1/2}t_n \circ f)_{n\in\mathbb{N}}$ mit $t_n \to \infty$ enthält deshalb aufgrund des Rellichschen Auswahlsatzes eine in $\mathcal{L}^2(K)$ konvergente Teilfolge. Wegen der bereits bewiesenen schwachen Konvergenz ist der Grenzwert Null. Mithin gilt in diesem Falle für jede Folge $t_n \to \infty$

$$\big\|w(t_n)\big\|_{\mathcal{L}^2(K)} = \big\|\cos A^{1/2}t_n \circ f\big\|_{\mathcal{L}^2(K)} \to 0.$$

Als letztes approximieren wir ein beliebiges $f \in \mathcal{H}$ durch $g \in \mathcal{D}(A^{1/2})$. □

10 Ergänzungen

Leider konnten in dieser Vorlesung bei weitem nicht alle interessanten Fragen der Funktionalanalysis aufgegriffen werden. In diesem letzten Kapitel sollen nur noch drei Ergänzungen gebracht werden.

10.1 Halbgruppen

Im letzten Kapitel haben wir durch den Spektralsatz eine schöne Darstellung selbstadjungierter Abbildungen gefunden. Nun sind aber bei weitem nicht alle linearen Abbildungen selbstadjungiert. Deshalb soll noch eine andere Methode zur Darstellung solcher Abbildungen und damit zum Lösen von Gleichungen angegeben werden.

Denken Sie an eine Gleichung der Form

$$u_t = Au \quad \text{mit } u(0) = u^0$$

in einem Banachraum $\mathcal{X}$. Dabei seien A eine lineare Abbildung

$$A \;:\; \mathcal{D}(A) \subset \mathcal{X} \longrightarrow \mathcal{X}$$

und $u \in \mathcal{C}(\mathbb{R}_0^+, \mathcal{D}(A)) \cap \mathcal{C}_1(\mathbb{R}_0^+, \mathcal{X})$ die *schwache Lösung* der Differentialgleichung mit

$$\forall\, \varphi \in \mathring{\mathcal{C}}\big(\mathbb{R}, \mathcal{D}(A^*)\big) \cap \mathcal{C}_1\big(\mathbb{R}, \mathcal{X}\big) \qquad 0 = \int_{\mathbb{R}+} \big(u, \varphi_t + A^*\varphi\big)(t)\, dt + (u^0, \varphi(0,\cdot)).$$

Im Falle $A = \mathrm{i}\, B$ mit selbstadjungiertem B erhalten wir die Lösung dieser Aufgabe durch

$$u(t) := \mathrm{e}^{At} u^0 = \mathrm{e}^{\mathrm{i}Bt} u^0 = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}\lambda t}\, dP(\lambda)\, u^0,$$

wobei $P(\lambda)$ die Spektralschar von B ist.

Es stellt sich nun die Frage, ob man den Lösungsoperator

$$\mathrm{e}^{At}$$

auch für allgemeinere A sinnvoll definieren kann. Es zeigt sich, daß das möglich ist, wenn man t auf $t \in \mathbb{R}_0^+$ beschränkt. Es soll also nur in die Zukunft und nicht in die Vergangenheit gerechnet werden – oder umgekehrt (denken Sie an Wärmeleitungspobleme). Wir wollen zeigen, daß dann *Halbgruppen*, siehe Definition 10.1.1, der Form e^{At} existieren.

Beginnen wir mit einem einfachen Beispiel: Es seien $\mathcal{X}$ ein Banachraum, $A \in \mathcal{CL}(\mathcal{X})$, und $z \in \mathbb{C}$. Dann können wir

$$\mathrm{e}^{Az} := \sum_{n=0}^{\infty} \frac{z^n}{n!}\, A^n \in \mathcal{CL}(\mathcal{X})$$

definieren. Die Reihe rechts konvergiert nämlich absolut in der Operatornorm und stellt mithin eine stetige Abbildung dar. Es ist sogar e^{Az} in z holomorph, und es gilt

$$\mathrm{e}^{A(z_1+z_2)} = \mathrm{e}^{Az_1}\, \mathrm{e}^{Az_2}$$

sowie

$$\frac{d}{dz}\, \mathrm{e}^{Az} = A\, \mathrm{e}^{Az} = \mathrm{e}^{Az}\, A.$$

Die Definition von e^{At} über eine Reihe wird viel schwieriger, wenn wir unbeschränkte A zulassen. Es seien also von nun an $\mathcal{X}$ ein Banachraum und

$$A \;:\; \mathcal{D}(A) \subset \mathcal{X} \longrightarrow \mathcal{X}$$

eine lineare Abbildung. Dann bildet im allgemeinen A nicht $\mathcal{D}(A)$ in sich ab. Bei der formalen Definition von e^{At} durch die Exponentialreihe erhält man also Schwierigkeiten mit dem Definitionsbereich;

$$\bigcap_{n=1}^{\infty} \mathcal{D}(A^n)$$

könnte sehr klein sein. Dieselbe Schwierigkeit tritt auf, wenn man an die Definition der Exponentialfunktion durch

$$\mathrm{e}^x = \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n,$$

also an

$$\mathrm{e}^{At} := \lim_{n\to\infty}\left(id+\frac{t}{n}A\right)^n$$

denkt. Erfolgversprechender ist es jedoch, die rechte Seite der Abschätzung

$$\forall\, n\in\mathbb{N}\quad \forall\, x,\, |x|<n, \qquad \left(1+\frac{x}{n}\right)^n \le \mathrm{e}^x \le \left(1-\frac{x}{n}\right)^{-n}$$

auszunutzen. Man erhält sie aus

$$1+x \le \mathrm{e}^x \le \frac{1}{1-x}.$$

Beide Seiten konvergieren bei festem x für $n\to\infty$ gegen e^x. Das legt die Definition

$$\mathrm{e}^{At} := \lim_{n\to\infty}\left(id-\frac{t}{n}A\right)^{-n} \qquad (*)$$

nahe. Dabei geht rechts die n-te Potenz der Resolvente von A an der Stelle n/t ein; es muß also $\mathbb{R}^+\subset\rho(A)$ sein. Im folgenden soll kurz gezeigt werden, daß sich auf diese Weise *Halbgruppen* e^{At} sinnvoll einführen lassen. Für die ausführliche Darstellung der Theorie sei aber auf die Literatur verwiesen, etwa auf das bereits zitierte Buch von *Kato [1966]* und vor allem auf das Buch von *Pazy [1983]*.

Eine hinreichende Bedingung für die Existenz des Limes in $(*)$ ist:

1. A ist eine abgeschlossene Abbildung mit in $\mathcal{X}$ dichtem Definitionsbereich $\mathcal{D}(A)$.
2. Es gilt $\mathbb{R}^+\subset\rho(A)$ und

$$\forall\,\lambda>0 \qquad \left\|(A-\lambda\, id)^{-1}\right\| \le \frac{1}{\lambda}.$$

Bevor wir das zeigen, holen wir etwas aus und beginnen mit

Definition 10.1.1: *Eine Schar* $\{H(t)\}_{t\ge 0}$, $H(t)\in\mathcal{CL}(\mathcal{X})$, *heißt* Halbgruppe, *wenn*

1. $H(0)=id$,
2. $H(s+t)=H(s)H(t)$ *für alle* $s,t\ge 0$ *ist.*

Definition 10.1.2: *Eine Halbgruppe* $\{H(t)\}$ *heißt*

1. gleichmäßig stetig, *wenn* $\lim\limits_{t\downarrow 0}\|H(t)-id\|=0$ *ist,*
2. stark stetig *oder* $\mathcal{C}_0$-Halbgruppe, *wenn* $\lim\limits_{t\downarrow 0}H(t)x=x$ *für alle* $x\in\mathcal{X}$ *gilt.*

Definition 10.1.3: *Es seien* $\{H(t)\}$ *eine Halbgruppe und*

$$A \;:\; \mathcal{D}(A)\subset\mathcal{X}\longrightarrow\mathcal{X}$$

eine lineare Abbildung mit

$$\mathcal{D}(A) := \left\{x\in\mathcal{X} \;\middle|\; \exists\, y\in\mathcal{X}\quad \lim_{t\downarrow 0}\frac{H(t)x-x}{t}=y\right\}$$

und $Ax := y$. *Dann heißt* A (infinitesimaler) Erzeuger *der Halbgruppe* $\{H(t)\}$.

10.1.1 Gleichmäßig stetige Halbgruppen

Als erstes behandeln wir gleichmäßig stetige Halbgruppen und zeigen

Lemma 10.1.4: *Die Halbgruppe* $\{H(t)\}$ *sei gleichmäßig stetig. Dann gilt*

$$\forall\, s,t\ge 0 \qquad \lim_{s\to t}\|H(s)-H(t)\|=0.$$

Beweis: Aus der gleichmäßigen Stetigkeit folgt

$$\forall\,\varepsilon>0 \quad \exists\,\delta>0 \quad \forall\,t\in[0,\delta) \qquad \|H(t)-id\|<\varepsilon$$

oder

$$\forall\,\varepsilon>0 \quad \exists\,\delta>0 \quad \forall\,t\in[0,\delta) \qquad \|H(t)\|<1+\varepsilon.$$

Für $T\in\mathbb{R}^+$ sei nun $t\in[0,T]$. Dann gibt es ein $n\in\mathbb{N}$ mit $t=n\delta+\tau$ und $\tau\in[0,\delta)$, also

$$\|H(t)\|\le\|H(\delta)\|^n\cdot\|H(\tau)\|<(1+\varepsilon)^{n+1}$$

oder

$$\forall\,T\in\mathbb{R}^+ \quad \exists\,C\in\mathbb{R}^+ \quad \forall\,t\in[0,T] \qquad \|H(t)\|\le C.$$

Mit $h>0$ folgt daraus

$$\|H(t+h)-H(t)\|\le\|H(t)\|\cdot\|H(h)-id\|\le C\,\|H(h)-id\|$$

und für $0<t\le T$, $0<h\le t$

$$\|H(t-h)-H(t)\|\le\|H(t-h)\|\cdot\|id-H(h)\|\le C\,\|H(h)-id\|. \qquad \square$$

Es folgt

Satz 10.1.5: *Eine lineare Abbildung A erzeugt genau dann eine gleichmäßig stetige Halbgruppe, wenn $A\in\mathcal{CL}(\mathcal{X})$ ist.*

Beweis:

$\Leftarrow$: Es seien $A\in\mathcal{CL}(\mathcal{X})$ und

$$H(t):=\mathrm{e}^{At}:=\sum_{n=0}^{\infty}\frac{(At)^n}{n!}.$$

Offenbar existiert $H(t)$, und es gelten $H(0)=id$,

$$H(s+t)=H(s)H(t)$$

sowie

$$\|H(t)-id\|\le t\cdot\|A\|\cdot\mathrm{e}^{t\|A\|}.$$

Mithin ist $\{H(t)\}$ gleichmäßig stetig, und es folgt

$$\left\|\frac{H(t)-id}{t}-A\right\|\le t\,\|A\|^2\,\mathrm{e}^{t\|A\|}\to 0.$$

$\Rightarrow$: Es sei $\{H(t)\}$ eine gleichmäßig stetige Halbgruppe. Wir wählen ein $\rho>0$ mit

$$\left\|id-\frac{1}{\rho}\int_0^\rho H(s)\,ds\right\|<1.$$

Dann ist

$$K:=\frac{1}{\rho}\int_0^\rho H(s)\,ds\in\mathcal{CL}(\mathcal{X})$$

invertierbar (Neumannsche Reihe), mithin ist auch

$$L:=\int_0^\rho H(s)\,ds\in\mathcal{CL}(\mathcal{X})$$

invertierbar mit $L^{-1}\in\mathcal{CL}(\mathcal{X})$. Nun gilt

$$\frac{1}{h}\big(H(h)-id\big)L=\frac{1}{h}\Big\{\int_h^{h+\rho}H(s)\,ds-\int_0^\rho H(s)\,ds\Big\}=\frac{1}{h}\Big\{\int_\rho^{\rho+h}H(s)\,ds-\int_0^h H(s)\,ds\Big\},$$

also für $h\to 0$ in der Operatornorm

$$\frac{1}{h}\big(H(h)-id\big)=\frac{1}{h}\Big\{\ \cdots\ \Big\}L^{-1}\to\Big\{H(\rho)-id\Big\}\,L^{-1}.$$

Mithin ist $A=\big\{H(\rho)-id\big\}L^{-1}\in\mathcal{CL}(\mathcal{X})$. $\square$

Satz 10.1.6: *Es seien $\{H(t)\}$ und $\{I(t)\}$ gleichmäßig stetige Halbgruppen, und es gelte bzgl. der Operatornorm*

$$A = \lim_{t \downarrow 0} \frac{H(t) - id}{t} = \lim_{t \downarrow 0} \frac{I(t) - id}{t}.$$

Dann ist $H(t) = I(t)$ für alle $t \geq 0$.

Beweis: Es sei $T > 0$ fest. Wir zeigen

$$\forall\, t \in [0,T] \qquad H(t) = I(t).$$

Wie wir bereits wissen, gibt es ein $C > 0$ mit

$$\forall\, s,t \in [0,T] \qquad \|H(t)\| \cdot \|I(s)\| \leq C,$$

und nach Voraussetzung gilt

$$\forall\, \varepsilon > 0 \quad \exists\, \delta > 0 \quad \forall\, 0 \leq h < \delta \qquad \frac{1}{h}\|H(h) - I(h)\| < \frac{\varepsilon}{TC}.$$

Zu $t \in [0,T]$ wählen wir nun ein $n \in \mathbb{N}$ mit $t/n < \delta$. Dann folgt

$$\|H(t) - I(t)\| \leq \sum_{i=0}^{n-1} \Big\| \underbrace{H\Big((n-i)\frac{t}{n}\Big)}_{=H\left((n-i-1)\frac{t}{n}\right)H\left(\frac{t}{n}\right)} I\Big(i\frac{t}{n}\Big) - H\Big((n-i-1)\frac{t}{n}\Big) \underbrace{I\Big((i+1)\frac{t}{n}\Big)}_{=I\left(i\frac{t}{n}\right)I\left(\frac{t}{n}\right)} \Big\| \leq$$

$$\leq \sum_{i=0}^{n-1} \Big\| H\Big((n-i-1)\frac{t}{n}\Big) \Big\| \cdot \Big\| H\Big(\frac{t}{n}\Big) - I\Big(\frac{t}{n}\Big) \Big\| \cdot \Big\| I\Big(i\frac{t}{n}\Big) \Big\| \leq n \cdot C \cdot \Big(\frac{t}{n} \cdot \frac{\varepsilon}{TC}\Big) \leq \varepsilon. \qquad \square$$

Folgerung 10.1.7: *Es sei $\{H(t)\}$ eine gleichmäßig stetige Halbgruppe. Dann gilt:*

1. *Es gibt genau ein $A \in \mathcal{CL}(\mathcal{X})$ mit $H(t) = \mathrm{e}^{At}$.*
2. *Es gibt ein $\omega \geq 0$ mit $\|H(t)\| \leq \mathrm{e}^{\omega t}$, nämlich $\omega := \|A\|$.*
3. *Die Abbildung $t \mapsto H(t)$ ist in der Operatornorm differenzierbar, und es gilt*

$$\frac{d}{dt} H(t) = AH(t) = H(t)A.$$

10.1.2 Stark stetige Halbgruppen

Nun kommen wir zum eigentlich interessanten Teil der Halbgruppentheorie und behandeln stark stetige Halbgruppen. Wir beginnen mit

Satz 10.1.8: *Es sei $\{H(t)\}$ eine $\mathcal{C}_0$-Halbgruppe. Dann gilt*

$$\exists\, \omega \geq 0 \quad \exists\, M \geq 1 \quad \forall\, t \in \mathbb{R}_0^+ \qquad \|H(t)\| \leq M\mathrm{e}^{\omega t}.$$

Beweis:

1. Wir zeigen

$$\exists\, \eta > 0 \quad \exists\, k \geq 0 \quad \forall\, t \in [0,\eta] \qquad \|H(t)\| \leq k.$$

Der Nachweis erfolgt indirekt: Es gebe eine Nullfolge (t_n) mit $\|H(t_n)\| > n$. Dann existiert ein $x \in \mathcal{X}$ mit

$$\big\|H(t_n)x\big\| \to \infty,$$

denn sonst wäre $H(t_n)$ wegen des Satzes von Banach-Steinhaus beschränkt. Aus der starken Stetigkeit folgt dann aber der Widerspruch

$$\big\|H(t_n)x\big\| \leq \big\|H(t_n)x - x\big\| + \big\|x\big\| \to \big\|x\big\|.$$

2. Es ist also $\|H(t)\| \leq k$ für alle $t \in [0,\eta]$, und wegen $H(0) = id$ ist $k \geq 1$. Es sei

$$\omega := \frac{\ln k}{\eta} \geq 0.$$

Dann gibt es zu $t \in \mathbb{R}_0^+$ ein $n \in \mathbb{N}_0$ mit

$$t = n\,\eta + \delta \quad \text{und } 0 \le \delta < \eta,$$

also

$$\big\|H(t)\big\| = \big\|H(\delta)\cdot\big(H(\eta)\big)^n\big\| \le k^{n+1} \le k\,k^{t/\eta} = k\,\mathrm{e}^{\omega t}. \qquad \square$$

Folgerung 10.1.9: *Es sei $\{H(t)\}$ eine $\mathcal{C}_0$-Halbgruppe. Dann ist für alle $x \in \mathcal{X}$ die Abbildung*

$$\begin{array}{ccc} \mathbb{R}_0^+ & \longrightarrow & \mathcal{X}, \\ t & \longmapsto & H(t)x \end{array}$$

stetig.

Beweis: Es seien $t, h \ge 0$. Dann ist

$$\|H(t+h)x - H(t)x\| \le \|H(t)\| \cdot \|H(h)x - x\| \le M\,\mathrm{e}^{\omega t}\,\|H(h)x - x\| \to 0$$

und

$$\|H(t-h)x - H(t)x\| \le \|H(t-h)\|\,\|x - H(h)x\| \le M\,\mathrm{e}^{\omega t}\,\|x - H(h)x\| \to 0. \qquad \square$$

Als nächstes zeigen wir

Satz 10.1.10: *Es seien $\{H(t)\}$ eine $\mathcal{C}_0$-Halbgruppe und A ihr Erzeuger. Dann gilt:*

1. $\forall\, x \in \mathcal{X} \qquad \lim\limits_{h\downarrow 0} \dfrac{1}{h}\displaystyle\int_t^{t+h} H(s)x\,ds = H(t)x.$
2. $\forall\, x \in \mathcal{X} \qquad \displaystyle\int_0^t H(s)x\,ds \in \mathcal{D}(A) \quad \textit{und} \quad A\int_0^t H(s)x\,ds = H(t)x - x.$
3. $\forall\, x \in \mathcal{D}(A) \qquad H(t)x \in \mathcal{D}(A) \quad \textit{und} \quad \dfrac{d}{dt}H(t)x = AH(t)x = H(t)Ax.$
4. $\forall\, x \in \mathcal{D}(A) \qquad H(t)x - H(s)x = \displaystyle\int_s^t H(\tau)Ax\,d\tau.$

Beweis:

1. Die erste Aussage ist klar, weil $H(t)x$ stetig ist.

2. Es sei $h > 0$. Dann folgt für $h \downarrow 0$

$$\frac{H(h) - id}{h}\int_0^t H(s)x\,ds = \frac{1}{h}\int_0^t \big(H(s+h)x - H(s)x\big)ds = \\ = \frac{1}{h}\int_t^{t+h} H(s)x\,ds - \frac{1}{h}\int_0^h H(s)x\,ds \to H(t)x - x.$$

3. Es sei $h > 0$. Dann folgt für $h \downarrow 0$

$$\frac{H(h) - id}{h}\,H(t)x = H(t)\,\frac{H(h) - id}{h}x \to H(t)Ax,$$

also $H(t)x \in \mathcal{D}(A)$ sowie $AH(t)x = H(t)Ax$ und

$$\lim_{h\downarrow 0}\frac{H(t+h)x - H(t)x}{h} = H(t)Ax.$$

Nun gilt wegen $\|H(t-h)\| \le M\,\mathrm{e}^{\omega t}$ auch

$$\lim_{h\downarrow 0}\Big\{\frac{H(t)x - H(t-h)x}{h} - H(t)Ax\Big\} = \\ = \lim_{h\downarrow 0} H(t-h)\Big\{\frac{(H(h) - id)x}{h} - Ax\Big\} + \lim_{h\downarrow 0}\Big\{H(t-h)Ax - H(t)Ax\Big\} = o.$$

4. Die vierte Aussage folgt aus der dritten durch Integration. $\square$

Folgerung 10.1.11: *Der Erzeuger A einer $\mathcal{C}_0$-Halbgruppe ist abgeschlossen und besitzt einen in $\mathcal{X}$ dichten Definitionsbereich.*

Beweis: Für $x \in \mathcal{X}$ und $t > 0$ sei

$$x_t := \frac{1}{t} \int_0^t H(s)x \, ds.$$

Dann ist $x_t \in \mathcal{D}(A)$, und aus der ersten Aussage des vorigen Satzes folgt

$$\lim_{t \downarrow 0} x_t = x.$$

Damit ist $\overline{\mathcal{D}(A)} = \mathcal{X}$. Nach Definition ist A eine lineare Abbildung. Es seien $x_n \in \mathcal{D}(A)$ mit $x_n \to x$ und $Ax_n \to y$. Dann folgt aus der vierten Aussage

$$H(t)x_n - x_n = \int_0^t H(s)Ax_n \, ds,$$

also

$$H(t)x - x = \int_0^t H(s)y \, ds$$

oder

$$\frac{H(t)x - x}{t} = \frac{1}{t} \int_0^t H(s)y \, ds \to H(0)y = y.$$

Damit ist $x \in \mathcal{D}(A)$ und $Ax = y$. □

Es gilt auch

Satz 10.1.12: *Es seien $\{H(t)\}$ und $\{I(t)\}$ $\mathcal{C}_0$-Halbgruppen mit demselben Erzeuger. Dann ist*

$$\forall\, t \geq 0 \qquad H(t) = I(t).$$

Beweis: Es seien $x \in \mathcal{D}(A)$, $0 \leq s \leq t$ und

$$\begin{aligned} C \;:\; \mathbb{R}_0^+ &\longrightarrow \mathcal{H}, \\ s &\longmapsto H(t-s)I(s)x. \end{aligned}$$

Dann folgt aus der dritten Aussage des letzten Satzes die Differenzierbarkeit von C, nach der Kettenregel also

$$\frac{d}{ds} H(t-s)I(s)x = -AH(t-s)I(s)x + H(t-s)AI(s)x = o.$$

Mithin ist C konstant, also

$$\forall\, x \in \mathcal{D}(A) \qquad H(t)x = I(t)x.$$

Nun ist $\mathcal{D}(A)$ in $\mathcal{X}$ dicht, und $H(t)$ sowie $I(t)$ sind beschränkt. Daraus folgt die Behauptung. □

Die letzte Folgerung hat gezeigt, daß $\mathcal{D}(A)$ in $\mathcal{X}$ dicht ist. Es gilt aber viel mehr, nämlich

Satz 10.1.13: *Es sei A der Erzeuger der $\mathcal{C}_0$-Halbgruppe $\{H(t)\}$. Dann gilt*

$$\bigcap_{n=1}^{\infty} \mathcal{D}(A^n) \;\; \textit{ist in } \mathcal{X} \;\; \textit{dicht.}$$

Beweis: Es seien $\mathcal{D} := \mathring{\mathcal{C}}_\infty\big((0,\infty), \mathbb{C}\big)$, $x \in \mathcal{X}$, $\varphi \in \mathcal{D}$ und

$$y := x(\varphi) := \int_0^\infty \varphi(s)\, H(s)x \, ds.$$

Dann gilt mit $h > 0$, h genügend klein,

$$\frac{H(h) - id}{h} y = \int_0^\infty \frac{\varphi(s-h) - \varphi(s)}{h} H(s)x \, ds,$$

also $y \in \mathcal{D}(A)$ und

$$\lim_{h \downarrow 0} \frac{H(h) - id}{h} y = -\int_0^\infty \varphi'(s) H(s)x \, dx = Ay.$$

Daraus folgt für $n \in \mathbb{N}$

$$A^n y = (-1)^n \int_0^\infty \varphi^{(n)}(s) H(s) x \, ds$$

und

$$y \in \bigcap_{n=1}^{\infty} \mathcal{D}(A^n).$$

Es sei

$$\mathcal{Y} := \operatorname{span}\left(\{ x(\varphi) \mid x \in \mathcal{X} \text{ und } \varphi \in \mathcal{D} \} \right).$$

Dann ist $\mathcal{Y}$ in $\mathcal{X}$ dicht. Andernfalls würde aus dem Satz von Hahn-Banach

$$\exists\, x' \in \mathcal{X}',\, x' \neq o, \quad \forall\, y \in \mathcal{Y} \qquad x'y = 0$$

folgen, also für alle $x \in \mathcal{X}$ und $\varphi \in \mathcal{D}$

$$0 = x'\Big(\int_0^\infty \varphi(s) H(s) x \, ds \Big) = \int_0^\infty \varphi(s) x' H(s) x \, ds.$$

Für alle $x \in \mathcal{X}$ würde deshalb

$$s \longmapsto x' H(s) x$$

auf $(0, \infty)$ verschwinden. Weil H stetig ist, würden wir schließlich

$$\forall\, x \in \mathcal{X} \qquad 0 = x' x$$

erhalten, also $x' = o$. □

Für $\mathcal{C}_0$-Halbgruppen wissen wir nach Satz 10.1.8

$$\exists\, \omega \geq 0 \quad \exists\, M \geq 1 \quad \forall\, t \in \mathbb{R}_0^+ \qquad \big\| H(t) \big\| \leq M \, \mathrm{e}^{\omega t}.$$

Wir unterscheiden nun:

Definition 10.1.14: *Es sei $\{H(t)\}$ eine $\mathcal{C}_0$-Halbgruppe. Dann heißt*

1. $\{H(t)\}$ gleichmäßig beschränkt, *wenn $\omega = 0$ ist,*
2. $\{H(t)\}$ $\mathcal{C}_0$-Kontraktionshalbgruppe, *wenn*

$$\forall\, t \in \mathbb{R}_0^+ \qquad \big\| H(t) \big\| \leq 1$$

ist.

Folgerung 10.1.15: *Die $\mathcal{C}_0$-Halbgruppe $\{H(t)\}$ sei gleichmäßig beschränkt, es sei A ihr Erzeuger und $x \in \mathcal{D}(A^2)$. Dann gilt*

$$\big\| Ax \big\|^2 \leq 4M^2 \big\| A^2 x \big\| \cdot \|x\|.$$

Beweis: Aus Satz 10.1.10 folgt

$$H(t)x - x = \int_0^t H(s) Ax \, ds = t\, A\, x + I$$

mit

$$\begin{aligned} I &:= \int_0^t \big(H(s) Ax - Ax \big) ds = \int_0^t \Big(\int_0^s H(\tau) A^2 x \, d\tau \Big) ds = \\ &= \int_0^t \Big(\int_0^t \chi_{(0,s)}(\tau) H(\tau) A^2 x \, d\tau \Big) ds = \int_0^t \Big(\int_0^t \chi_{(0,s)}(\tau) H(\tau) A^2 x \, ds \Big) d\tau = \\ &= \int_0^t \Big(\int_\tau^t H(\tau) A^2 x \, ds \Big) d\tau = \int_0^t (t - \tau) H(\tau) A^2 x \, d\tau. \end{aligned}$$

Daraus folgt

$$\|Ax\| \leq \frac{1}{t} \Big\{ \|H(t)x\| + \|x\| + \int_0^t (t - s) \big\| H(s) A^2 x \big\| \, ds \Big\} \leq \frac{2M\|x\|}{t} + \frac{Mt}{2} \big\| A^2 x \big\|.$$

Ist $A^2 x = o$, dann erhalten wir für $t \to \infty$ daraus $Ax = o$. Andernfalls wählen wir

$$t := 2\sqrt{\|x\|} / \sqrt{\|A^2 x\|}.$$

□

Beispiel 10.1.16: *Es seien* $\mathcal{X} := \mathcal{BC}(\mathbb{R})$ *und*

$$\big(H(t)x\big)(s) := x(t+s).$$

Dann ist $\{H(t)\}$ eine $\mathcal{C}_0$-Halbgruppe mit $\|H(t)\| \le 1$ und $\mathcal{D}(A) = \mathcal{C}_1(\mathbb{R})$ sowie $Ax = x'$. Aus der letzten Folgerung erhalten wir deshalb für $f \in \mathcal{C}^2(\mathbb{R})$ die *Landausche Ungleichung*

$$\big\|f'\big\|_\infty^2 \le 4\,\big\|f''\big\|_\infty \cdot \big\|f\big\|_\infty,$$

benannt nach Edmund Landau, 1877–1938.

Wir kommen nun zum Hauptresultat dieses Abschnitts, nämlich dem

Satz von Hille-Yosida: *Eine lineare Abbildung A erzeugt genau dann eine $\mathcal{C}_0$-Kontraktionshalbgruppe, wenn folgendes gilt:*

1. *A ist abgeschlossen mit $\overline{\mathcal{D}(A)} = \mathcal{X}$.*
2. *Die Resolventenmenge $\rho(A)$ enthält $\mathbb{R}^+$, und es ist*

$$\forall\,\lambda > 0 \qquad \big\|(A - \lambda\,id)^{-1}\big\| \le \frac{1}{\lambda}.$$

Der Satz ist nach Einar Hille, 1894–1980, und Kôsaku Yosida, 1909–1990, benannt.

Beweis:

$\Rightarrow$: A erzeuge die $\mathcal{C}_0$-Halbgruppe $\{H(t)\}$. Dann ist A nach Folgerung 10.1.11 abgeschlossen und hat einen dichten Definitionsbereich. Es seien $\lambda > 0$ und $x \in \mathcal{X}$. Dann existiert

$$R(\lambda)x := \int_0^\infty \mathrm{e}^{-\lambda t} H(t)x\,dt,$$

weil $\{H(t)\}$ gleichmäßig beschränkt ist. Die Abbildung $R(\lambda)$ ist linear mit

$$\|R(\lambda)x\| \le \int_0^\infty \mathrm{e}^{-\lambda t}\,\|H(t)x\|\,dt \le \frac{\|x\|}{\lambda}.$$

Für $h > 0$ gilt

$$\frac{H(h) - id}{h}\,R(\lambda)x = \frac{\mathrm{e}^{\lambda h} - 1}{h}\int_0^\infty \mathrm{e}^{-\lambda t} H(t)x\,dt - \frac{\mathrm{e}^{\lambda h}}{h}\int_0^h \mathrm{e}^{\lambda t} H(t)x\,dt.$$

Für $h \downarrow 0$ folgt daraus

$$\forall\,x \in \mathcal{X} \quad \forall\,\lambda > 0 \qquad R(\lambda)x \in \mathcal{D}(A)$$

und

$$A\,R(\lambda)x = \big(\lambda\,R(\lambda) - id\big)x$$

oder

$$\big(\lambda\,id - A\big)R(\lambda) = id.$$

Nun ist für $x \in \mathcal{D}(A)$

$$R(\lambda)Ax = \int_0^\infty \mathrm{e}^{-\lambda t} H(t)Ax\,dt = A\int_0^\infty \mathrm{e}^{-\lambda t} H(t)x\,dt = A\,R(\lambda)x,$$

weil A abgeschlossen ist. Wir erhalten also

$$\forall\,x \in \mathcal{D}(A) \qquad R(\lambda)\,(\lambda\,id - A)x = x.$$

Mithin existiert für alle $\lambda > 0$

$$(A - \lambda\,id)^{-1} = -R(\lambda),$$

und es ist $\|R(\lambda)\| \le \frac{1}{\lambda}$. □

Bevor wir nun die andere Richtung der Aussage des Satzes von Hille-Yosida beweisen, zeigen wir drei Hilfssätze:

Lemma 10.1.17: *Die Abbildung A genüge den beiden Bedingungen im Satz von Hille-Yosida. Dann gilt*

$$\forall\, x \in \mathcal{X} \qquad \lim_{\lambda\to\infty} \lambda\big(A - \lambda\, id\big)^{-1} x + x = o.$$

Beweis: Für alle $x \in \mathcal{D}(A)$ gilt

$$\begin{aligned}\big\|\lambda(A-\lambda)^{-1}x + x\big\| &= \big\| -(A-\lambda)(A-\lambda)^{-1}x + A(A-\lambda)^{-1}x + x\big\| = \\ &= \big\|A(A-\lambda)^{-1}x\big\| = \big\|(A-\lambda)^{-1}Ax\big\| \le \frac{1}{\lambda}\big\|Ax\big\| \to 0.\end{aligned}$$

Daraus folgt die Behauptung, weil $\mathcal{D}(A)$ in $\mathcal{X}$ dicht und $\|\lambda(A-\lambda)^{-1}\| \le 1$ ist. □

Definition 10.1.18: *Die Folge (A_λ), $\lambda \in \mathbb{R}^+$, mit*

$$A_\lambda := -\lambda\, A(A - \lambda\, id)^{-1} \in \mathcal{CL}(\mathcal{X})$$

heißt Yosida-Approximierende *von A.*

Es ist

$$A_\lambda = -\lambda\left\{(A-\lambda)(A-\lambda)^{-1} + \lambda(A-\lambda)^{-1}\right\} = -\lambda\, id - \lambda^2(A-\lambda)^{-1}.$$

Lemma 10.1.19: *Die Abbildung A genüge den beiden Bedingungen im Satz von Hille-Yosida. Dann gilt*

$$\forall\, x \in \mathcal{D}(A) \qquad \lim_{\lambda\to\infty} A_\lambda x = Ax.$$

Das folgt unmittelbar aus dem vorigen Lemma.

Lemma 10.1.20: *Die Abbildung A genüge den beiden Bedingungen im Satz von Hille-Yosida. Dann erzeugt A_λ die gleichmäßig stetige Kontraktionshalbgruppe $\mathrm{e}^{A_\lambda t}$, und es gilt*

$$\forall\, x \in \mathcal{X} \quad \forall\, \lambda, \mu > 0 \qquad \big\|\mathrm{e}^{A_\lambda t}x - \mathrm{e}^{A_\mu t}x\big\| \le t\,\big\|A_\lambda x - A_\mu x\big\|.$$

Beweis: Nach Definition ist $A_\lambda \in \mathcal{CL}(\mathcal{X})$ und erzeugt deshalb die gleichmäßig stetige Halbgruppe $\mathrm{e}^{A_\lambda t}$. Es gilt

$$\big\|\mathrm{e}^{A_\lambda t}\big\| = \mathrm{e}^{-\lambda t}\big\|\mathrm{e}^{-\lambda^2 t(A-\lambda)^{-1}}\big\| \le \mathrm{e}^{-\lambda t}\mathrm{e}^{\lambda^2 t\|(A-\lambda)^{-1}\|} \le 1.$$

Mithin ist $\left\{\mathrm{e}^{A_\lambda t}\right\}$ eine Kontraktionshalbgruppe. Nach Definition kommutieren $\mathrm{e}^{A_\lambda t}$, $\mathrm{e}^{A_\mu t}$, A_λ und A_μ. Daraus folgt

$$\begin{aligned}\big\|\mathrm{e}^{A_\lambda t}x - \mathrm{e}^{A_\mu t}x\big\| &= \left\|\int_0^1 \frac{d}{ds}\left(\mathrm{e}^{A_\lambda ts}\mathrm{e}^{A_\mu t(1-s)}x\right) ds\right\| \le \\ &\le \int_0^1 t\left\|\mathrm{e}^{A_\lambda ts}\mathrm{e}^{A_\mu t(1-s)}\big(A_\lambda x - A_\mu x\big)\right\| ds \le t\,\big\|A_\lambda x - A_\mu x\big\|.\end{aligned}$$ □

Wir beenden nun den Beweis des Satzes von Hille-Yosida:

⇐: Die Abbildung A genüge den Bedingungen im Satz von Hille-Yosida. Dann ist

$$\big\|\mathrm{e}^{A_\lambda t}x - \mathrm{e}^{A_\mu t}x\big\| \le t\,\|A_\lambda x - A_\mu x\| \le t\,\|A_\lambda x - Ax\| + t\,\|Ax - A_\mu x\|,$$

und aus Lemma 10.1.19 folgt, daß für $x \in \mathcal{D}(A)$ die Folge $\left(\mathrm{e}^{A_\lambda t}x\right)$ gegen $H(t)x$ konvergiert, und zwar gleichmäßig in festen Intervallen $t \in [0,T]$. Es ist $\mathcal{D}(A)$ in $\mathcal{X}$ dicht und $\|\mathrm{e}^{A_\lambda t}\| \le 1$. Daraus folgt

$$\forall\, x \in \mathcal{X} \qquad \lim_{\lambda\to\infty} \mathrm{e}^{A_\lambda t}x = H(t)x, \tag{$*$}$$

und zwar gleichmäßig in $[0,T]$. Ferner ist $H(0) = id$, $H(s+t) = H(s)H(t)$ und $\|H(t)\| \le 1$, weil $\left\{\mathrm{e}^{A_\lambda t}\right\}$ eine Kontraktionshalbgruppe ist. Die Abbildung

$$t \longmapsto H(t)x$$

ist für $t \ge 0$ stetig, als gleichmäßiger Limes stetiger Abbildungen $t \mapsto \mathrm{e}^{A_\lambda t}x$. Mithin ist $H(t)$ eine $\mathcal{C}_0$-Kontraktionshalbgruppe.

Es bleibt zu zeigen, daß A Erzeuger von $H(t)$ ist: Es sei $x \in \mathcal{D}(A)$. Dann folgt aus Gl. $(*)$

$$H(t)x - x = \lim_{\lambda\to\infty} \left(\mathrm{e}^{A_\lambda t}x - x\right) = \lim_{\lambda\to\infty} \int_0^t \mathrm{e}^{A_\lambda s} A_\lambda x \, ds = \int_0^t H(s)Ax \, ds.$$

Es sei B Erzeuger von $\{H(t)\}$. Dann folgt aus Satz 10.1.10

$$\forall\, x \in \mathcal{D}(B) \qquad H(t)x - x = \int_0^t H(s)Bx \, ds,$$

mithin ist $A \subset B$. Aus dem bereits bewiesenen Teil unseres Satzes folgt $1 \in \rho(B)$, und aus der zweiten Bedingung auch $1 \in \rho(A)$. Die Abbildungen A und B sind abgeschlossen. Damit erhalten wir

$$(B - id)\mathcal{D}(A) = (A - id)\mathcal{D}(A) = \mathcal{X},$$

also

$$\mathcal{D}(B) = \big(B - id\big)^{-1}\mathcal{X} = \mathcal{D}(A)$$

oder $A = B$. □

Es folgen drei Korollare.

Korollar 10.1.21: *Es seien A Erzeuger der $\mathcal{C}_0$-Kontraktionshalbgruppe $\{H(t)\}$ und (A_λ) die Yosida-Approximierende von A. Dann gilt*

$$\forall\, x \in \mathcal{X} \qquad H(t)x = \lim_{\lambda\to\infty} \mathrm{e}^{A_\lambda t}x.$$

Beweis: Der Limes existiert und definiert eine Halbgruppe. Die Eindeutigkeit folgt aus Satz 10.1.12. □

Korollar 10.1.22: *Es sei A Erzeuger der $\mathcal{C}_0$-Kontraktionshalbgruppe $\{H(t)\}$. Dann gilt*

$$\big\{\lambda \in \mathbb{C} \;\big|\; \mathrm{Re}\,\lambda > 0\big\} \subset \rho(A)$$

und

$$\forall\, \lambda, \mathrm{Re}\,\lambda > 0, \qquad \big\|(A - \lambda)^{-1}\big\| \le \frac{1}{\mathrm{Re}\,\lambda}.$$

Beweis: Es ist

$$R(\lambda)x := \int_0^\infty \mathrm{e}^{-\lambda t} H(t)x \, dt$$

für alle $x \in \mathcal{X}$ und $\lambda \in \mathbb{C}$ mit $\mathrm{Re}\,\lambda > 0$ wohldefiniert. Wie im Beweis des Satzes von Hille-Yosida folgt $R(\lambda) = -(A - \lambda\, id)^{-1}$ und

$$\|R(\lambda)x\| \le \|x\|/\mathrm{Re}\,\lambda.$$ □

Korollar 10.1.23: *Ein linearer Operator A erzeugt genau dann eine $\mathcal{C}_0$-Halbgruppe $\{H(t)\}$ mit $\|H(t)\| \le \mathrm{e}^{\omega t}$, $\omega \ge 0$, wenn folgendes gilt:*

1. *A ist abgeschlossen und $\overline{\mathcal{D}(A)} = \mathcal{X}$.*
2. $$\Big\{\lambda \in \mathbb{C} \;\Big|\; \mathrm{Im}\,\lambda = 0 \text{ und } \lambda > \omega\Big\} \subset \rho(A)$$

   und

   $$\big\|(A - \lambda)^{-1}\big\| \le \frac{1}{\lambda - \omega} \quad \text{für solche } \lambda.$$

Beweis: Es ist

$$I(t) := \mathrm{e}^{-\omega t} H(t)$$

eine Kontraktionshalbgruppe. Ihr Erzeuger ist

$$A - \omega\, id.$$ □

Beispiel 10.1.24: *In $\mathbb{R}_0^+ \times G$, $G \subset \mathbb{R}$ betrachten wir die* Thermoelastizitätsgleichungen

$$u_{tt} - u_{xx} + \gamma\,\theta_x = o$$
$$\theta_t - \theta_{xx} + \gamma\, u_{tx} = o$$

mit dem Kopplungsparameter $\gamma \in \mathbb{R}$ *und vorgegebenen Anfangswerten $u(0) = u_0$, $u_t(0) = u_1$ und $\theta(0) = \theta_0$.*

Es sei

$$V := \begin{pmatrix} u_x \\ u_t \\ \theta \end{pmatrix}.$$

Dann ist mit $d := d/dx$

$$V_t = \begin{pmatrix} u_{tx} \\ u_{xx} - \gamma\,\theta_x \\ \theta_{xx} - \gamma\,u_{tx} \end{pmatrix} = \begin{pmatrix} 0 & d & 0 \\ d & 0 & -\gamma\,d \\ 0 & -\gamma\,d & d^2 \end{pmatrix} V,$$

und mit

$$A := \begin{pmatrix} 0 & d & 0 \\ d & 0 & -\gamma\,d \\ 0 & -\gamma\,d & d^2 \end{pmatrix}$$

erhalten wir

$$V_t = AV \quad \text{mit } V(0) = V_0 = \begin{pmatrix} du_0 \\ u_1 \\ \theta_0 \end{pmatrix}.$$

Dabei seien $\mathcal{H} := \left(\mathcal{L}^2(G)\right)^3$, $V_0 \in \mathcal{H}$ und

$$\mathcal{D}(A) := \left\{ V \in \mathcal{H} \;\middle|\; V_2 \in \mathring{\mathcal{H}}_1(G),\ V_3 \in \mathring{\mathcal{H}}_1(G) \text{ und } AV \in \mathcal{H} \right\}.$$

Wir betrachten also die Dirichletsche Randwertaufgabe. Dann ist

$$(AV, V) = -(dV_3, dV_3) - 2\,\mathrm{i}\,\mathrm{Im}\left\{(V_1, dV_2) + \gamma(dV_2, V_3)\right\},$$

also

$$\mathrm{Re}\,(AV, V) = -\,\|dV_3\|^2 \le 0.$$

Die Abbildung A ist abgeschlossen und besitzt einen dichten Definitionsbereich ist. Für alle λ mit $\mathrm{Re}\,\lambda > 0$ folgt aus $(A - \lambda)V = o$

$$\|dV_3\|^2 = -\mathrm{Re}\,(AV, V) = -\mathrm{Re}\,\lambda\,\|V\|^2,$$

also $V = o$. Analog sieht man $\mathcal{N}(A^* - \bar{\lambda}) = \{o\}$, mithin ist

$$\mathcal{H} = \overline{R(A - \lambda\,id)}.$$

Aus $(A - \lambda)V = F \in \mathcal{H}$ erhalten wir schließlich

$$\|dV_3\|^2 = -\mathrm{Re}\,(AV, V) = -\mathrm{Re}\,\lambda\,\|V\|^2 - \mathrm{Re}\,(V, F),$$

also

$$\|V\| \le \|F\|/\mathrm{Re}\,\lambda$$

oder

$$\left\|(A - \lambda)^{-1} F\right\| \le \|F\|/\mathrm{Re}\,\lambda.$$

Mithin erzeugt A eine $\mathcal{C}_0$-Kontraktionshalbgruppe $\{H(t)\}$,

$$H(t) = \mathrm{e}^{At},$$

und unsere Anfangsrandwertaufgabe wird durch

$$V(t) = \mathrm{e}^{At} V_0$$

für $t \ge 0$ gelöst. □

10.1.3 Stark stetige Gruppen

Im folgenden wollen wir noch kurz Gruppen von beschränkten Abbildungen besprechen.

Definition 10.1.25: *Eine Schar* $\{H(t)\}_{t\in\mathbb{R}}$, $H(t)\in\mathcal{CL}(\mathcal{X})$, *heißt* stark stetige Gruppe *oder* $\mathcal{C}_0$*-Gruppe, wenn folgendes gilt:*

1. $H(0)=id$,
2. $H(s+t)=H(s)H(t)$ *für alle* $s,t\in\mathbb{R}$.
3. *Für alle* $x\in\mathcal{X}$ *ist*
$$\lim_{t\to 0} H(t)x = x$$

Definition 10.1.26: *Es seien* $\{H(t)\}$ *eine Gruppe und*

$$A \;:\; \mathcal{D}(A)\subset\mathcal{X}\longrightarrow\mathcal{X}$$

eine lineare Abbildung mit

$$\mathcal{D}(A) := \left\{x\in\mathcal{X} \;\middle|\; \exists\, y\in\mathcal{X} \quad \lim_{t\to 0}\frac{H(t)x-x}{t}=y\right\}$$

und $Ax := y$. *Dann heißt* A (infinitesimaler) Erzeuger *der Gruppe* $\{H(t)\}$.

Für $t\geq 0$ sind offenbar $H(t)$ und $H(-t)$ Halbgruppen mit den Erzeugern A bzw. $-A$. Erzeugen umgekehrt A die Halbgruppe $H_+(t)$ und $-A$ die Halbgruppe $H_-(t)$, dann erzeugt A die Gruppe

$$H(t) := \begin{cases} H_+(t) & \text{für } t\geq 0\\ H_-(-t) & \text{für } t\leq 0.\end{cases}$$

Es sei nun $\mathcal{X}$ ein Hilbertraum. In Hilberträumen spielen die stark stetigen unitären Gruppen $\{U(t)\}$ eine besondere Rolle. Es soll also $U^*(t)=U^{-1}(t)$ für alle t sein. In physikalischen Anwendungen beschreiben solche Gruppen Bewegungen bzw. ihre Invarianten. Ist u^0 die Anfangslage, dann kann $U(t)u^0$ zum Beispiel die Energie darstellen, und es ist $\|U(t)u^0\|=\|u^0\|$ für alle t. Es gilt der

Satz von Stone: *Eine lineare Abbildung* A *erzeugt genau dann eine stark stetige unitäre Gruppe* $\{U(t)\}$, *wenn* $\mathrm{i}\,A$ *selbstadjungiert ist, wenn also* $A^*=-A$ *ist.*

Der Beweis folgt aus dem Vorangegangenen:

$\Leftarrow$: Es ist
$$-Ax=\lim_{t\downarrow 0}\frac{U(-t)x-x}{t}=\lim_{t\downarrow 0}\frac{U^*(t)x-x}{t}=A^*x.$$

$\Rightarrow$: Es sei $B:=\mathrm{i}\,A$ und $B^*=B$. Dann ist B dicht definiert und damit auch A. Es ist $\sigma(B)\subset\mathbb{R}$, und aus $(B-\lambda)x=y$, $\mathrm{Im}\,\lambda\neq 0$, folgt
$$-|\mathrm{Im}\,\lambda|\,\|x\|\leq\|y\|.$$

Für A bedeutet das $\mathbb{R}^+,\mathbb{R}^-\subset\rho(A)$ und

$$\forall\,\mu\in\mathbb{R}\setminus\{0\}\qquad \|(A-\mu)^{-1}\|\leq\frac{1}{|\mu|}.$$

Mithin erzeugen sowohl A als auch $-A$ eine $\mathcal{C}_0$-Kontraktionshalbgruppe. A erzeugt deshalb eine stark stetige Gruppe $\{H(t)\}$ mit
$$\|H(t)\|\leq 1.$$

Ferner ist
$$H^{-1}(t)=H(-t)=H^*(t).$$

Damit ist $\{H(t)\}$ eine unitäre Gruppe. □

10.2 Ein Darstellungssatz

Es sei

$$\mathcal{C}^0(\mathbb{R}) := \Big\{ f \in \mathcal{C}(\mathbb{R}) \;\Big|\; \lim_{x \to \pm\infty} f(x) = 0 \Big\}.$$

$\mathcal{C}^0(\mathbb{R})$ ist ein normierter Raum, ausgestattet mit der Supremumsnorm $\| \cdot \|_\infty$. Ein Funktional $\Lambda \in \mathcal{C}^0(\mathbb{R})'$ heißt *positiv*, wenn

$$\forall\, \varphi \in \mathcal{C}^0(\mathbb{R}),\ \varphi \geq 0, \qquad \Lambda\varphi \geq 0$$

ist. Dann gilt der folgende

Darstellungssatz für positive Funktionale aus $\mathcal{C}^0(\mathbb{R})'$: *Zu jedem positiven $\Lambda \in \mathcal{C}^0(\mathbb{R})'$ gibt es genau ein beschränktes positives Borelmaß mit*

$$\forall\, \varphi \in \mathcal{C}^0(\mathbb{R}) \qquad \Lambda\varphi = \int_{\mathbb{R}} \varphi \, d\mu$$

und $\mu(\mathbb{R}) = \|\Lambda\|$.

Dem Beweis des Darstellungssatzes liegt das folgende Resultat von F. Riesz aus dem Jahre 1909 zugrunde (vgl. den Beweis zu Satz 4.2.1):

Satz 10.2.1: *Es seien $I = [a,b]$, $\mathcal{X} = \mathcal{C}(I, \mathbb{C})$ und $x' \in \mathcal{X}'$. Dann gibt es genau ein $v \in \mathcal{NBV}(I, \mathbb{C})$ mit $\|v\|_{\mathcal{NBV}} = \|x'\|$ und*

$$\forall\, x \in \mathcal{X} \qquad x'x = \int_I x(t)\, dv(t).$$

Nun ist bekannt, daß sich jede Funktion mit beschränkter Variation als Differenz zweier nicht negativer, monoton wachsender Funktionen schreiben läßt. Daraus folgt insbesondere

Satz 10.2.2: *Es seien $I = [a,b]$, $\mathcal{X} = \mathcal{C}(I, \mathbb{C})$ und $x' \in \mathcal{X}'$ ein positives Funktional. Dann gibt es genau eine monoton wachsende Funktion $\mu \in \mathcal{NBV}(I, \mathbb{C})$ mit $\|\mu\|_{\mathcal{NBV}} = \|x'\|$ und*

$$\forall\, x \in \mathcal{X} \qquad x'x = \int_I x(t)\, d\mu(t).$$

Einen Beweis dieser Sätze findet man in *Taylor & Lay [1980], S. 146f* und *Heuser [1986], S. 317f.* Allgemeinere Aussagen finden sich in *Rudin [1987], S. 40f und S. 130f.*

Wir skizzieren nun den Beweis des Darstellungssatzes:

1. Wir zeigen die Eindeutigkeit. Es seien

$$\forall\, \varphi \in \mathcal{C}^0(\mathbb{R}) \qquad 0 = \int_{\mathbb{R}} \varphi \, d\mu$$

und $\mu(a) < \mu(b)$. Dann wählen wir ein $\varphi \in \mathring{\mathcal{C}}(\mathbb{R})$ mit $\varphi \,|\, [a,b] = 1$, $0 \leq \varphi \leq 1$, und erhalten den Widerspruch $0 = \mu(a) = \mu(b)$.

2. Es sei nun $\mathcal{Y} := \mathring{\mathcal{C}}(\mathbb{R}, \mathbb{C})$. Dann gilt die Satz 10.2.2 entsprechende Aussage auch für $\mathcal{Y}$: Es sei nämlich $y' \in \mathcal{Y}'$ ein positives Funktional und $y \in \mathcal{Y}$. Dann gibt es ein $I_n := [-n, n]$ mit $\operatorname{supp} y \subset I_n$ und eine monoton wachsende Funktion ν_n mit

$$y'y = \int_{I_n} y(t)\, d\nu_n(t)$$

und $\nu_n(I_n) \leq \|y'\|$. Es seien

$$\mu_n(t) := \begin{cases} 0 & \text{für } t < -n \\ \nu_n(t) & \text{für } -n \leq t \leq n \\ \nu_n(n) & \text{für } n < t. \end{cases}$$

Dann ist auch

$$y'y = \int_{\mathbb{R}} y(t)\, d\mu_n(t)$$

mit $\mu_n(\mathbb{R}) \leq \|y'\|$. Es sei nun für $M \subset \mathbb{R}$

$$\mu_n(M) := \int_{\mathbb{R}} \chi_M(t)\, d\mu_n(t).$$

Dann gilt

$$\forall\, M \subset \mathbb{R} \quad \exists\, n(M) \quad \forall\, n \geq n(M) \qquad \mu_n(M) = \mu_{n(M)}(M).$$

Mithin existiert

$$\mu(M) := \lim_{n\to\infty} \mu_n(M),$$

und es ist μ ein positives Maß mit $\mu(\mathbb{R}) = \|y'\|$.

3. Weil $\mathring{\mathcal{C}}(\mathbb{R},\mathbb{C})$ bezüglich der Supremumsnorm in $\mathcal{C}^0(\mathbb{R},\mathbb{C})$ dicht ist, folgt die Behauptung.

10.3 Spektraldarstellung unitärer Abbildungen

seien $\mathcal{H}$ ein komplexer Hilbertraum und

$$U \;:\; \mathcal{H} \longrightarrow \mathcal{H}$$

eine *unitäre Abbildung*, also $U \in \mathcal{CL}(\mathcal{H})$, $\mathcal{R}(U) = \mathcal{H}$ und $UU^* = U^*U = id$. Offenbar ist U invertierbar, erhält das Skalarprodukt, und es ist $\|U\| = 1$.

Ferner ist $\sigma(U) \subset \partial B(o,1)$. Wegen $\|U\| = 1$ ist nämlich der Spektralradius von U gleich Eins, also $\sigma(U) \subset \bar{B}(0,1)$, und wegen der Bijektivität von U gilt sogar

$$\sigma(U) \subset \bar{B}(0,1)\backslash\{o\}.$$

Es sei nun $\lambda \in \mathbb{C}$ mit $0 < |\lambda| < 1$. Dann ist $1/|\lambda| > 1$ und somit $1/\lambda \in \rho(U^*)$ und

$$\lambda\, id - U = \lambda\, U\big(U^* - \tfrac{1}{\lambda} id\big)$$

als Produkt zweier bijektiver Abbildungen ebenfalls bijektiv. Somit gehört λ zu $\rho(U)$, und damit ist $\sigma(U) \subset \partial B(0,1)$.

Man kann nun nachrechnen (vgl. *Heuser [1986], S. 562*), daß es zu einer unitären Abbildung U eine selbstadjungierte Abbildung $A \in \mathcal{CL}(\mathcal{H})$ mit $\sigma(A) \subset [-\pi,\pi]$ und

$$U = \mathrm{e}^{\mathrm{i}A}$$

gibt. Es sei $P(\lambda)$ die Spektralschar von A. Dann erhält man die Spektraldarstellung von U durch

$$Ux = \int_{-\pi-0}^{\pi} \mathrm{e}^{\mathrm{i}\lambda}\, dP(\lambda)x.$$

1997

Literaturverzeichnis

Adams, R. A. [1975]: Sobolev spaces. Academic Press, New York.

Alt, H. W. [1985]: Lineare Funktionalanalysis. Springer-Verlag, Berlin et al.

Banach, S. et H. Steinhaus [1927]: Sur le principe de condensation de singularités. Fundamenta Math. 9, 50–61.

Barner, M. und F. Flohr [1982]: Analysis II. de Gruyter. Berlin et al.

Brown, R. [1993]: A Topological Introduction to Nonlinear Analysis. Birkhäuser Verlag, Basel et al.

Courant, R. [1920]: Über die Eigenwerte bei den Differentialgleichungen der mathematischen Physik. Math. Z. 7, 1–57.

Courant, R. und D. Hilbert [1993]: Methoden der mathematischen Physik. Springer-Verlag, Berlin et al.

Day, M. [1941]: Reflexive Banach spaces not isomorphic to uniformly convex spaces. Bull. AMS 47, 313–317.

Dieudonné, J. [1981]: History of Functional Analysis. North-Holland, Amsterdam et al.

Dunford, N. and J. T. Schwartz [1988]: Linear Operators I – III. John Wiley & Sons, New York et al.

Fredholm, I. [1903]: Sur une classe d'équations fonctionnelles. Acta Math. 27, 365–390.

Friedrichs, K. O. [1944]: The identity of weak and strong extensions of differential operators. Trans. AMS 55, 132–151.

Halmos, P. [1974]: A Hilbert Space Problem Book. Springer-Verlag, Berlin et al.

Hellwig, G. [1960]: Partielle Differentialgleichungen. B.G. Teubner, Stuttgart.

Heuser, H. [1986]: Funktionalanalysis. B.G. Teubner Verlag, Stuttgart.

Hilbert, D. [1912]: Grundzüge einer allgemeinen Theorie der Integralgleichungen. B.G. Teubner, Leipzig.

Hirzebruch, F. und W. Scharlau [1971]: Einführung in die Funktionalanalysis. Bibliographisches Institut, Mannheim.

Jörgens, K. [1970]: Lineare Integraloperatoren. B.G. Teubner, Stuttgart.

Kasuga, T. [1957]: On Sobolev-Friedrichs generalisation of derivatives. Proc. Jap. Ac. 33, 596–599.

Kato, T. [1966]: Perturbation Theory for Linear Operators. Springer-Verlag 1966, Berlin et al.

Leinfelder, H. [1979]: A geometric proof of the spectral theorem for unbounded self-adjoint operators Math. Ann. 242, 85–96.

Meyers, N. G. and J. Serrin [1964]: $H = W$. Proc. Nat. Acad. Sci. USA 51; 1055–1056.

Neumann, C. [1877]: Untersuchungen Über das logarithmische und Newtonsche Potential. B.G. Teubner; Leipzig.

Neumann, J. von [1929]: Allgemeine Eigenwerttheorie Hermitescher Funktionaloperatoren Math. Ann. 102, 49–131

Nirenberg, L. [1955]: Remarks on strongly elliptic differential equations. Comm. Pure Appl. Math. 8, 649–675.

Pazy, A. [1983]: Semigroups of Linear Operators and Applications to Partial Differential Equations. Springer-Verlag, Berlin et al.

Pedersen, G. [1989]: Analysis Now. Springer-Verlag, Berlin et al.

Reed, M. and B. Simon [1980]: Methods of Modern Mathematical Physics I; Functional Analysis. Academic Press, New York.

Riesz, F. [1918]: Über lineare Funktionalgleichungen. Acta Math. 41, 71–98.

Riesz, F. und B. Sz.-Nagy [1956]: Vorlesungen über Funktionalanalysis. VEB Deutscher Verlag der Wissenschaften, Berlin.

Rudin, W. [1987]: Real and complex analysis. McGraw-Hill Publ. Comp., New York et al.

Schauder, J. [1930]: Über lineare, vollstetige Funktionaloperationen. Studia Math. 2, 183–196.

Stone, M.H. [1932]: Linear transformations in Hilbert space. Amer. Math. Soc. Coll. Publ. 15; New York.

Taylor, A. and D. Lay [1980]: Introduction to Functional Analysis. John Wiley, New York et al.

Weidmann, J. [1976]: Lineare Operatoren in Hilberträumen. B.G. Teubner Verlag, Stuttgart.

Weierstraß, K. [1870]: Über das sogenannte Dirichlet'sche Princip. Gelesen in der Königl. Akademie der Wissenschaften am 14. 7. 1870.

Weyl, H. [1940]: The method of orthogonal projection in potential theory. Duke Math. J. 7, 411–444.

Witsch, K. J. [1993]: A remark on a compactness result in electromagnetic theory. Math. Meth. in the Appl. Sci. 16, 123–129.

Yosida, K. [1965]: Functional Analysis. Springer-Verlag, Berlin et al.

Bezeichnungen

Räume von Abbildungen und deren Normen:

$\{A;\ \|\cdot\|\}^{\sim}$, Vervollständigung von A unter der Norm $\|\cdot\|$
$\mathcal{B}(E,Y)$, 18
$\mathcal{BC}_k(\Omega)$, $\|\cdot\|_{\mathcal{BC}_k(\Omega)}$, 94
$\mathcal{BV}(I)$, $\mathcal{NBV}(I)$, 30, 68, 123
$\mathcal{C}_e(E,Y)$, 20
$\mathcal{C}(\Omega)$, 93
$\mathcal{C}^0(\mathbb{R})$, 193, 219
$\mathring{\mathcal{C}}(\Omega)$, $\mathring{\mathcal{C}}_m(\Omega)$, 94
$\mathcal{C}_\infty(\Omega)$, 93
$\mathcal{C}_{k,\alpha}(\bar{\Omega})$, $\|\cdot\|_{\mathcal{C}_{k,\alpha}(\bar{\Omega})}$, 93
$\mathcal{C}_k^{p*}(\Omega)$, 112
$\mathcal{CL}(\mathcal{X},\mathcal{Y})$, 31
$\mathcal{CL}(\mathcal{X}) = \mathcal{CL}(\mathcal{X},\mathcal{X})$
$\mathcal{D}(\Omega)$, $\mathcal{D}_K(\Omega)$, $p_{m,K}(\cdot)$, 106
$\mathcal{D}'(\Omega)$, 107
$\mathcal{D} = \mathring{\mathcal{C}}_\infty(\mathbb{R}^n)$, 110
$\mathcal{D}'$, 110
$\mathcal{E} = \mathcal{C}_\infty(\mathbb{R}^n)$, 110
$\mathcal{E}'$, 110
$\mathcal{F}(E,Y)$, 18
$\mathcal{H}_m^p(\Omega)$, $\|\cdot\|_{\mathcal{H}_m^p(\Omega)}$, 90, 112
$\mathring{\mathcal{H}}_m^p(\Omega)$, $\|\cdot\|_{\mathring{\mathcal{H}}_m^p(\Omega)}$, 99, 112
$\mathcal{KL}(\mathcal{X},\mathcal{Y})$, 143
$\mathcal{KL}(\mathcal{X}) = \mathcal{KL}(\mathcal{X},\mathcal{X})$
$\mathcal{K}_\alpha(\mathcal{X})$, 173
$\mathcal{L}^1(\Omega)$, $\mathcal{L}^1_{loc}(\Omega)$, 97
$\mathcal{L}^p(\Omega)$, $\|\cdot\|_{\mathcal{L}^p}$, 30, 95
$\mathcal{L}^\infty(\Omega)$, 30, 95
$\mathcal{L}^p_{loc}(\Omega)$, 95
$\mathcal{L}^p_k(\Omega)$, $\|\cdot\|_{\mathcal{L}^p_k}$, 95
$\mathcal{L}^2_{\mu_x}$, 197
$\mathcal{L}(\mathcal{X},\mathcal{Y})$, 6
$\mathcal{L}(\mathcal{X}) = \mathcal{L}(\mathcal{X},\mathcal{X})$
ℓ^2, $\|\cdot\|_{\ell^2}$, 5
ℓ^p, $\|\cdot\|_{\ell^p}$, 30
$\mathcal{S} = \mathcal{S}(\mathbb{R}^n)$, 104
$\mathcal{S}'$, 105
$S_{\mathcal{X}} := \{x \in \mathcal{X} \mid \|x\| = 1\}$
$\mathcal{W}_m^p(\Omega)$, $\|\cdot\|_{m,p,\Omega}$, 111

Funktionen und Distributionen:

$D(\cdot)$, $E(\cdot,\cdot)$, $J(\cdot)$, 128
δ_x, 108
e^{-iAt}, 126, 207
$E(t)$, 171, 205
$f * g$, 98
$F(\cdot)$, $\hat{f}$, 103
H, 108
j, j_ϵ, 100
J_ϵ, 100
$J_{\mathcal{X}}$, 50
π_μ, 74
T_f, 105
$[u]$, 107

Symbole:

A^{tr}, 7
A', 47, 83
A^*, 83
$\mathcal{B}$, 191
$B(x,\epsilon)$, 14
codim, 153
$C\sigma(A)$, $P\sigma(A)$, $R\sigma(A)$, 131
$d(\cdot,\cdot)$, d_e, d_∞, d_p, 15
$\delta(\cdot)$, 20
$\mathrm{dist}(\cdot,\cdot)$, 21
$\mathrm{conv}(F)$, 56, 71
$\mathcal{D}(A)$, $\mathcal{N}(A)$, $\mathcal{R}(A)$, 6
ess sup, 30
$G(A)$, $\|\langle\cdot,\cdot\rangle\|_{G(A)}$, 77
G_a, G, ∂G, 157
$H(t)$, 208
$\mathcal{H}_p$, $\mathcal{H}_c$, $\mathcal{H}_{pp}$, $\mathcal{H}_{ac}$, $\mathcal{H}_{sc}$, 200, 202
$\mathrm{höl}_\alpha(f,\bar{\Omega})$, 94
ind, 160
$M^\perp$, 139
$\mu_x, |\mu|, \mu_{x,y}$ 187, 191, 192
$\mathcal{N}_n$, $\mathcal{R}_n$, 146
N_ϵ, 23
ν_n, 169
$\mathcal{P}(\mathcal{X})$, Potenzmenge von $\mathcal{X}$
$P(\lambda)$, $P(I)$, 185
$\hat{P}(f)$, 196
$\Psi(A)$, 161
$r(A)$, 149
$R(A,\lambda)$, $R(\lambda)$, 134
$\rho(A)$, $\sigma(A)$, 131
$S = \partial B(o,1)$
$\Sigma(A)$, 161
span, 5
supp, 94, 109
$\mathcal{T}$, 13, 75
$U(x,E,\epsilon)$, 75
w-$\lim_{n\to\infty}$, w*-$\lim_{n\to\infty}$, 64
$\mathcal{X}^f$, 7
$\mathcal{X}'$, 47
$\bar{Y}$, $\mathring{Y}$, ∂Y, 13, 14
$|\alpha|$, $\alpha!$, $\binom{\alpha}{\beta}$, 89, 93

Δ, 89
∂^α, ∂_k^α, $\frac{\partial}{\partial n}$ 89, 93, 157
o, 5
$\mathbb{R}^{\mathbb{N}}$, 93
$\hookrightarrow$, 119
$\hookrightarrow\hookrightarrow$, 121
$\Subset$, 21, 94
$\leq$, 9
$\ll$, $\perp$ 201
$\langle\cdot,\cdot\rangle$, 18
$\perp$, 38
$(\cdot,\cdot)$, 37
$(\cdot,\cdot)_k$, $|\cdot|_k$, $\|\cdot\|_k$, 89, 90, 110
$\|\cdot\|$, 27, 31
$\|\cdot\|_A$, 77
$\|\langle\cdot,\cdot\rangle\|$, 38

Namen- und Sachverzeichnis

Abbildung
  abgeschlossene –, 78
  abschließbare –, 80–82
  adjungierte –, 83–89
  beschränkte –, 18, 30
  Fortsetzung einer linearen –, 10
  halbbeschränkte –, 182–184
  hermitesche –, 163
  kanonische –, 28
  kompakte –, 143–147, 172–174
  konjugierte –, 50–51, 83
  kontrahierende –, 54
  lineare –, 6
  maximal symmetrische –, 179
  nichtnegative –, 164
  normale –, 163
  offene –, 15
  positive –, 164
  Räume von –en, 18–20
  Satz von der inversen –, 63
  Satz von der offenen –, 62
  selbstadjungierte –, 87
  Spektrum einer kompakten –, 160–163
  Spektrum einer kompakten normalen –, 163–172
  stetige –, 15, 20
  stetige lineare –, 31–33
  symmetrische –, 87
  transponierte –, 9
  wesentlich selbstadjungierte –, 87
abgeschlossen, 13
  –e Operatoren, 77–92
  Satz vom –en Graphen, 79–80
Ableitung
  – im Distributionensinn, 109
  schwache –, 82, 111
  starke –, 82, 112
abschließbar, 78, 80
Abschluß, 14, 80
absolut stetiger Teilraum, 202
absolut stetiges Maß, 201
Abstand, 21
abzählbare Basis, 29
adjungiert, 84
adjungierter Raum, 47
Alaoglu
  Satz von –, 75
Algebra, 33
  Banach–, 33
  kommutative –, 33
  normierte –, 33
algebraisch
  – konjugierter Raum, 7
  – reflexiv, 8
  –e Strukturen, 5–13
Annihilator, 139, 153
Approximationssatz, 39, 73
Approximieren, 112
äquivalente Normen, 32
Arzelà, Cesare, 24
Arzelà-Ascoli
  Satz von –, 24–26
Ascoli, Giulio, 24
Auswahlaxiom, 10
Auswahlsatz
  Rellichscher –, 121
Außenraumaufgabe, 157

Baire, René Louis, 59
Bairescher Kategoriensatz, 59
Banach
  –algebra, 33
  –raum, 28
Banach, Stefan, 28, 48, 61, 74
Banach-Steinhaus
  Satz von, 61–62
Basis
  – einer Topologie, 14
  abzählbare –, 29
  Hamel-, 6
Berührungspunkt, 15
beschränkt
  – linear, 31
  –e Abbildung, 18
  –e Menge, 20
Besselsche Ungleichung, 41
Bidualraum, 8
Borelmaß
  beschränktes, 191
  komplexes, 191
  positives –, 191
Borelmenge, 191

Calderón, Alberto Pedro, 123
Calderónscher Fortsetzungssatz, 123
Cauchy-Schwarzsche Ungleichung, 38
Cayley, Arthur, 175
Cayleytransformation, 175–182
Cayleytransformierte, 177
$\mathcal{C}_0$-Gruppe, 218
$\mathcal{C}_0$-Halbgruppe, 208
Courant, Richard, 169
Courantsches Maximum-Minimum Prinzip, 169

Darstellungssatz
  – für positive Funktionale aus $\mathcal{C}^0(\mathbb{R})'$, 193, 219
Defektindizes, 179
Definitionsbereich, 6
Definitionskern, 80
Δ-Operator, 89–92, 181
differenzierbar

schwach –, 111
stark –, 112
Dipolbelegung, 157
Dirac, Paul, 101
Dirac-Folgen, 101–102
Dirac-Funktion, 108
Dirichletsche Randbedingung, 92
Dirichletsche Randwertaufgabe, 1, 128, 157, 182
Distribution, 106–110
– endlicher Ordnung, 108
– mit kompaktem Träger, 110
Dirac–, 108
Heaviside–, 108
Ordnung einer –, 108
reguläre –, 107
temperierte –, 105, 110
Träger einer –, 109
von f erzeugte –, 107
$\mathcal{D}$-Nullfolge, 106
$\mathcal{D}$-stetig, 107
Dualraum, 47–53
Durchmesser, 20

Eigen
–funktion, 131
–vektor, 131
–wert, 131
Eigenfunktion
verallgemeinerte –, 203
Eigenwertaufgabe, 170
Einbettungen
kompakte –, 121–123
stetige –, 118–121
Einbettungssatz
Sobolevscher –, 114, 118
endlichdimensionale Räume, 35–37
Endomorphismus, 6
Energie, 172
kinetische –, 172
Lösung mit endlicher –, 205
potentielle –, 172
Prinzip vom lokalen Abklingen der –, 205
ε-Gitter, 23
Ergodensatz, 126
statistischer –, 126
Erzeuger
– einer Gruppe, 218
– einer Halbgruppe, 208
ess sup, 30
Eulersche Differentialgleichung, 129

Faltungen, 98–99
Fast
–metrik, 17
–norm, 27
fastmetrischer Raum, 17
Fixpunkt
–eigenschaft, 54
–gleichung, 54
Fixpunktsatz, 54
Banachscher –, 54
Brouwerscher –, 54–56
Schauderscher –, 56–58
Flächenbelegung
doppelte –, 157
einfache –, 157
Folge
konvergente –, 75
schwach beschränkte –, 60
folgenkompakt, 23
Fortsetzung, 77
Friedrichssche –, 182
kleinste selbstadjungierte –, 184
selbstadjungierte –, 175–184
stetige –, 32
triviale –, 175
Fortsetzungssatz, 10
Fouriertransformation, 103–106, 205
Fredholm, Ivar, 7, 147
Fredholm
–menge, 161
–operator, 160
–sche Alternative, 155
Fredholmsche Alternative, 7
Friedrichs, Kurt Otto, 100, 113, 182
Friedrichsscher Mollifier, 100–101
Fréchet, René Maurice, 28
Fréchetraum, 28
Funktional
beschränkt lineares –, 47–53
lineares –, 7
positives –, 193, 219
sublineares –, 12
Funktionenräume
klassische –, 93–94
Lebesguesche –, 95–96
spezielle –, 93–130

geordnet
halb –, 9
vollständig –, 9
glatter Rand, 157
gleichartig stetig, 25
gleichmäßig stetige
– Gruppe, 218
– Halbgruppe, 208
Graph, 77
–ennorm, 78
Gruppe, 5
$\mathcal{C}_0$–, 218
gleichmäßig stetige –, 218
stark stetige –, 218

Hadamard, Jacques, 1
Hahn, Hans, 48
Hahn-Banach
Satz von –, 48
Halb

–metrik, 17
–norm, 27
–ordnung, 9
halbbeschränkte Abbildung, 182–184
Halbgruppe, 126, 207–218
$\mathcal{C}_0$–, 208
gleichmäßig stetige –, 208
stark stetige –, 208
halbmetrischer Raum, 17
halbstetig, 76
Hamel, George, 6
Hamel-Basis, 6
Hauptvektor, 152
Hausdorff, Felix, 14, 17, 59, 124
Hausdorffsches Momentenproblem, 124
– vom positiven Typ, 124
Heaviside, Oliver, 108
Heaviside-Distribution, 108
hermitesche Abbildung, 163
Hilbert-Schmidt Operator, 173
Hilbertkubus, 58
Hilbertraum, 37–47
Hilbertraummethode, 2
Hille, Einar, 214
Hille-Yosida
Satz von –, 214
Hindernisaufgabe, 130
Hölder, Otto, 94
Höldersche Ungleichung, 30, 95
hölderstetig, 94
Homomorphismus, 6
homöomorph, 15
Homöomorphismus
linearer –, 33–35

Index eines Fredholmoperators, 160
Induktion
transfinite –, 9, 10
induktiver Limes, 106
infinitesimaler Erzeuger
– einer Gruppe, 218
– einer Halbgruppe, 208
Innenraumaufgabe, 157
innerer Punkt, 13
Integralgleichung
Volterrasche –, 34
Integralgleichungsmethode, 2, 157–160
isolierter Punkt, 14
Isometrie, 17
isometrisch, 17
– isomorph, 29
isomorph, 7
isometrisch –, 29
topologisch –, 29

kanonische Abbildung, 8, 28
Kategorie
erste, 59
zweite, 59
Kegeleigenschaft, 98, 121
strikte –, 123
klassische Lösung, 172
koerzitiv, 79, 92
kompakt, 20, 21
folgen–, 23
prä–, 23
relativ –, 23
kompakte Abbildung, 143–147
Spektrum –, 160–163
kompakte Mengen, 20–24
kongruent, 29
konjugiert
–e Abbildung, 50
algebraisch –er Raum, 7
topologisch –er Raum, 47
kontinuierliches Spektrum, 131
konvex, 39, 55, 71
–e Mengen, 71–74
–e Hülle, 71
gleichmäßig –, 73
konvexe Hülle, 56
Körper, 5

Landau, Edmund, 214
Landausche Ungleichung, 214
Laplace-Operator, 89–92
Leibniz, Gottfried Wilhelm, 117
Leibnizsche Regel, 117
Leray, Jean, 58
Leray-Schaudersche Randbedingung, 58
linear homöomorph, 29
Lipschitz, Rudolf, 94
Lipschitz-Eigenschaft, 121
lipschitzstetig, 94
Lösung
– endlicher Energie, 172, 205
– im Distributionensinn, 172
klassische –, 172
schwache –, 172
starke –, 172

maximal symmetrische Abbildung, 179
maximales Element, 9
Maximum-Minimum Prinzip
Courantsches –, 169
Mazur, Stanislaw, 74
Maß
absolut stetiges –, 201
singulär stetiges –, 201
Metrik, 15
Fast –, 17
Halb –, 17
Supremums–, 18
metrischer Raum, 15–18
Minimalprobleme der Variationsrechnung, 128–130
Minkowskische Ungleichung, 30
Möbiustransformation, 175

Multiindizes, 93

Neumann, Carl, 33
Neumann, John von, 175, 185
Neumannsche Randbedingung, 92
Neumannsche Randwertaufgabe, 157, 182
Neumannsche Reihe, 33
nichtnegative Abbildung, 164
Norm, 27
  –konvergenz, 64
  äquivalente –en, 32
  Fast–, 27
  Halb–, 27
  induzierte –, 38
  k–, 90
  Operator–, 31
  Quasi–, 27
normale Abbildung, 163
normale kompakte Abbildung
  Spektrum –, 163–172
normierte Räume, 27–31
normierter Vektorraum, 27
nuklearer Operator, 173
Nullraum, 6

offen, 13
  –e Abbildung, 15
offener Kern, 14
Operator
  Hilbert-Schmidt –, 173
  nuklearer –, 173
  Spurklasse –, 173
Operatornorm, 31
orthogonal, 38
  –es Komplement, 38
Orthonormalsystem, 40–43
  vollständiges –, 42
Osgood, William, 59

Parallelogrammgleichung, 38
Parsevalsche Gleichung, 43
Partition der Eins, 109, 115
Poincaré, Henri, 98
Poincarésche Abschätzung
  erste –, 98
  zweite –, 98
Polarisationsformel, 38
Polarisieren, 38
positive Abbildung, 164
positives Funktional, 193, 219
Potential
  mit einer Dipolbelegung, 157
  mit einer doppelten Belegung, 157
  mit einer einfachen Belegung, 157
Potentialtheorie, 157–160, 181–182
Prähilbertraum, 37
präkompakt, 23
Prinzip vom lokalen Abklingen der Energie, 205
Prinzip von der gleichmäßigen Beschränktheit, 60–61
Projektionssatz, 40
Projektor, 45–47
  orthogonaler –, 46
Punktspektrum, 131

Quasi
  –norm, 27
Quotientenraum, 28

Randbedingung
  Dirichletsche –, 92
  Neumannsche –, 92
Randpunkt, 15
Randwertaufgabe
  Dirichletsche –, 128, 157, 182
  dritte –, 182
  Neumannsche –, 157, 182
Raum
  normierter –, 27–31
  normierter Vektor–, 27
  reflexiver, 51
  topologisch linearer, 27
  topologischer –, 13
Raum der beschränkten Abbildungen, 30
reflexiv
  algebraisch –, 8
  topologisch –, 51
Regularitätsaussagen, 106, 113
Relation, 9
relativ kompakt, 23
Rellich, Franz, 121
Rellicher Auswahlsatz, 121
Residuenspektrum, 131
Resolvente, 131–143
  –nformel, 134
  –nmenge, 131
Retrakt, 55
Retraktion, 55
Riesz, Frigyes, 52, 147
Riesz
  –abbildung, 52
  –scher Darstellungssatz, 52
  –sches Lemma, 36
  –zahl, 149
Riesz-Schauder Theorie, 147–156

Saks, Stanislaw, 74
Satz
  – vom abgeschlossenen Graphen, 79–80
  – vom abgeschlossenen Wertebereich, 142
  – von Alaoglu, 75
  – von Arzelà-Ascoli, 24–26
  – von Baire-Hausdorff, 59–63
  – von Banach-Steinhaus, 61
  – von Hahn-Banach, 48
  – von Hille-Yosida, 214
  – von Pythagoras, 38

– von Stone, 218
– von der inversen Abbildung, 63
– von der offenen Abbildung, 62
– von der Neumannschen Reihe, 33
Schauder, Juliusz Pawel, 29, 56, 58, 147
Schauder-Basis, 6, 29
Schmidt, Erhard, 40
Schmidtsches Orthogonalisierungsverfahren, 40
schnell fallend, 103
Schrödingeroperator, 175, 184
schwach
– differenzierbar, 111
– konvergent, 64–76
–e Ableitung, 82, 111
–e Cauchyfolge, 66
–e Lösung, 172, 207
–e Topologie, 74–76
– folgenabgeschlossen, 72
– folgenkompakt, 66
– folgenvollständig, 66
– kompakt, 66–68
– konvergent, 64
schwach-stern
– dicht, 83
– folgenkompakt, 66
– konvergent, 64
Schwartz, Laurent, 103
Schwartzscher Raum, 103
Schwarzsche Ungleichung, 30
Segmenteigenschaft, 98
strikte –, 120
selbstadjungiert, 87
Seminorm, 27
separabel, 29, 41
separable Hilberträume, 43–44
separabler Raum, 41
Separation der Variablen, 170
Sesquilinearform, 37
σ-Algebra, 191
singulär stetiger Teilraum, 202
singulär stetiges Maß, 201
singulärer Teilraum, 202
singulärer Wert einer Abbildung, 172
Skalarprodukt, 37–39
Sobolev
–räume, 90, 110–123
–scher Einbettungssatz, 114, 118
Sobolev, Sergei Lvovich, 110
Spektraldarstellung, 125, 185, 186
– unitärer Abbildungen, 220
Spektralradius, 135
Spektralsatz, 2, 125
für kompakte Abbildungen, 163
für kompakte normale Abbildungen, 165
für selbstadjungierte Abbildungen, 196
Spektralschar, 126, 185–191
beschränkte –, 185
konstante, 185
Träger einer –, 185
Spektrum, 131–143
– kompakter Abbildungen, 160–163
– kompakter normaler Abbildungen, 163–172
kontinuierliches –, 131
Punkt–, 131
Residuen–, 131
wesentliches –, 161
Spitzeneigenschaft
strikte p–, 123
Sprungrelation, 157
Spur, 120
–satz, 120
Spurklasse-Operator, 173
stark
– differenzierbar, 112
–e Ableitung, 82, 112
–e Lösung, 172
– konvergent, 64
stark gleich schwach, 115–118
stark stetige
– Gruppe, 218
– Halbgruppe, 208
stehende Welle, 170
Steinhaus, Hugo, 61
stetig
–e Abbildung, 15, 20
absolut –es Maß, 201
gleichartig –, 25
singulär –es Maß, 201
stetiger Teilraum, 200
Stieltjes, Thomas, 192
Stieltjessche Umkehrformel, 192
Stone, Marshall, 185
Stone
–sche Formel, 196
Satz von –, 218
Streifenbedingung, 99
Streutheorie, 205
Subbasis, 14
Supremumsmetrik, 18
symmetrisch, 87

Teilraum
absolut stetiger –, 202
singulär stetiger –, 202
singulärer –, 202
stetiger –, 200
unstetiger –, 200
testen, 12, 112
Testfunktion, 94
Thermoelastizitätsgleichungen, 216
Topologie, 13
diskrete –, 14
euklidische –, 14
induzierte –, 15
relative –, 15
schwache –, 74–76
triviale –, 14

topologisch
– isomorph, 29
– konjugierter Raum, 47
–e Räume, 13–15
lineare Räume, 27
totalbeschränkt, 23
totale Teilmenge von $\mathcal{X}'$, 141
totale Teilmenge von X', 83
Totalvariation, 192
Träger, 94
transfinite Induktion, 9, 10
transponierte Abbildung, 9
Trennungsaxiom, 14
Trennungssatz, 11, 49, 71

Überdeckung, 21
endliche –, 21
lokal finite, 115
offene –, 21
Umgebung, 14
–sbasis, 14
–sfilter, 14
unitäre Abbildung
Spektraldarstellung –, 220
unstetiger Teilraum, 200

Variation, 31
beschränkte –, 30
Total–, 192
totale –, 31
Variationsaufgabe, 128
Variationsungleichung, 129
Vektorraum, 5
verallgemeinerte Eigenfunktion, 203
Vervollständigung, 17
vollstetig, 57
Volterra, Vito, 35
Volterrasche Integralgleichung, 34
Volterrascher Integraloperator, 132, 136

Wasserstoffatom, 175, 184
Wellengleichung, 205
Wertebereich, 6
wesentlich selbstadjungiert, 87
wesentliches Spektrum, 161
Weyl, Hermann, 113
Weylsches Lemma, 113
Wohlordnungssatz, 10

Yosida, Kôsaku, 214
Yosida-Approximierende, 215

Zermelo, Ernst, 9
Zorn, Max, 9
Zornsches Lemma, 9
Zustandsdiagramm, 139